

PRÉFACE

Alain Berthoz

La réalité virtuelle va jouer un rôle décisif dans les années à venir et ce traité est donc très utile pour rassembler un certain nombre de connaissances, de théories et de techniques. On peut, en effet, prévoir une explosion de son utilisation dans les domaines suivants.

D'abord l'éducation où elle va apporter beaucoup à de nombreuses disciplines par le développement de l'enseignement à distance et par l'apprentissage des processus industriels. Les plus grandes sociétés développent aujourd'hui des méthodes basées sur la réalité virtuelle pour entraîner leur personnel à des opérations de manipulation, l'aider dans la compréhension de processus complexes et dans la navigation en environnements hostiles ou lointains. De façon plus générale, la réalité virtuelle haptique et visuelle (car la réalité virtuelle n'est pas seulement visuelle mais multimodale) sera un outil indispensable pour le développement des simulateurs de demain, qu'ils concernent le pilotage d'avions ou de voitures, la simulation de méthodes chirurgicales ou la simulation d'activités sportives. On peut dire que nous tenons là un formidable outil d'apprentissage. Cependant, son usage ne sera pas limité à l'enseignement ou à l'entraînement. La réalité virtuelle est aussi entrée dans le monde du travail par la téléopération qui exige une restitution de la présence des objets manipulés et de leur environnement. Véritable prolongement de la main, de la vision, du sens du toucher, la réalité virtuelle permet aujourd'hui d'allonger le bras de l'opérateur au-delà des routes et des océans.

L'utilisation de la réalité virtuelle ne s'arrête toujours pas là. On entrevoit maintenant son emploi dans le domaine médical pour l'aide au diagnostic de certaines maladies (par exemple le vertige ou les troubles de l'orientation spatiale) et pour l'étude des bases neurales de ces maladies, où la réalité virtuelle s'est introduite car elle permet de manipuler les paramètres du monde visuel, haptique ou acoustique. La réalité virtuelle peut être aussi exploitée pour la réhabilitation, voire le traitement, des troubles liés à l'âge (par exemple par des méthodes de thérapies comportementales).

Enfin son usage va se répandre dans la science par le biais de la simulation des processus et de la construction de modèles explicatifs dans de très nombreux domaines. Il est inutile de mentionner l'immense champ d'application que constitue le domaine des jeux, qu'ils soient basés ou non sur des techniques de réalité virtuelle.

Ces exemples donnent une idée de l'enjeu. Face à lui, les méthodes et les techniques, comme on le lira ici, se développent très rapidement. Il faut toutefois rester conscient du chemin qui reste à parcourir. Et qui risque d'être long. Ce chemin est d'abord technologique. En effet, la technologie qui permettra une restitution fidèle de la réalité n'est pas encore achevée et le savoir-faire des ingénieurs sera précieux. Les jeunes ont devant eux un immense champ de recherche et de réalisation. Ensuite, sur le plan de la fusion des modalités de la réalité virtuelle, il est nécessaire de combiner vision, sens haptique et acoustique, comme le fait si habilement le cerveau. Ceci est encore un grand défi à relever. Pour cela, je conseille sans aucun doute d'accroître la trop faible coopération entre les spécialistes des sciences de la cognition et les ingénieurs. Pas parce que les processus qu'utilise le cerveau, lorsqu'ils sont connus, sont nécessairement à copier

mais parce qu'un échange réciproque ne peut qu'enrichir les connaissances des uns et des autres. De telles coopérations sont trop rares en France, et je remercie les auteurs de ce livre de s'être adressé à un Physiologiste pour préfacer leur ouvrage.

Le chemin est ensuite théorique. Nous savons peu de choses sur les processus cérébraux qui permettent de reconnaître avec une grande finesse la forme d'un objet en mouvement. Au fond, nous savons peu de chose sur la perception du mouvement et sur les mécanismes qui permettent au cerveau de mettre en correspondance le toucher et la vision. Nous savons enfin peu de chose sur la façon dont le cerveau résout les problèmes de délai. J'ai eu l'impression, en lisant ce livre, que des problèmes théoriques majeurs sont encore à résoudre dans le domaine de l'algorithmique, en particulier concernant les problèmes tridimensionnels. La représentation de l'espace 3D, la reconnaissance des objets, posent des problèmes de géométrie encore non résolus, les capteurs sont encore à perfectionner mais aussi les théories sur le traitement de leurs données. Nous savons que le cerveau peut résoudre le problème du grand nombre de degrés de liberté à contrôler, qu'il anticipe, présélectionne et fait des hypothèses sur les données des sens. Ces capacités de prédiction, d'anticipation, d'utilisation de la mémoire pour prédire le futur permettent de simplifier la "neurocomputation". Nous savons aussi que l'analyse du monde extérieur utilise des "modèles internes" des lois de la mécanique et des lois physiques. Faut-il intégrer de tels processus dans les systèmes de contrôle et d'analyse des dispositifs de réalité virtuelle ? Faut-il en inventer de meilleurs ? Voilà le défi et ce livre sera utile à tous ceux qui veulent le relever.

Alain Berthoz est Professeur au Collège de France, il est le directeur du Laboratoire de Physiologie de la Perception et de l'Action CNRS Collège de France. Il est notamment l'auteur de l'ouvrage «Le sens du mouvement» publié chez Odile Jacob en 1997.

AVANT PROPOS

Depuis une dizaine d'années, la réalité virtuelle s'est développée en France. Celle-ci ouvre potentiellement de nouvelles perspectives pour notre société. Mais soyons réalistes, la réalité virtuelle induit d'abord bien des défis scientifiques pour les chercheurs et pour les industriels. Nous avons participé avec enthousiasme à l'éclosion de la réalité virtuelle en France, conscients de l'immensité de la tâche. Durant cette décennie, nous avons développé, entre autres, des connaissances pour l'interfaçage en environnement virtuel et conçu des méthodes pour la réalité virtuelle, tant sur le plan technique que sur le plan humain.

Parallèlement à nos travaux de recherche, nous avons souhaité faire une tâche d'information sur les avancées du domaine. Dans les années 90, Jean-Paul Papin a présidé le groupe de recherche OTAN sur les applications de la réalité virtuelle dans le milieu militaire. Plusieurs ateliers internationaux ont été organisés et ont publié des rapports scientifiques. En 1996, Philippe Fuchs rédigea un premier livre sur la réalité virtuelle (Les Interfaces de la Réalité Virtuelle). Guillaume Moreau a exposé dans son mémoire de thèse ses recherches sur l'animation d'êtres humains dans les mondes virtuels.

Conscient de l'évolution des techniques, des nouvelles approches cognitives de la réalité virtuelle et de l'absence de livre récent en français, nous avons décidé de rédiger un nouveau livre plus ambitieux, couvrant les différents aspects du domaine. Son objectif est de fournir à toute personne un document de référence sur la Réalité Virtuelle, d'où son titre «Le traité de la réalité virtuelle».

Mais personne ne peut avoir une vision globale et précise de tous les secteurs de la réalité virtuelle. Nous souhaitons vivement que d'autres chercheurs participent à la rédaction du traité dans une prochaine version, soit pour écrire de nouveaux chapitres, soit pour compléter des chapitres existants. Notre ambition actuelle est de proposer une première version, sachant qu'une réactualisation périodique est possible grâce à la réactivité des Presses de l'Ecole des Mines de Paris et grâce aux facilités d'échange par Internet. Ainsi nous désirons profiter de l'essor des nouveaux moyens d'information : en proposant une version consultable sur notre site web (et téléchargeable gratuitement¹ pour les étudiants), nous espérons une plus grande interactivité avec le lecteur qui pourra donner en ligne son point de vue et ses critiques, que nous accepterons volontiers. Et en étant optimiste, dans un futur plus ou moins proche, quand le livre électronique sera performant, exploitable et exploité à la place du livre papier, l'interactivité pourra être plus grande entre auteurs et lecteurs. Ceci induira un nouveau moyen de communication : l'INTERLIVRE, puisqu'il sera INTERactif, téléchargeable sur INTERNet et rédigé INTERauteurs.

Nous tenons à remercier toutes les personnes qui nous ont aidés dans nos recherches et dans la réalisation de ce livre, en particulier le Professeur Claude Laurgeau, directeur du Centre de Robotique de l'Ecole des Mines de Paris, tous nos collègues, en particulier Fawzi Nashashibi et Domitile Lourdeaux, les entreprises nous ayant fourni des photographies ainsi que Madame Catherine Delamare, responsable des Presses de

¹La version électronique n'est gratuite que pour les étudiants. Nous espérons que cette règle sera respectée par tous.

l'Ecole des Mines de Paris. Nous exprimons notre gratitude aux différents relecteurs dont les conseils ont contribué à améliorer notre manuscrit : les Professeurs Jean Serra, Bruno Arnaldi, Gérard Subsol et Jacques Tisseau.

Nos remerciements s'adressent aussi au Professeur Alain Berthoz du Collège de France, qui a eu la gentillesse de préfacer notre livre.

Nous espérons, cher lecteur, que vous apprécierez de lire ce traité, comme nous avons pris plaisir à le rédiger et à transmettre nos idées.

Les Auteurs,

TABLE DES MATIÈRES

I	Les concepts de base de la réalité virtuelle	1
1	La réalité virtuelle et ses applications	3
1.1	Fondement de la réalité virtuelle	3
1.1.1	Introduction	3
1.1.2	Définitions de la réalité virtuelle	4
1.1.3	Aspects philosophique et social de la réalité virtuelle	11
1.1.4	Définition et taxonomie de la réalité augmentée	12
1.1.5	Historique succinct de la réalité virtuelle	17
1.1.6	Les domaines précurseurs de la réalité virtuelle	18
1.1.7	Le développement de la réalité virtuelle en France	18
1.1.8	Un petit bêtisier de la réalité virtuelle	21
1.2	Les applications de la réalité virtuelle	22
1.2.1	Taxonomie théorique des applications RV	22
1.2.2	Les applications potentielles de la réalité virtuelle	23
1.2.3	Les domaines économiques des applications actuelles	25
1.3	Références bibliographiques	35
2	Approche théorique et pragmatique de la réalité virtuelle	39
2.1	Comportement de l'homme dans un environnement réel	39
2.2	Conception des interfaces comportementales	41
2.2.1	Conception matérielle	41
2.2.2	Transparence d'une interface	42

2.2.3	Interfaces commercialisées et interfaces spécifiques	43
2.2.4	Comparaison entre interfaces en réalité virtuelle et en téléopération . . .	45
2.3	L'approche «instrumentale» pour l'immersion et l'interaction	45
2.3.1	Les concepts fondamentaux pour l'interfaçage comportemental	45
2.3.2	Interfaces comportementales, schèmes et métaphores	49
2.4	Méthode de conception et d'évaluation d'un environnement RV	53
2.4.1	Le schéma de référence en RV	53
2.4.2	Les Primitives Comportementales Virtuelles et les Aides Logicielles Comportementales	56
2.4.3	Démarche de conception	57
2.4.4	Démarche de l'évaluation	59
2.4.5	Ouverture d'une porte	61
2.4.6	Magasin virtuel d'expérimentation commerciale	61
2.4.7	Formation en réalité virtuelle sur des infrastructures ferroviaires	67
2.5	Discussion sur l'immersion et l'interaction du sujet	70
2.5.1	Approche instrumentale	70
2.5.2	Présence et immersion	72
2.6	Perspectives et conclusions de notre approche	73
2.7	Références bibliographiques	74

II Les interfaces comportementales 77

3 L'être humain dans les environnements réel et virtuel 79

3.1	Introduction	79
3.2	L'être humain, son environnement naturel	79
3.3	Les facteurs humains et le facteur humain	83

3.3.1	Définitions	83
3.3.2	Démarche systémique	84
3.3.3	L'homme en action	84
3.3.4	Facteurs influents	89
3.4	Conclusion	93
3.5	Références bibliographiques	93
4	Les sens et les réponses motrices de l'homme	95
4.1	Le fonctionnement des sens	95
4.2	La vision	97
4.2.1	Le système visuel humain	97
4.2.2	Traitement de l'information dans le système visuel	102
4.2.3	Prise visuelle d'informations	107
4.2.4	La perception visuelle de la profondeur	108
4.2.5	Les caractéristiques psychophysiques de la vision	115
4.3	L'ouïe	122
4.3.1	Introduction	122
4.3.2	Physique du son	122
4.3.3	Le système auditif	125
4.3.4	Physiologie de l'audition	126
4.3.5	Principe de localisation spatiale d'une source sonore	128
4.3.6	La perception auditive	130
4.4	Sensibilité cutanée	131
4.4.1	Physique de la matière	132
4.4.2	La peau	133
4.5	Les ambiances chimiques : l'odorat et le goût	136

4.5.1	Les odeurs	136
4.5.2	Les saveurs	139
4.6	Proprioception	140
4.6.1	Introduction	140
4.6.2	Physique de la gravité et des accélérations	141
4.6.3	L'appareil vestibulaire et les voies kinesthésiques	141
4.7	Réponses motrices	144
4.8	La main, organe d'information et d'action	145
4.8.1	Anatomie de la main	145
4.8.2	Fonction de la main	147
4.9	Classification des interfaces	149
4.10	Références bibliographiques	149

III Les interfaces motrices 153

5	Les capteurs de localisation	155
5.1	Généralités	155
5.1.1	La localisation spatiale	155
5.1.2	Capteur de localisation et interface de commande	156
5.2	Principes de positionnement	157
5.2.1	Principe par mesures de distance entre points	157
5.2.2	Principe de mesure avec source émettrice	158
5.2.3	Principe de mesure sans source émettrice artificielle	158
5.2.4	Principe de mesure de caractéristiques du mouvement	158
5.2.5	Principe par mesures d'angles entre solides	159
5.3	Traqueurs mécaniques	159

5.3.1	Traqueurs mécaniques mesurant des distances	159
5.3.2	Capteurs mécaniques déterminant une orientation, une vitesse ou une accélération	162
5.4	Traqueurs électromagnétiques	166
5.4.1	Traqueurs électromagnétiques à champ alternatif	166
5.4.2	Traqueurs électromagnétiques à champ impulsionnel	168
5.4.3	Caractéristiques des traqueurs électromagnétiques	170
5.4.4	Compas	173
5.5	Traqueurs acoustiques	174
5.6	Traqueurs optiques	177
5.6.1	Traqueurs avec récepteurs ponctuels	177
5.6.2	Traqueurs avec récepteurs plans	178
5.7	Comparaison entre les différents types de traqueurs	183
5.8	Annexe : Calcul des paramètres du capteur électromagnétique Fastrak . .	185
5.8.1	Principe	185
5.8.2	Détermination des courants dans le récepteur	187
5.8.3	Détermination des changements en position et orientation	189
5.9	Références bibliographiques	191
6	Les interfaces spécifiques de la localisation corporelle	193
6.1	Introduction	193
6.2	Interfaces de localisation du corps «motion capture»	194
6.2.1	Deux principes de «motion capture»	194
6.2.2	Interfaces portables de localisation du corps	195
6.2.3	Interfaces de localisation optique	197
6.3	Interfaces de localisation des membres supérieurs ou inférieurs	201

6.4	Interfaces spécifiques de locomotion	205
6.4.1	Interfaces de locomotion par patins	207
6.4.2	Interfaces de locomotion par tapis	208
6.4.3	Interfaces de locomotion en position assise	211
6.5	Interfaces de detection de L'expression du visage ou des lèvres	212
6.5.1	Détection optique de l'expression du visage ou des lèvres	212
6.5.2	Détection mecanique de l'expression du visage ou des lèvres	213
6.5.3	Conclusion	215
6.6	Oculomètres	215
6.6.1	Introduction	215
6.6.2	Oculomètres optiques	215
6.6.3	Oculomètres électromagnétiques	217
6.7	Capteurs physiologiques	217
6.8	Conclusion	217
6.9	Références bibliographiques	218
7	Les interfaces manuelles motrices	221
7.1	Introduction	221
7.1.1	Capteur de localisation et gant de données	221
7.1.2	Capteur de localisation et interface de commande	221
7.2	Gants de données	222
7.2.1	Gants à fibres optiques	222
7.2.2	Gants à variation de résistances	224
7.2.3	Gants à effet Hall	226
7.2.4	Cas particulier : gant de commandes binaires	226
7.2.5	Conclusion	228

7.3	Interfaces de commande	228
7.3.1	Souris 3D	228
7.3.2	Interfaces non-manuelles de commande	230
7.4	Références bibliographiques	231

IV Les interfaces sensori-motrices 233

8 Les interfaces manuelles sensorimotrices, interfaces à retour d'effort 235

8.1	Introduction	235
8.2	Etude de la création des retours d'effort	237
8.2.1	Réalisation physique des efforts	237
8.2.2	Architecture mécanique de l'interface	237
8.3	Les différentes interfaces à retour d'effort	240
8.3.1	Retour d'effort à réaction externe	240
8.3.2	Retour d'effort à réaction interne	260
8.4	Conclusion	265
8.5	Références bibliographiques	266

V Les interfaces sensorielles 269

9 Les interfaces manuelles sensorielles, les interfaces à sensibilité cutanée 271

9.1	Introduction	271
9.2	Interface à retour thermique	272
9.3	Interfaces à retour tactile	274
9.3.1	Retour tactile pneumatique	274
9.3.2	Retour tactile à matrice d'aiguilles	274

9.3.3	Retour tactile par vibrations	276
9.3.4	Retour tactile par manette	278
9.3.5	Retour tactile par électrodes cutanées	278
9.3.6	Le cas particulier du mixage de phénomènes virtuel et réel	279
9.4	Conclusion	280
9.5	Références bibliographiques	280
10	Les interfaces visuelles	283
10.1	Introduction aux interfaces visuelles	283
10.2	Interfaces visuelles à support fixe	284
10.2.1	Ecrans d'ordinateur monoscopiques	284
10.2.2	Ecrans d'ordinateur stéréoscopiques	285
10.2.3	Fenêtre immersive	291
10.2.4	Visiobureaux ou bureaux immersifs	291
10.2.5	Interfaces visuelles à taille humaine : visiosalle (salle immersive) et visiocube	298
10.2.6	Différents types de vidéoprojecteurs	305
10.3	Interfaces visuelles portables	306
10.3.1	Architecture d'un visiocasque	307
10.3.2	Visiocasques avec écrans à tube cathodique	310
10.3.3	Visiocasques avec écrans à cristaux liquides	312
10.3.4	Modèle optique d'un visiocasque et problèmes connexes	314
10.3.5	Vidéolunettes	317
10.3.6	Visiocasque et dispositif semi transparents	320
10.4	Critères de choix d'une interface visuelle	321
10.5	Tests visuels sur les interfaces	323

10.5.1 Tests visuels sur les écrans stéréoscopiques	323
10.5.2 Tests visuels sur les écrans visiocasques	324
10.6 Création d'images stéréoscopiques	326
10.6.1 Principe	327
10.6.2 Choix des paramètres stéréoscopiques	333
10.6.3 Création d'images en relief pour la téléopération	334
10.6.4 Limitation de la fatigue oculaire en vision stéréoscopique	338
10.6.5 Création d'images en relief pour une revue de projet	341
10.7 Évaluation des techniques stéréoscopiques	342
10.7.1 Intérêts de la vision en relief	342
10.7.2 Choix des paramètres de la vision en relief	342
10.7.3 Comparaison des solutions à un écran	343
10.7.4 Comparaison entre les écrans des visiocasques	344
10.8 Conclusion	344
10.9 Annexes	345
10.9.1 Restitution par images volumiques	345
10.9.2 Dispositifs pour la polarisation de lumière	345
10.9.3 Modélisation d'un visiocasque	347
10.9.4 Perception du relief sur une feuille	349
10.9.5 Formules de la vision stéréoscopique	350
10.10 Références bibliographiques	359
 11 Les interfaces à simulation de mouvement et les interfaces à simulation de climat	 363
11.1 Généralités	363
11.2 Sièges de simulation pour mono-utilisateur	366

11.2.1 Généralités	366
11.2.2 Les différents produits de simulation mono-utilisateur	367
11.3 Cabines de simulation pour multi utilisateurs	371
11.3.1 Evolution des simulateurs	371
11.3.2 Plates-formes mobiles	373
11.3.3 Les différentes cabines de simulation	376
11.4 Conclusion sur les interfaces à simulation de mouvement	378
11.5 Interface à simulation de climat	378
11.6 Références bibliographiques	379
 12 Les interfaces olfactives	 381
12.1 Introduction	381
12.2 Les interfaces olfactives commercialisées	381
12.2.1 L'interface olfactive de Olfacom	381
12.2.2 L'interface olfactive de Ruetz	382
12.2.3 L'interface olfactive de AromaJet	383
12.2.4 L'interface olfactive de Aerome	384
12.2.5 L'interface olfactive de Digiscents	384
12.3 Conclusion	385
12.4 Références bibliographiques	385
 13 Les primitives comportementales virtuelles	 387
13.1 Rappel sur notre démarche en réalité virtuelle	387
13.1.1 Concepts de base	387
13.1.2 Taxonomie des PCV	389
13.2 Primitives Comportementales Virtuelles d'observation	390

13.2.1	Classification	390
13.2.2	Observation visuelle	391
13.2.3	Observation acoustique	392
13.2.4	Observation tactile	392
13.3	Primitives Comportementales Virtuelles de déplacement	392
13.3.1	Problématique d'un déplacement dans un monde virtuel	392
13.3.2	Les différentes interfaces comportementales exploitables	394
13.4	Primitives Comportementales Virtuelles pour agir sur le monde virtuel	395
13.4.1	Introduction	395
13.4.2	Manipuler et orienter un objet	396
13.4.3	Assembler des objets	398
13.4.4	Déformer un objet	398
13.4.5	Agir sur un être vivant	398
13.5	Primitives Comportementales Virtuelles de communication avec autrui	399
13.6	Conclusion	400
13.7	Références bibliographiques	400

VI Modélisation des mondes virtuels 401

14 Modèles géométriques des environnements virtuels 403

14.1	Introduction	403
14.1.1	Types d'objets	404
14.1.2	Propriétés des modèles	404
14.2	Modèles volumiques	406
14.2.1	Énumération spatiale	406

14.2.2	Constructive Solid Geometry	409
14.3	Modèles surfaciques	411
14.3.1	Utilisation de surfaces planes	411
14.3.2	Utilisation de surfaces non planes	412
14.4	Géométrie algorithmique	416
14.4.1	Transformation d'un volume en surface	416
14.4.2	Maillage polygonal d'un nuage de points	417
14.4.3	Décimation de maillages	421
14.5	Optimisation des modèles pour la réalité virtuelle	424
14.5.1	Texturation	425
14.5.2	Niveaux de détails	426
14.6	Références bibliographiques	428
15	Comportements des objets du monde virtuel	431
15.1	Introduction	431
15.2	Modèles descriptifs	432
15.2.1	Modèle	432
15.2.2	Animation par keyframing	432
15.2.3	Cinématique directe et inverse	434
15.3	Modèles générateurs	435
15.3.1	Modèle	435
15.3.2	Rappels sur la mécanique des solides	436
15.3.3	Mécanismes rigides articulés	437
15.3.4	Simulation de systèmes déformables	438
15.3.5	Discussion sur le rôle des modèles générateurs	441
15.4	Modèles comportementaux	442

15.4.1	Modèles de transformation internes	443
15.4.2	Modèles de transformation externes	446
15.5	Approche stimulus-réponse	446
15.5.1	Les réseaux SAN	446
15.5.2	Contrôle Stimulus-Réponse de créatures	447
15.5.3	Sense Control Action	448
15.5.4	Résumé	449
15.6	Approche à base de règles	449
15.6.1	Nuées, troupes et bancs	449
15.6.2	Approche «écologique»	450
15.6.3	In VitroAm	450
15.6.4	Discussion	451
15.7	Approche optimale	452
15.8	Approche automate	452
15.8.1	Automate d'états finis	452
15.8.2	Piles d'automates	453
15.8.3	Concurrence et hiérarchie : HPTS	455
15.9	Autonomie sous contrôle : scénarios d'animation	460
15.10	Cas particulier : le motion capture	461
15.11	Références bibliographiques	463
VII	Aspects généraux de la réalité virtuelle	469
16	Les effets sur l'homme des environnements	471
16.1	Les effets sur l'homme des environnements naturels	471
16.1.1	La lumière	471

16.1.2	Le bruit	471
16.1.3	La chaleur	473
16.1.4	Le froid	476
16.1.5	Les accélérations	477
16.1.6	Les vibrations des solides et leurs effets généraux sur l'organisme humain	478
16.2	Les effets sur l'homme des environnements virtuels	481
16.2.1	Les effets visuels	481
16.2.2	Les effets musculo-squelettiques	482
16.2.3	Cinétoses	483
16.2.4	Effets psychologiques	484
16.2.5	Facteurs influents	486
16.2.6	Outils d'analyse des environnements virtuels et des effets de l'immersion	487
16.3	Références bibliographiques	489
17	Ethique, droit et réalité virtuelle	491
17.1	Introduction	491
17.1.1	Loi sur la protection des personnes se prêtant à des recherches biomédicales	491
17.1.2	Recommandations aux promoteurs et aux chercheurs	492
17.1.3	Autorisation des lieux de recherches biomédicales sans bénéfice individuel direct	493
17.2	Ethique et droit	494
17.2.1	Perspectives juridiques et éthiques selon Huriet	494
17.2.2	Ethique et droit en télé médecine	494
17.3	Références bibliographiques	495
	Vocabulaire	497

Table des matières	xix

Table des illustrations	503

Première partie

Les concepts de base de la réalité virtuelle

1 LA RÉALITÉ VIRTUELLE ET SES APPLICATIONS

Philippe Fuchs

1.1 FONDEMENT DE LA RÉALITÉ VIRTUELLE

1.1.1 INTRODUCTION

Depuis toujours l'homme aspire à s'échapper de la réalité quotidienne pour des raisons personnelles, professionnelles, artistiques ou culturelles. L'évolution des techniques l'aidant, l'homme a pu satisfaire ce besoin par des représentations principalement visuelles ou sonores, mais *figées* du monde. *Figées* dans le sens où l'utilisateur ne peut observer la représentation qu'en *spectateur*, que ce soit une peinture, une photographie, un film d'images réelles ou d'images de synthèse. La réalité virtuelle lui offre une dimension supplémentaire en lui procurant un environnement virtuel dans lequel il devient *acteur*. Que le lecteur ne s'y méprenne pas, la nouveauté n'est pas dans la création d'environnements virtuels plus performants dans leur représentation¹, mais bien dans la possibilité de pouvoir «agir virtuellement» dans un monde artificiel (ou «interagir», vu plus sous l'angle technique).

La réalité virtuelle ne peut être envisagée que depuis peu, grâce à l'augmentation importante de la puissance intrinsèque des ordinateurs, en particulier la possibilité de créer en temps réel des images de synthèse et de permettre une interactivité, toujours en temps réel, entre l'utilisateur et le monde virtuel. Le lecteur doit bien noter qu'à la base de la réalité virtuelle, ce sont des évolutions techniques qui ont permis son essor, et donc, par déduction, qui en contraignent sa portée, due aux limites inhérentes de la technique. N'écoutons pas les dires de certains, qui ont proclamé un peu vite, que le «Grand Jour du Virtuel» était arrivé ! Essayons simplement de relever ce challenge pour franchir cette nouvelle étape («agir dans le virtuel») en sachant que la route sera longue et que les évolutions seront lentes. Que de désillusions pour certains qui proclamaient, il y a dix ans, une révolution grâce à la réalité virtuelle. C'était l'époque où la réalité virtuelle faisait la Une des journaux... Mais Internet l'a remplacée !

Et la réalité virtuelle n'est pas née spontanément il y a quinze ans. Comme toute nouvelle technique, elle a eu des antécédents qui ne s'appelaient pas «réalité virtuelle». Dans quel type d'application savait-on permettre à une personne d'agir dans un environnement virtuel ? Principalement dans les simulateurs de transport², qui ont permis à des professionnels d'interagir avec un environnement partiellement³ virtuel, depuis 50 ans environ (On faisait de la réalité virtuelle sans le savoir, comme M. Jourdain, qui comme chacun sait, faisait de la etc.).

Mais si les ordinateurs permettent de simuler des mondes virtuels, l'interaction de l'homme avec ceux-ci n'est possible qu'au travers de logiciels et d'interfaces maté-

¹ A ce propos, il est étonnant de voir certains faire ressortir l'exploitation des images en relief, comme la nouveauté liée à la réalité virtuelle. Rappelons que l'on sait créer des images en relief (stéréoscopiques) depuis 150 ans environ.

² On précisera plus loin les liens entre les simulateurs de transport et la réalité virtuelle.

³ Le poste de conduite de la voiture, de l'avion ou du train n'est pas simulé mais réel.

rielles qui soient adaptés à cet objectif. Si la réalité virtuelle s'appuie sur l'informatique pour son essor, elle doit être à plus ou moins long terme un secteur de recherche et d'activité indépendant. Elle ne doit pas être considérée comme une simple branche de l'informatique, vu les potentialités novatrices qu'elle propose.

1.1.2 DÉFINITIONS DE LA RÉALITÉ VIRTUELLE

1.1.2.1 Préliminaire et image réductrice de la réalité virtuelle

Le terme «réalité virtuelle» est employé depuis dix ans, mais le domaine qu'il recouvre n'a jamais été bien défini, à notre avis. Ce terme est discutable et a été discuté par certains. Cet oxymoron vient de l'expression anglaise «virtual reality», introduite aux États-Unis dans les années 80 par Jaron Lanier. Cette appellation étant maintenant courante, il est vain de vouloir la modifier. Cependant, comme le rappelle J.P. Papin, en anglais, virtual signifie «de fait», «pratiquement». La traduction française ne rend donc pas compte de cette signification. Il aurait fallu parler de tenant lieu de réalité ou de réalité vicariante ou mieux encore d'environnement vicariant. Le mot vicariant est utilisé en physiologie pour désigner un organe qui se supplée à un autre. Ce mot vient de vicaire, le remplaçant du curé. Le vicaire prend les fonctions du curé en son absence mais il est loin d'en être toujours le sosie.

Certains ont employé à tort ce terme ou souvent lui ont donné une signification restrictive. On trouve dans la littérature des définitions qui mélangent la finalité de la réalité virtuelle, ses fonctions, ses applications et ses techniques sur lesquelles elle repose. Certains vont jusqu'à définir la réalité virtuelle par **l'image réductrice** qui a été véhiculée dans les médias : une personne équipée d'un visiocasque⁴ avec différentes commandes pour interagir (gant de données, manette, volant, etc.) et reliées à un ordinateur (Figure 1.1). Il ne faut donc pas prendre ces approches. On ne définit pas l'informatique par son matériel : clavier, écran, souris ; ni en parlant de ses fonctions : calculs, raisonnement logique, mémorisation ; ni par ses types d'applications.



Figure 1.1 : Simulation de la conduite d'une voiture avec visiocasque ou sur écran de simulateur, copyright Photo PSA Peugeot Citroën

⁴Visiocasque : traduction de «Head Mounted Display». On peut aussi employer l'expression «casque immersif».

1.1.2.2 Finalité de la réalité virtuelle

*La finalité de la réalité virtuelle est de permettre à une ou à plusieurs personnes des activités **sensori-motrices** et donc mentales dans un monde artificiel, qui est soit imaginaire, soit une simulation de certains aspects du monde réel.*

Dans le deuxième cas, *une simulation de certains aspects du monde réel*, ceux-ci sont à déterminer lors de la conception de l'application. Nous verrons par la suite que cette phase initiale de conception est fondamentale et doit être analysée explicitement. L'erreur, souvent rencontrée, est celle du concepteur qui recherche le plus grand «degré de réalisme» (en pratique limité par des contraintes techniques et économiques). Cette approche erronée est entreprise sans se soucier de savoir précisément quels sont les aspects de la réalité que l'application exige. L'idée naïve de vouloir, si possible, le comportement du monde virtuel entièrement identique à celui du monde réel est absurde. Si on veut faire une réalité «virtuelle», c'est bien dans le but de modifier des aspects de la «vraie» réalité. Par exemple :

- on souhaite former du personnel dans un environnement virtuel pour leur éviter des risques réels ;
- on souhaite étudier un futur produit, qui n'est donc pas encore matérialisé.

Il ne faut pas penser que c'est seulement à cause de difficultés techniques que la simulation virtuelle est différente de la réalité. Mais réciproquement, si une simulation ne représente que certains phénomènes physiques réels, la réalité virtuelle permet aussi de simuler différemment ces phénomènes, le réalisme «allant au-delà» du réel :

- des phénomènes physiques, invisibles par les sens humains (radioactivité, infrarouges, ...) peuvent être représentés virtuellement ;
- la représentation visuelle des objets peut être plus compréhensible s'ils sont affichés en fil de fer, en éclaté ou à une échelle plus grande que nature.

Dans le premier cas de la finalité de la réalité virtuelle, *monde imaginaire*, la virtualité peut être exploitée soit pour créer un monde irréel (art, sciences fiction,...), soit pour représenter des concepts ou des entités symboliques, associés au monde réel simulé. Ceux-ci permettent à l'utilisateur de se faire une meilleure représentation mentale de son environnement. Par exemple :

- Nous pouvons ajouter dans le monde virtuel des informations schématiques pour l'utilisateur, lui permettant de mieux saisir la structure d'un mécanisme ou la planification d'une tâche à accomplir ;
- Nous pouvons représenter la potentialité d'un danger par le simple changement de couleur d'objets, virant au rouge.

Il ne faut donc pas parler du «degré de réalisme», ce terme laissant supposer qu'au mieux le virtuel est identique au réel. Nous préférons parler du domaine de transposition entre les mondes réel et virtuel, qui est l'intersection de deux ensembles :

- l'ensemble des phénomènes réels ;

- l'ensemble de phénomènes représentés virtuellement.

Au moment de la conception d'un dispositif en réalité virtuelle, l'analyse de l'application doit permettre de déterminer ce domaine de transposition entre les mondes réel et virtuel.

1.1.2.3 Définition fonctionnelle

Vers les années 1995 environ, à l'époque des premiers congrès sur la réalité virtuelle à Montpellier, nous essayions de classer les domaines et les fonctions où la réalité virtuelle était pertinente. J'ai proposé pour ma part une taxonomie basée sur ses fonctionnalités «théoriques». Je vous livre ma réflexion, qui est discutable et qui reste théorique et sans beaucoup d'utilité pratique : Face à sa propre perception de la réalité, l'homme a conceptualisé les notions de temps et d'espace sur lequel il ne peut interagir que suivant des lois physiques immuables. D'après Piaget, dont nous reparlerons, l'intelligence sensori-motrice organise le réel en construisant les grandes catégories de l'action que sont les schèmes de l'objet permanent, de l'espace, du temps et de la causalité.

*La réalité virtuelle va lui permettre de s'extraire de la réalité physique pour changer virtuellement de **temps**, de **lieu** et(ou) de type d'**interaction** : interaction avec un environnement simulant la réalité ou interaction avec un monde imaginaire ou symbolique.*

Cette définition fait allusion à l'exigence inverse des auteurs des tragédies du XVII^e siècle prônant la règle des trois unités de temps, de lieu et d'action. Ces potentialités théoriques mais non utopiques à long terme permettent à certains de prédire une révolution importante pour l'homme face à cette nouvelle «réalité». Nous ne débattons pas ici de ce propos, mais nous utiliserons simplement cette approche qui nous permet une taxonomie fonctionnelle des applications de la réalité virtuelle selon les combinaisons de ces trois fonctions (développée en 1.2.1).

1.1.2.4 Définition technique

*Les techniques de la réalité virtuelle sont fondées sur l'**interaction**⁵ en **temps réel** (1) avec un **monde virtuel** (2), à l'aide d'**interfaces comportementales** (3) permettant l'**immersion «pseudo-naturelle»** (4) de(s) l'utilisateur(s) dans cet environnement.*

Seules les installations basées sur ces quatre conditions sont intrinsèquement des systèmes de réalité virtuelle.

1. l'**interaction en temps réel** est obtenue si l'utilisateur ne perçoit pas de décalage temporel entre son action sur l'environnement virtuel et la réponse sensorielle de ce dernier. En pratique, la limite du délai de réaction est de l'ordre de 100 ms ou

⁵A notre avis, la navigation, citée par certains auteurs, n'est qu'une interaction particulière : l'utilisateur interagit en se déplaçant virtuellement par une commande spécifique : le changement de point de vue de la scène virtuelle. S'il ne commande pas lui-même le déplacement, il est alors un passif spectateur et il ne s'agit plus dans ce cas de réalité virtuelle.

moins pour avoir l'impression de temps réel. Cet ordre de grandeur est fonction des sens et des réponses motrices exploités, nous en reparlerons. Donnons un exemple : si l'utilisateur tourne la tête, les images correctes doivent être affichées avec un délai inférieur à 50 ms environ pour que ce dernier ne soit pas perçu ;

2. une des deux problématiques majeures de la réalité virtuelle est la modélisation et la numérisation du **monde virtuel**, l'autre problématique étant celle de l'interfaçage entre le sujet et ce monde virtuel. Nous pouvons noter le cas particulier d'associer un monde réel avec un monde virtuel (techniques de la réalité augmentée que nous présentons succinctement par la suite) ;
3. nous exploitons des interfaces matérielles de la réalité virtuelle, que j'appelle «**interfaces comportementales**» (je préciserai plus loin pourquoi cette nouvelle appellation). Elles sont composées «d'interfaces sensorielles» et «d'interfaces motrices». Les sensorielles informent l'utilisateur par ses sens de l'évolution du monde virtuel. Les motrices informent l'ordinateur des actions motrices de l'homme sur le monde virtuel. Le nombre et le choix de ces interfaces dépendent de l'objectif poursuivi de l'application. Nous détaillerons au chapitre suivant le concept d'interfaçage comportemental ;
4. l'utilisateur doit être en «**immersion pseudo-naturelle**» la plus efficace possible dans le monde virtuel. L'immersion ne peut être naturelle car nous avons appris à agir *naturellement* dans un monde *réel* et non virtuel (il y a aura des biais sensorimoteurs et mentaux, d'où le terme *pseudo*). Cette sensation est une notion en partie subjective qui dépend de l'application et du dispositif utilisé (interfaces, logiciel...). Nous parlerons plus en détail par la suite des concepts d'immersion et d'interaction qui doivent être analysés à plusieurs niveaux.

Ces quatre conditions sont rarement réalisables parfaitement par rapport à l'application envisagée. C'est plutôt un objectif à atteindre. Par contre, elles doivent être en partie réalisées, même modestement, pour parler d'un système basé sur les techniques de réalité virtuelle.

Les schémas suivants (Figure 1.2, Figure 1.3) explicitent la différence essentielle entre l'homme observateur d'un monde virtuel figé (peinture, cinéma, modélisation en CAO, ...) et l'homme acteur dans un monde virtuel interactif.

Les techniques de la réalité virtuelle permettent à toute personne d'agir par l'intermédiaire de ses muscles sur un environnement virtuel, ce dernier réagissant en retour sur les organes sensoriels de l'utilisateur. Les développements récents de certains dispositifs sensoriels et moteurs permettent déjà une première approche. Le monde virtuel étant simulé sur du matériel informatique, la réalité virtuelle implique un changement dans l'interaction entre l'homme et l'ordinateur. Ce dernier ne peut plus être seulement interfacé avec un écran monoscopique, un clavier et une souris 2D si l'on désire une immersion relativement naturelle. Nous préférons employer le terme d'**interfaces comportementales** plutôt que d'interfaces matérielles ou d'interfaces homme machine (IHM), pour bien différencier celles-ci de celles-là. Et nous utiliserons encore moins le terme de «périphérique», qui a une signification négative par rapport à notre approche anthropocentrique et non technocentrique. Pour notre définition, *l'interface comportementale est un dispositif qui exploite un comportement humain, naturel et sans (ou très peu) acquis préalable* (nous préciserons plus loin ces critères). Autrement dit, une interface n'est pas comportementale si elle implique une période d'apprentissage importante et complexe pour pouvoir être utilisée. Au niveau théorique, les «**interfaces sensorielles**» sont conçues pour transmettre des stimuli sensoriels de l'ordinateur vers

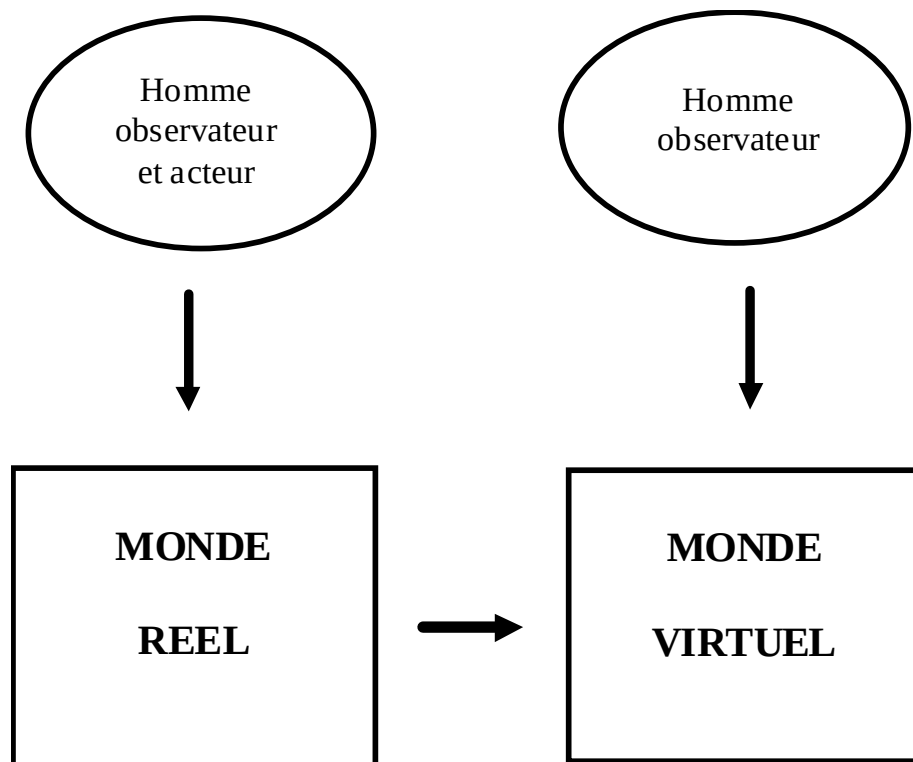


Figure 1.2: Comparaison entre un monde réel et un monde virtuel figé

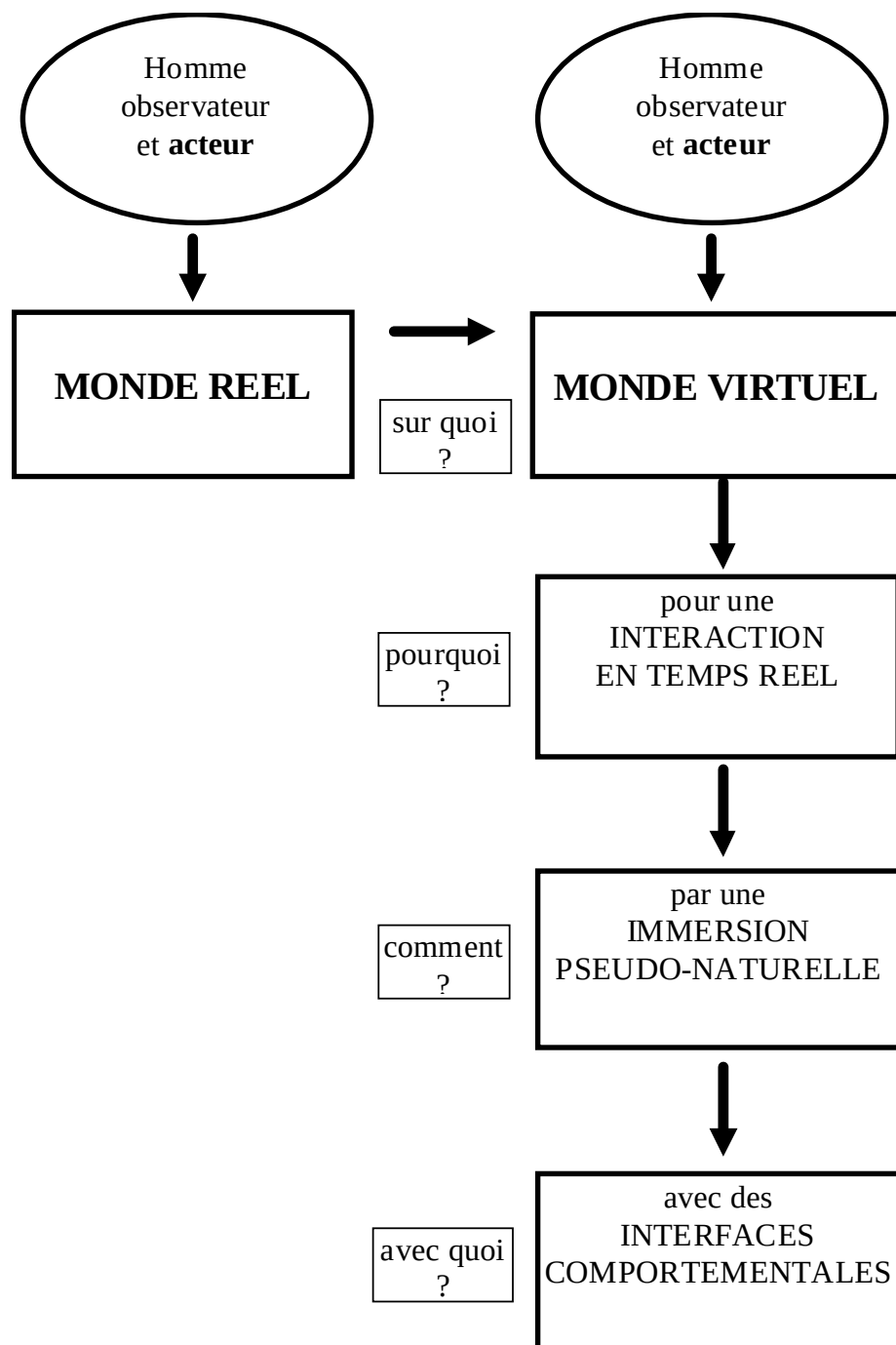


Figure 1.3 : Comparaison entre un monde réel et un monde virtuel interactif

l'homme et inversement les «**interfaces motrices**» pour transmettre des réponses motrices de l'homme vers l'ordinateur. Certaines interfaces mixtes, «**interfaces sensori-motrices**», transmettent des réponses motrices et en réaction des stimuli sensoriels sont renvoyés par l'ordinateur, comme les gants de données à retour d'effort.

Comment les interfaces doivent-elles transmettre les stimuli sensoriels et les réponses motrices ? Comment les utiliser pour permettre un interfaçage comportemental efficace ? Nous expliciterons notre approche et notre méthodologie dans le chapitre suivant.

Nous terminons ce paragraphe en faisant un tour d'horizon sur les définitions de la réalité virtuelle données par d'autres spécialistes. On retrouve toujours les termes d'Immersion et d'Interaction dans les définitions voisines proposées par Coiffet et Burdea [Coiffet, 1993] et par Pimentel et Teixera [Pimentel, 1994]. Par contre, je trouve inapproprié d'y associer le terme «Imagination» (même si cela complète symboliquement les deux premiers «I» de la définition de Coiffet et Burdea). L'imagination (de l'utilisateur) n'est pas indispensable et spécifique à la réalité virtuelle. Elle n'est pas au fondement de la réalité virtuelle comme l'immersion et l'interaction du sujet dans son environnement virtuel : comme je l'expliquerai plus loin, il y a trois niveaux d'immersion et d'interaction. L'imagination ne se situe qu'au niveau mental de l'immersion et de l'interaction. Mais il était si tentant d'associer trois «I». Jean-Paul Papin, [Huriet, 1997], a même donné la «*règle des 5 I : Immersion d'un sujet dans un monde Imaginé et créé Informatiquement qui, grâce à des Interfaces, donne l'Illusion de percevoir et d'agir comme dans le monde réel*». Le dossier de presse de l'INRIA sur la réalité virtuelle [http1] donne pour définition : «*la réalité virtuelle rassemble l'ensemble des techniques et systèmes qui procurent à l'homme le sentiment de pénétrer dans des univers synthétiques créés sur ordinateur avec la possibilité d'y effectuer en temps réel un certain nombre d'actions définies par un ou plusieurs programmes informatiques. Ces techniques lui donnent la possibilité d'éprouver physiquement un certain nombre de sensations et celle de pouvoir opérer dans ces mondes par des moyens d'action naturels*». Dans ces définitions, les auteurs associent des fonctions avec des aspects techniques et matériels de la réalité virtuelle, ce qui est regrettable. Il ne faut pas non plus analyser le domaine de la réalité virtuelle comme une extension en tridimensionnel (3D) du domaine informatique des IHM. Dans ce dernier domaine, l'objectif est de faire communiquer une personne avec une machine «intelligente» pour diverses activités cognitives. En réalité virtuelle, il s'agit fondamentalement de transposer dans un espace virtuel des activités physiques (sensori-motrices). Pour terminer sur les «I», on peut citer l'intéressante approche cognitive de Verna et Grumbach qui parlent des 4 «I» (4 types d'interactions) suivant les différentes combinaisons d'interaction :

- interaction réflexe ou non ;
- interaction physique sur les interfaces ou mentale sur l'environnement virtuel.

Dans un article [Verna, 1999], ces auteurs proposent une approche cognitive pour définir le virtuel. Nous parlerons des différentes Interactions et Immersions dans le chapitre suivant, aussi bien sur les aspects physiques, matériels, sensori-moteurs, mentaux que fonctionnels.

1.1.3 ASPECTS PHILOSOPHIQUE ET SOCIAL DE LA RÉALITÉ VIRTUELLE

La finalité de la réalité virtuelle est de permettre à toute personne des expériences sensori-motrices et donc cognitives dans un monde artificiel, qui peut être ou non réaliste. Mais que pouvons nous dire d'un point de vue philosophique ou social de la «réalité virtuelle»? A la réalité virtuelle, on peut opposer la réalité «concrète» mais il est plus judicieux d'opposer le virtuel au réel. Qu'est ce que le réel, comparativement au virtuel? Sans entrer dans ce débat philosophique, on peut faire deux remarques :

- la différence ne se situe pas au niveau sensori-moteur puisque le réel comme le virtuel sont perçus par nos sens humains et que nous interagissons sur eux par nos réponses motrices ;
- le virtuel se différencie physiquement du réel comme étant une pure création de l'homme, n'ayant pas de substance propre. Le virtuel n'est qu'une modélisation éphémère : concrètement une base de données informatique ayant une certaine interactivité avec l'homme. Le virtuel n'a de réalité temporaire que par l'intermédiaire d'une machine informatique et d'interfaces matérielles⁶ ;

Deux questions importantes se posent sur l'impact social de la réalité virtuelle :

- quels sont les apports et quelles sont les craintes suscitées par la réalité virtuelle pour notre société ?
- quelles similitudes et quelles différences peut-on faire entre le monde réel et un monde virtuel interactif ? Quelles perceptions en a le sujet, présent dans ce nouveau monde virtuel ?

Ces questions renvoient en premier à l'interprétation philosophique du monde réel et à la notion de présence dans celui-ci. Les lecteurs intéressés par ce sujet pourront lire l'article de Sheridan [Sheridan, 1999], comparant les approches philosophiques de la Réalité par Descartes, Heidegger et Gibson, ainsi que l'article de Mantovani et Riva [Mantovani, 1999] sur les différentes présences (physique, subjective et objective) d'un sujet dans un monde réel ou virtuel.

Quelle est la relation de l'homme face aux mondes virtuels ? Comment peut-il se situer par rapport à des entités physiques artificielles dans des espaces et des temps virtuels ? L'espace peut être euclidien pour être assimilable à l'espace réel, mais il peut aussi être géométriquement aberrant, procurant des paradoxes à l'observateur. En particulier, l'unicité de lieu est remise en question : plusieurs utilisateurs en communication virtuelle (télévirtualité) se trouvent chacun corporellement dans un lieu géographique différent et en même temps cognitivement dans un lieu virtuel commun. Celui-ci peut être perçu différemment par chaque personne, même s'il leur est commun. Cette ouverture et ce renouveau de l'espace ne sont pas seulement théoriques, ils peuvent déboucher sur des applications concrètes comme le télétravail, sous réserve de l'adaptabilité de l'homme face à ces nouvelles situations.

L'homme peut jouer virtuellement avec le temps, revenir dans un temps passé ou aller vers le futur. Il peut aussi ralentir ou accélérer le temps pour mieux l'appréhender,

⁶Coupez le courant électrique et tout le virtuel disparaît ! Ce qui est à l'opposé du critère «lavoisien» de la matière réelle.

comme dans la simulation de croissance des plantes ornementales en vue d'une étude esthétique. L'homme perçoit dans le monde virtuel des objets, des entités par leur forme visuelle, tactile ou auditive. Ceux-ci n'ont pas de consistance propre, mais sont seulement des images, et reposent sur des modèles mathématiques, dont l'homme n'en discerne que des représentations. Comme l'écrit Philippe Quéau : «*Le modèle et l'image sont constitués l'un par l'autre. Il y a un aller-retour permanent entre l'intelligibilité formelle du modèle et la perception sensible de l'image*».

Des voix se sont élevées sur les dangers que pouvait induire la réalité virtuelle sur la société et sur nos concitoyens. Si cette crainte est justifiée en théorie, beaucoup n'ont pas vu que les réalisations réelles basées sur la réalité virtuelle impliquent des techniques relativement complexes et donc d'un coût élevé, qui les mettent souvent hors de portée du grand public. Si celui-ci y a seulement accès dans les parcs d'attraction, ce n'est que lors d'un temps d'utilisation court qui n'implique pas une accoutumance et des nuisances effectives. Les jeux vidéo, les images de synthèse et Internet, exploitant aussi le virtuel, ont à cet égard de plus grands impacts sur notre société. A ce propos, les lecteurs intéressés par l'aspect moral et juridique des images de synthèse et du monde virtuel doivent consulter le rapport du sénateur Huriet : «*Images de synthèse et monde virtuel : techniques et enjeux de société*», [Huriet, 1997], ainsi que le chapitre de ce livre, écrit par Jean Paul Papin «*Ethique et droit en réalité virtuelle*».

Ces quelques remarques n'ont pas la prétention de faire un tour d'horizon du débat sur le virtuel. Elles ne sont que des réflexions introductives aux questions soulevées, les réponses dépassant largement le cadre de ce chapitre.

1.1.4 DÉFINITION ET TAXONOMIE DE LA RÉALITÉ AUGMENTÉE

Nous pouvons souhaiter dans certaines applications associer le monde réel avec un monde virtuel. C'est le domaine de la réalité augmentée, dont nous présentons une taxonomie. La réalité augmentée regroupe l'ensemble des techniques permettant d'associer un monde réel avec un monde virtuel, spécialement en utilisant l'intégration d'Images Réelles (IR) avec des Entités Virtuelles (EV) : images de synthèse, objets virtuels, textes, symboles, schémas, graphiques, etc. Plusieurs classifications de la Réalité Augmentée peuvent être envisagées selon les approches technique, fonctionnelle ou tournée vers les applications. Nous proposons une approche fonctionnelle de l'association d'images réelles avec des entités virtuelles : quels types d'association peut-on réaliser, sans s'occuper particulièrement des techniques d'affichage ou des catégories des applications concernées ? Différemment, une approche technique est donnée par le professeur Milgram de l'université de Toronto [Milgram, 1996], qui privilégie les différentes techniques d'affichage ou de vision du réel : on peut parler aussi de réalité augmentée quand l'observateur voit directement l'environnement réel à travers un visiocasque semi-transparent sans passer par des caméras et des images vidéo. Dans notre classification, nous ne ferons pas de différence entre la vision du monde réel par l'intermédiaire de caméras, en vision indirecte, et celle par l'intermédiaire de visiocasques semi-transparents (HMD see-through), en vision directe. Dans le cas de vision indirecte, nous nous limitons aux images vidéo de la scène réelle qui sont transmises en temps réel (en contre exemple, les films avec trucages numériques en images de synthèse ne font pas partie de notre classification).

Les catégories fonctionnelles sont basées sur l'importance de la relation entre les images réelles avec les entités virtuelles affichées sur un même écran (ou vues à travers un

écran, dans le cas de visiocasques semi-transparents). On peut donc écarter le cas où les images réelles (IR) et les entités virtuelles (EV) sont affichées sur un même écran, mais sans aucun rapport entre elles. Dans ce cas, l'utilisateur emploie un seul écran pour deux affichages indépendants (On peut classer ce cas en fonctionnalité 0 !). La fonctionnalité 0 sera définie par l'absence de relation entre le contenu des EV et des IR. Les IR et les EV peuvent être affichées soit simultanément sur l'écran total, soit sur deux fenêtres du même écran ou soit sur une fenêtre et en fond d'écran. Pour regrouper ces différents cas, on emploie indifféremment le terme de «cadre d'affichage».

Remarque préliminaire : Nous présentons à la suite des fonctionnalités numérotées de 1 à 5, mais il n'y a pas intrinsèquement une hiérarchie entre elles : chacune de celles-ci a des objectifs différents. On peut toutefois dire qu'elles induisent secondairement des difficultés scientifiques et techniques d'ordre approximativement croissant de la fonctionnalité 1 à la fonctionnalité 5.

1.1.4.1 *Fonctionnalité 1 : «Réalité Documentée» ou «Virtualité Documentée»*

C'est la fonctionnalité minimum de la réalité augmentée où les IR et les EV sont dans deux cadres d'affichage différents, mais ayant un rapport d'information entre eux. En général, ce sont les EV (du texte principalement) qui permettent de mieux comprendre les IR. On peut dans le cas de cette fonctionnalité parler de «Réalité Documentée».

Mais l'inverse peut être envisagé dans quelques cas particuliers, par exemple, le synoptique d'un processus industriel (EV) avec incorporation en temps réel d'une ou de plusieurs fenêtres visualisant des parties réelles du processus (IR). Dans ce cas particulier, on parle de «Virtualité Documentée».

1.1.4.2 *Fonctionnalité 2 : «Réalité à Compréhension ou Visibilité Augmentée»*

Dans cette fonctionnalité les IR et les EV sont dans un même d'affichage. On peut faire deux sous-niveaux selon l'apport de l'augmentation :

Fonctionnalité 2.1 : «Réalité à Compréhension Augmentée»

Il s'agit dans ce cas de l'augmentation de la *compréhension* des images de la scène réelle par incrustation d'informations sémantiques passives (Figure 1.4). Les EV (titres, légendes, symboles), plus ou moins proches visuellement des objets réels, donnent des informations complémentaires sur ces derniers (fonctions, références,...). L'inversion des IR et des EV dans cette fonctionnalité n'est pas envisageable.

Fonctionnalité 2.2 : «Réalité à Visibilité Augmentée»

Il s'agit dans ce cas de l'augmentation de la *visibilité* des images de la scène réelle (Figure 1.5). Les EV (modèle «fil de fer» des objets réels) sont en correspondance géométrique précise avec les contours des objets réels. Ils permettent de surligner les objets pour mieux les voir. Cela n'est bien sûr possible que si l'on connaît les modèles des objets et que l'on peut établir une correspondance géométrique correcte entre le



Figure 1.4: «Réalité à Compréhension Augmentée», photo Ecole des Mines de Paris

monde virtuel et le monde réel. Ceci implique, entre autres, un calibrage des caméras. On peut distinguer deux cas :

- on améliore la netteté des images en surlignant les contours apparents des objets pour mieux les percevoir ;
- on améliore la compréhension des objets en représentant virtuellement les contours visibles et les contours invisibles.

1.1.4.3 Fonctionnalité 3 : «ssociation Visuelle du Réel et du Virtuel»

Dans cette fonctionnalité, on ajoute de nouveaux objets virtuels à la scène réelle. On peut faire la différence entre 2 cas, selon le niveau d'intégration des objets virtuels dans la scène réelle :

- fonctionnalité «Incrustation d'Objets Virtuels sur des Images Réelles». Les objets virtuels sont incrustés (overlay) par-devant les objets réels, donc sans aucune occultation des objets virtuels par les objets réels. On peut dire qu'il s'agit d'une *association par superposition* ;
- fonctionnalité «Intégration d'Objets Virtuels dans une Scène Réelle». Les objets virtuels sont intégrés avec les objets réels (Figure 1.6). Ces derniers occultent les objets virtuels qui sont positionnés derrière eux. On peut dire qu'il s'agit d'une *association tridimensionnelle*. Mais l'intégration avec occultation en temps réel reste un verrou technologique important. On peut consulter à ce propos les recherches de D. Ernadote, [Ernadote, 1997].

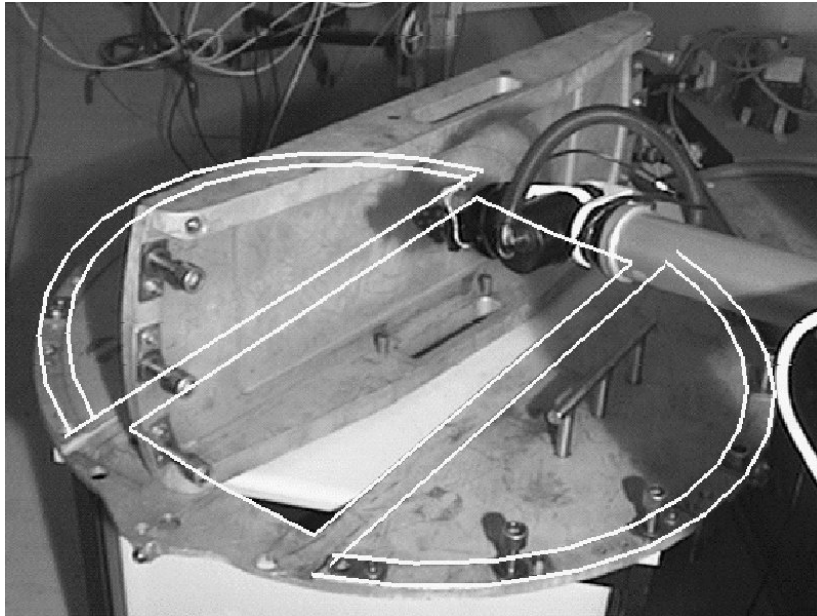


Figure 1.5 : Réalité à Visibilité Augmentée avec représentation des contours visibles et invisibles, photo Ecole des Mines de Paris

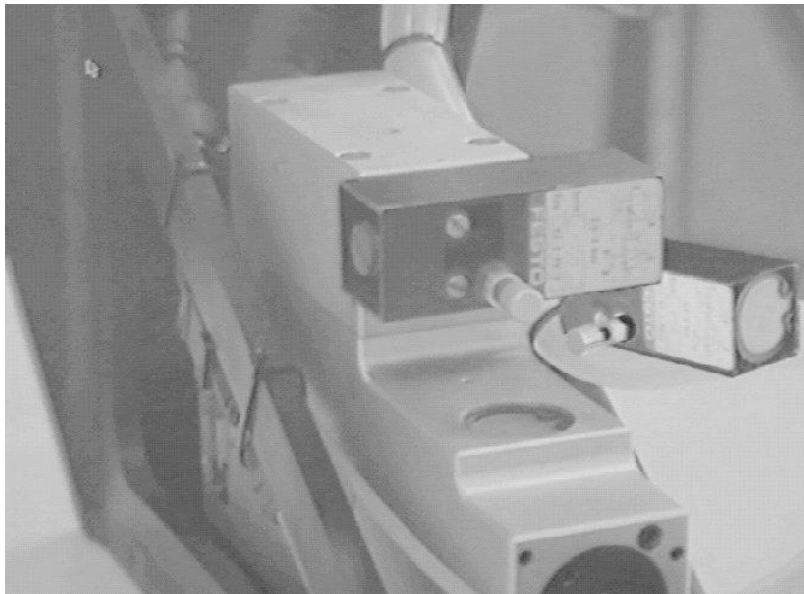


Figure 1.6 : Association d'un distributeur virtuel autour d'un mécanisme réel ayant le même distributeur (Simulation de montage du mécanisme), photo Ecole des Mines de Paris

1.1.4.4 Fonctionnalité 4 : «Association Comportementale du Réel et du Virtuel»

En plus des conditions de la fonctionnalité précédente, on modélise sémantiquement les objets virtuels en tenant compte de leurs propriétés physiques suivant les lois de gravité, de contact, d'élasticité, de fluidité, etc., ceci pour enrichir la scène. On utilise les connaissances a priori sur la scène réelle et de ses objets. Cette fonctionnalité définit les interactions géométriques et physiques entre les objets réels et les objets virtuels. En exemple, on peut citer les fonctions de comportement et d'attraction des objets virtuels avec des objets réels (voir les recherches de D. Maman, [Maman, 1998]).

1.1.4.5 Fonctionnalité 5 : «Substitution du Réel par le Virtuel» ou «Réalité Virtualisée»

Si on connaît la modélisation géométrique de la scène réelle observée, on peut remplacer l'affichage de l'image vidéo de la scène réelle par l'image de synthèse du modèle, déterminée à partir du même point de vue. En passant de l'image vidéo à l'image de synthèse, et réciproquement, on augmente la compréhension de l'environnement. On parle dans ce cas de «Réalité Virtualisée», [Boulanger, 1998]. On peut alors aussi changer le point de vue pour l'image de synthèse, sans déplacer la caméra filmant la scène réelle, ce qui permet une meilleure compréhension de la scène.

En exemple, nous pouvons citer les travaux de téléopération effectués à la DER d'EDF, aux quels nous avons collaborés. Dans leur dispositif de visualisation, on ne remplace pas l'affichage de l'image vidéo de la scène réelle par l'image de synthèse du modèle, mais on les propose sur *deux* écrans voisins. Le robot, que l'opérateur est en train de manipuler, est affiché simultanément en images réelles en relief et en images de synthèse en relief (Figure 1.7). Quel que soit le point de vue des caméras réelles, l'opérateur peut toujours observer sur un écran en «Réalité virtualisée» la configuration mécanique de son robot avec des images de synthèse animées en temps réel.



Figure 1.7 : Le robot téléopéré est affiché simultanément en images réelles en relief et en images de synthèse en relief, photo Ecole des Mines de Paris

1.1.5 HISTORIQUE SUCCINCT DE LA RÉALITÉ VIRTUELLE

La naissance des techniques de la réalité virtuelle est étroitement liée à l'évolution de l'informatique, en particulier à celle du domaine de la synthèse des images. On peut rappeler certains jalons des progrès techniques ayant permis d'envisager par la suite les systèmes de réalité virtuelle :

- 1962 : le précurseur des systèmes d'immersion virtuelle, le Sensorama. Ce dispositif permettait à un utilisateur d'être immergé visuellement et auditivement dans une scène réelle, préalablement filmée. De plus l'utilisateur percevait des odeurs et des vibrations. Le but du système était de simuler un déplacement virtuel en véhicule. Reposant sur des techniques non informatiques, l'utilisateur était spectateur et il resta à l'état de prototype ;
- vers 1960 : c'est le début de la micro-informatique, avec en particulier la création d'une interface «comportementale» entre l'homme et la machine, la souris. C'est la première interface matérielle entre la machine et l'homme qui exploite un de ses comportements naturels : le mouvement de la main ;
- 1962 : le premier logiciel graphique permettant la simulation visuelle de données est créé, ce qui permet l'essor de la CAO, la Conception Assistée par Ordinateur ;
- vers 1970 : des logiciels sont conçus pour la navigation et l'exploration de mondes virtuels visualisés sur un écran ;
- 1970 : le premier «visiocasque» est réalisé par Daniel Vickers de l'Université d'Utah. Ce dispositif permet une immersion visuelle dans un monde virtuel. L'utilisateur, coiffé d'un casque supportant deux mini-écrans, peut observer une scène virtuelle tout en tournant la tête ;
- 1982 : une nouvelle interface comportementale apparaît : le gant de données. Ce dispositif mesure les déplacements des doigts et de la main de l'opérateur et permet à ce dernier de communiquer avec l'ordinateur ;
- 1980/1985 : c'est le début de la commercialisation de matériels et de logiciels spécifiquement dédiés à la réalité virtuelle ;
- 1990/1995 : Le grand public prend connaissance des premiers développements de la réalité virtuelle grâce aux médias ;

Les lecteurs intéressés par l'historique de la réalité virtuelle peuvent consulter le livre de Rheingold, [Rheingold, 1993], qui fut le premier à écrire sur le développement de la réalité virtuelle, spécialement aux États-Unis.

Ce sont principalement les chercheurs américains qui sont à l'origine des techniques de la réalité virtuelle. Cette dernière a donné naissance à un grand nombre de nouveaux défis scientifiques que la communauté des chercheurs souhaite résoudre. Parallèlement, le développement de produits fiables et à prix accessibles est un challenge pour le milieu industriel. D'importants efforts de recherche ont déjà été entrepris depuis les années 80-90 pour concevoir des interfaces matérielles permettant une meilleure immersion, et pour développer des logiciels puissants dédiés à la réalité virtuelle. Au début, dans la décennie 80, des logiciels spécifiques pour la réalité virtuelle ont été développés par de petites sociétés (World Tool Kit de Sense 8, Clovis de Médialab, etc.). Mais, vers la fin des années 90, plusieurs sociétés du secteur de la réalité virtuelle ont été rachetées par les grandes sociétés de la CAO, montrant l'intérêt de ces dernières pour incorporer dans leur logiciel CAO les fonctionnalités de la réalité virtuelle. Nous expliciterons plus en

détail les liens entre CAO et RV. Les interfaces sont souvent fabriquées par de petites sociétés (principalement américaines et japonaises), le marché ne se développant pas aussi rapidement que certains l'espéraient au début des années 90. Et pour terminer ce rapide résumé historique, quelques anecdotes : vers les années 94-97, de grandes sociétés fabricant les jeux vidéo ont fait plusieurs années de suite, au moment des fêtes de fin d'année, des annonces de ventes de visiocasques (HMD) pour les jeux. Ils n'ont jamais été commercialisés, montrant ainsi que les contraintes techniques, économiques et de santé publique étaient plus importantes que prévues dans le développement des interfaces. Le gant de données, Powerglove de Mattel, n'a été vendu que quelques mois pour les jeux car il était peu fiable et exploitable qu'avec quelques jeux vidéo.

1.1.6 LES DOMAINES PRÉCURSEURS DE LA RÉALITÉ VIRTUELLE

La réalité virtuelle ne s'est pas développée isolément, indépendamment de domaines voisins. Comme nous l'avons indiqué précédemment, les simulateurs de transport furent des dispositifs de réalité virtuelle avant la lettre. Son essor repose sur les progrès enregistrés dans différents domaines dont elle dépend. Elle s'appuie sur le domaine de la simulation «interactive»⁷, sur le domaine de l'infographie pour la création d'images de synthèse en temps réel et sur le domaine de la téléopération pour l'immersion de l'utilisateur à partir d'interfaces. Les interfaces entre le téléopérateur et le poste de contrôle, situées hors de la zone de téléopération, ont les mêmes objectifs que ceux des interfaces de réalité virtuelle : permettre une immersion et une interaction en temps réel, soit sur une zone réelle distante ou soit sur un environnement virtuel (Figure 1.8).

1.1.7 LE DÉVELOPPEMENT DE LA RÉALITÉ VIRTUELLE EN FRANCE

Les recherches en France ont commencé au début des années 90 dans des centres de recherche du domaine de l'informatique et de la robotique (l'INRIA, notre Centre de Robotique de l'École des Mines de Paris, le Laboratoire de Robotique de Paris, l'ACROE, le Laboratoire d'Informatique Industrielle de l'ÉNIB de Brest, ...) et du domaine des sciences cognitives (le Laboratoire de Physiologie de la Perception et de l'Action du Collège de France). Rapidement, début 1994, les centres de recherche français se sont regroupés au sein du Groupe de Travail sur la Réalité Virtuelle (GT-RV), sur l'initiative de l'équipe ÉPIDAURE de l'INRIA Sophia-Antipolis. Le groupe de travail s'est donné comme but de fédérer les efforts de recherche des équipes universitaires et industrielles françaises travaillant dans les différents domaines d'application de la réalité virtuelle. Depuis, le GT-RV organise des journées de travail qui comprennent des exposés, des démonstrations et des visites. Ces journées se sont transformées en congrès internationaux (Laval 99, Brest 2000). Le Groupe édite une lettre électronique d'informations, *Rêveries*, qui permet de maintenir un contact permanent entre les participants du groupe de travail. Le GT-RV est animé par Gérard Subsol (ÉPIDAURE, INRIA, Sophia Antipolis) et par Jean-Luc Dugelay (EURÉCOM, Sophia Antipolis), Adresse : [http2].

Au début des années 90, nous avons axé nos recherches à l'École des Mines de Paris, comme d'autres, sur les aspects techniques de la Réalité virtuelle : les interfaces et les

⁷A bien différencier de la simulation de lois physiques sur un produit ou processus, calculée en temps différé.

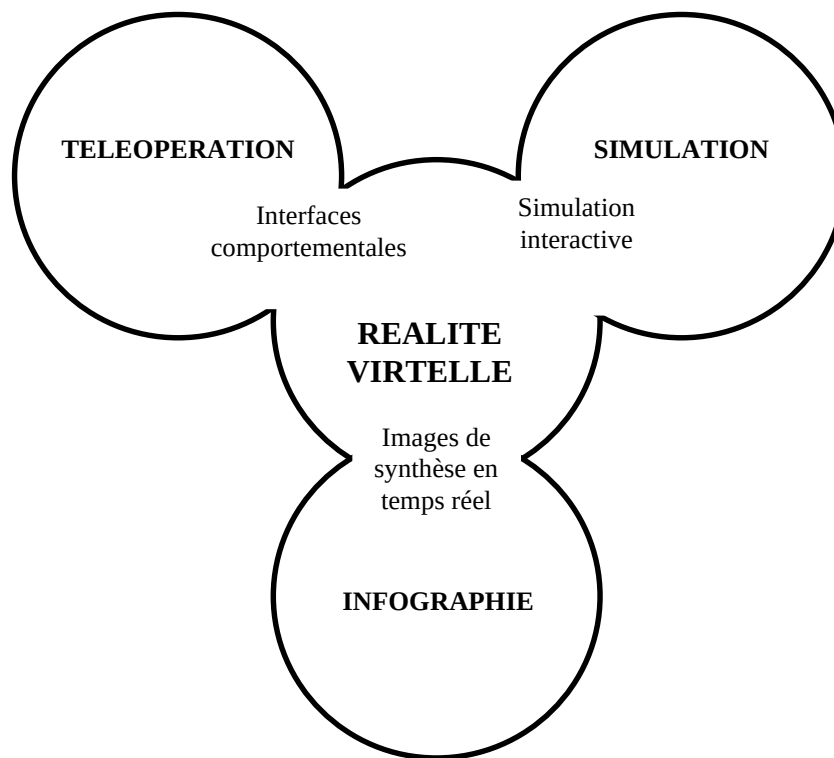


Figure 1.8 : Domaines précurseurs de la réalité virtuelle

logiciels RV. Mais, depuis quelques années, la recherche se développe aussi sur une autre facette de la réalité virtuelle, à laquelle nous avons contribué⁸ : l'étude du comportement humain et de son interfacement comportemental dans les environnements virtuels. Cette nouvelle approche nécessite l'apport de connaissances des chercheurs en sciences cognitives, avec qui nous collaborons.

La réalité virtuelle ne s'est pas non plus développée rapidement et exponentiellement en France et dans le monde, comme certains le pensaient. De jeunes entreprises ont dû changer d'orientation, le marché du début des années 90 n'étant pas assez porteur. Malgré des promesses lancées un peu vite par l'entreprise Simulis, aucun visiocasque n'a été fabriqué en France à ce jour. Les premiers congrès sur la réalité virtuelle à Montpellier ont eu des périodes florissantes mais celles-ci ne se sont pas pérennisées (dernier congrès de Montpellier en 1997). Le club de réalité virtuelle de l'IRIAM de Marseille ne fonctionne plus. Actuellement, en France, ce sont principalement les grandes entreprises, spécialement du domaine du transport et de l'énergie, ainsi que le secteur médical, qui consacrent une part importante de leurs moyens de recherche et de développement. Début 2001, un consortium d'une quinzaine d'industriels et de centres de recherche français mène une action d'envergure sur trois ans pour développer une plateforme en réalité virtuelle (PERF RV) dans le cadre des projets RNTL, subventionnés par le gouvernement [http3].

Comme nous l'avons indiqué, les premiers congrès français du domaine se sont tenus à Montpellier de 1992 à 1997, «L'interface des mondes réels et virtuels». Les fameux «Voyages Virtuels», sponsorisés par «les Trois Suisses» au début des années 90, informèrent le grand public. Après, des congrès sur la réalité virtuelle ont eu lieu à Nîmes (98), Laval (1999 à 2001), [http4] et à Brest (2000), [http5]. Les «Rencontres de la réalité virtuelle» ont lieu chaque année à l'Institut Image de Chalon sur Saône depuis 98, [http6], animées en outre par Jean Ségura, journaliste du domaine. A noter que dans cette ville, un Centre de Création en Réalité Virtuelle (CCRV) vient d'être créé pour développer des activités artistiques avec les techniques RV. Des journées ARÉVI se tiennent aussi chaque année à l'ÉNIB de Brest depuis 1996 [http7] et en 1999 un colloque, RÉVICO [http8], a regroupé les chercheurs de la réalité virtuelle et des sciences cognitives à l'ENST de Paris. Depuis cinq ans, l'exposition MICAD (domaines de la CAO et de la FAO) [http9], consacre un espace à la réalité virtuelle ainsi que des cours, que nous animons, faisant le lien entre la CAO et la RV. En juillet 2000, s'est tenu le congrès Virtual Worlds [http10].

Des formations sur la réalité virtuelle sont proposées par différents instituts : les masters en réalité virtuelle de l'ÉNIB de Brest et de l'Institut Image de Chalon sur Saône, le DEA RV&MSC, organisé par les Universités d'Évry, de Versailles et l'INSTN du CEA [http11]. L'École des Mines de Paris, participe au DEA RV&MSC et propose des enseignements en réalité virtuelle auprès de ses étudiants et auprès des industriels [http12]. Une formation DESS en innovation technologique et réalité virtuelle a débuté en septembre 2000 à Laval et est organisée par l'ISTIA d'Angers [http13].

Nous ne pouvons citer toutes les entreprises françaises fournisseurs de produits logiciels ou matériels pour la réalité virtuelle. Mais nous tenons à mentionner deux entreprises dynamiques dont les deux directeurs, Pascal Chevalier pour Sim Team [http14] et Christophe Chartier pour Immersion S.A. [http15] oeuvrent, depuis dix ans déjà, à l'expansion en France de la réalité virtuelle. Ces deux entreprises distribuent la plu-

⁸Notre principale contribution est présentée dans le chapitre suivant.

part du matériel de la réalité virtuelle et proposent leur service pour développer des applications professionnelles.

Il n'y a pas de revue spécialisée en français sur le domaine, malgré les efforts entrepris par certains. Seul, le GT-RV édite la lettre électronique d'informations «Rêveries». Les lecteurs peuvent consulter les quelques livres en français : ceux de Rheingold [Rheingold, 1993], de Coiffet et Burdea [Coiffet, 1993], de Cadoz [Cadoz, 1994], de Quéau [Quéau, 1993] et le rapport du sénateur Huriet [Huriet, 1997]. Mon livre, «*Les interfaces de la réalité virtuelle*» [Fuchs, 1996], n'a plus d'intérêt pour vous puisque vous lisez ce texte, dont il est issu et actualisé.

1.1.8 UN PETIT BÊTISIER DE LA RÉALITÉ VIRTUELLE

Je ne peux résister à l'envie de rédiger ce petit bêtisier dans l'espoir que chacun⁹ ne commette plus ces petites «bêtises» que l'on entend encore quelques fois. Au commencement de ce nouveau millénaire, il serait bon de ne pas continuer à propager ces erreurs :

1. **«Les systèmes de réalité virtuelle sont ceux qui utilisent un visiocasque».** Que de déception, au début de la réalité virtuelle, chez les professionnels qui voulaient à tout prix voir dans un visiocasque. On avait beau leur dire que ce n'était guère montrable et peu enthousiasmant pour des raisons techniques, ils voulaient voir...et repartaient bien souvent déçus, les visiocasques n'étant pas au point (l'évolution des visiocasques est lente, voir le chapitre «interfaces visuelles»). Mais soyons honnêtes, nous ne pouvons pas reprocher aux spécialistes des simulateurs de transport (automobile, aéronautique, etc.), faisant de la «réalité virtuelle» avant l'invention du terme RV, d'associer la RV aux visiocasques qui sont apparus en même temps, d'où la confusion qui perdure...malgré nos efforts chez un constructeur automobile ;
2. **«La réalité virtuelle est immersive quand on utilise un visiocasque».** Étonnant : l'immersion du sujet ne dépendrait toujours que d'une seule caractéristique psychophysique (l'immersion du regard¹⁰) d'un seul sens humain, la vision. Notons, et nous en reparlerons, que l'obtention d'un grand champ de vision ne se fait pas que dans un visiocasque. Nous avons souvent de plus grands champs de vision avec des grands écrans. Nous avons réalisé, comme d'autres, des immersions très efficaces sans visiocasque ;
3. **«La réalité virtuelle est définie comme un ensemble de logiciels qui...».** Définition bien restrictive ! S'il n'y avait que des problèmes logiciels à résoudre ;
4. **«Les interfaces de la réalité virtuelle sont appelées interfaces sensorielles (ou perceptives)».** Comme nous l'avons expliqué, l'innovation qu'apporte la réalité virtuelle est justement de permettre à un utilisateur «d'agir virtuellement», ceci n'étant donc possible qu'avec des interfaces...«motrices» !
5. **«La réalité virtuelle peut (doit) exploiter les cinq sens de l'homme».** Mais l'homme a plus de cinq sens¹¹ et ce sont souvent les sens oubliés (kinesthésie,

⁹Mea culpa. ...J'ai moi-même fait l'erreur 4 à mes débuts.

¹⁰L'immersion du regard : l'utilisateur voit toujours la scène virtuelle, même s'il tourne sa tête (son regard), voir chapitre «interfaces visuelles».

¹¹Cette erreur perdure dans le grand public.. Les professeurs de SVT devraient plus insister sur la connaissance du corps humain.

perception proprioceptive des efforts) qui posent des problèmes complexes dans bien des applications RV (les simulateurs de transport, par exemple). Une présentation des sens est exposée au chapitre «Les sens et les réponses motrices de l'home» ;

6. «**La réalité virtuelle utilise obligatoirement les images en relief**». Et pourtant bien des dispositifs RV fonctionnent parfaitement sans image en relief (voir le chapitre "interfaces visuelles pour l'apport et les contraintes des images en relief) ;
7. «**Nous percevons la profondeur, l'espace 3D, grâce aux images en relief, grâce à nos deux yeux**». Le lecteur non convaincu est prié de fermer un œil et de regarder autour de soi : l'environnement est-il plat, est-il perçu qu'en deux dimensions ? (On en reparlera aussi) ;
8. «**Grâce à la réalité virtuelle, on testera demain virtuellement la conduite de sa nouvelle voiture avant de l'acheter**», dixit un journaliste. Quand on connaît le prix d'un simulateur de voiture qui coûte au moins dix à cinquante fois plus cher qu'une voiture et les difficultés qu'impose la simulation, on peut plus que douter de cette affirmation. Mais certains journalistes aiment tant le sensationnel ;
9. «**La réalité virtuelle va pénétrer les jeux : en 1995, 1% des jeux pouvaient être qualifiés de jeux de réalité virtuelle. Cette proportion devrait atteindre 20% en 2001**», d'après un certain cabinet d'études anglais. A moins que tous jeux exploitant les images en 3D soient déclarés être des jeux basés sur la réalité virtuelle, on est bien loin du compte. Mais il faut reconnaître qu'il est difficile de faire des pronostics sur l'évolution de la réalité virtuelle et de son matériel : je pronostiquais bien à tort, il y a cinq ans, une évolution plus rapide des gants de données (voir le chapitre correspondant) ;
10. **A vous, lecteur, de m'indiquer la dixième.**

1.2 LES APPLICATIONS DE LA RÉALITÉ VIRTUELLE

1.2.1 TAXONOMIE THÉORIQUE DES APPLICATIONS RV

Face à la multitude d'applications potentielles, une taxonomie fonctionnelle est souhaitable, plutôt que de passer en revue les secteurs d'activité de l'homme, qui sont tous susceptibles d'exploiter les techniques de réalité virtuelle. Cette classification est critiquable, mais elle a le mérite d'être indépendante des domaines d'activité. Elle est donc basée sur **les fonctions intrinsèques de la réalité virtuelle** : l'utilisateur peut s'extraire de son propre espace, du temps présent pour interagir dans un monde artificiel. Ces trois possibilités de changement induisent des états différents qui peuvent se combiner (Figure 1.9).

- déconnecté du temps présent (T_0), l'utilisateur peut se mouvoir dans le passé (T_-) ou le futur (T_+) ;
- il peut désirer se projeter soit dans un lieu inaccessible (lieu géographique ou espace microscopique par exemple) ($L_{\rightarrow}$), soit s'associer à plusieurs dans un lieu virtuel ($L_{\cup}$). La notation (L_0) indique que le lieu est inchangé ou indifférent pour l'application envisagée ;

- la réalité virtuelle crée des interactions soit dans un monde simulant la réalité (IA_r), soit dans un monde imaginaire ou symbolique (IA_i). La notation (IA_0) indique que les interactions se font dans le monde réel.

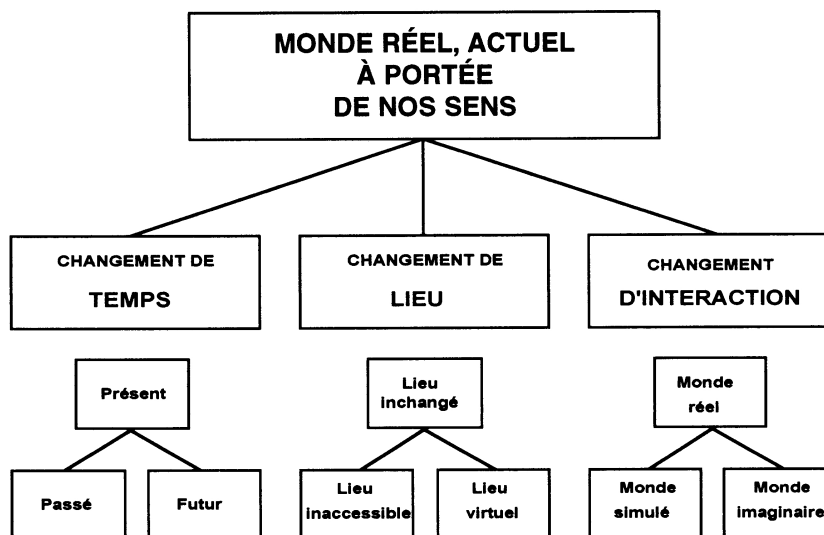


Figure 1.9: Les fonctions intrinsèques de la réalité virtuelle

Remarques :

La réalité virtuelle implique des interactions IA_r ou IA_i . La téléprésence et la téléopération (interactions à distance avec un monde réel : $IA_0.T_0.L_{\rightarrow}$), qui ne font pas partie de la réalité virtuelle, exploitent en partie des interfaces de mêmes types. Une remarque identique s'applique aux télécommunications par téléphone, visiophone ou vidéoconférence classique ($IA_0.T_0.L_{\cup}$).

Il y a théoriquement $2 \times 3^2 = 18$ combinaisons possibles pour les deux classes IA_r et IA_i , mais l'association de plusieurs utilisateurs ($L_{\cup}$) ne peut se faire qu'au présent (T_0), d'où 4 combinaisons qui ne peuvent être envisagées : $(IA_r + IA_i).L_{\cup}.(T_{-} + T_{+})$.

1.2.2 LES APPLICATIONS POTENTIELLES DE LA RÉALITÉ VIRTUELLE

Nous obtenons ainsi une taxonomie des applications de la réalité virtuelle qui regroupe tous les domaines, sans faire intervenir des considérations technico-économiques.

1.2.2.1 Interactions avec un monde simulant la réalité

L'objectif de ces applications est d'offrir la simulation de certains aspects de la réalité pour mieux les appréhender.

- **IA_r.T₀.L₀** **Activité virtuelle** : le lieu et le temps étant indifférents, l'utilisateur agit physiquement sur une scène virtuelle pour ses loisirs, pour une formation incluant un apprentissage physique. *Exemples : Jeux, sports et loisirs virtuels, formation virtuelle dans les secteurs suivants : médecine, chirurgie, rééducation, armée, industries aéronautique et spatiale...*
- **IA_r.T₀.L_→** **Transfert virtuel** : le temps étant indifférent, l'utilisateur est transporté dans un autre lieu ayant une scène simulant le réel pour son travail, ses loisirs, ses activités quotidiennes. *Exemples : Tourisme virtuel, étude dans un monde microscopique ;*
- **IA_r.T₀.L_u** **Téléassociation virtuelle** : l'association de plusieurs personnes dans un lieu virtuel permet à celles-ci de se rencontrer. *Exemples : Jeux et sports d'équipes virtuels, téléconférence avec outils de réalité virtuelle, cité virtuelle ;*
- **IA_r.T₊.L₀** **Conception virtuelle** : La réalité virtuelle permet de concevoir et d'expérimenter de futurs produits. Elle ajoute une dimension supplémentaire par rapport à la CAO : tester le comportement de l'homme pour des produits qui ne sont pas encore fabriqués. *Exemples : Architecture virtuelle, conception de produits industriels, de vêtements ;*
- **IA_r.T₊.L_→** **Aménagement virtuel** : même possibilité que pour la conception virtuelle, mais concernant des espaces à aménager. *Exemples : Conception d'environnements urbains ou paysagistes, ergonomie d'un poste de travail, d'un atelier ;*
- **IA_r.T₋.L₀** **Exposition virtuelle** : Il s'agit dans cette application de recréer et d'observer des objets qui n'existent plus. *Exemples : Études de produits anciens ou d'œuvres d'art détruits ;*
- **IA_r.T₋.L_→** **Événement virtuel** : en recréant d'anciens événements, on permet à l'utilisateur de mieux les appréhender et les comprendre. *Exemples : études d'anciennes civilisations, visite d'anciens lieux, enquêtes policières par reconstitution virtuelle.*

1.2.2.2 Interactions avec un monde imaginaire ou symbolique

L'objectif de ces applications est de proposer à l'utilisateur soit des mondes imaginaires à des fins ludiques ou artistiques, soit des mondes symboliques pour concrétiser des phénomènes, des concepts grâce à des métaphores.

- **IA_i.T₀.L₀** **Création virtuelle** : Création virtuelle : la réalité virtuelle permet de créer soit des œuvres éphémères et en interaction avec le spectateur, soit des métaphores. *Exemples : Art virtuel, musée virtuel, cinéma interactif, vie artificielle, représentation symbolique d'un marché financier, navigation dans une base de données ;*
- **IA_i.T₀.L_→** **Navigation symbolique** : l'utilisateur peut se déplacer dans un graphe, un réseau ou un arbre représentant symboliquement des structures réelles. *Exemple : navigation dans un réseau de communication ;*
- **IA_i.T₀.L_u** **Télévirtualité** : relation entre plusieurs utilisateurs par l'intermédiaire de clones. *Exemples : Télécommunication par clonage, communauté virtuelle ;*
- **IA_i.T₊.(L₀+L_→)** **Science-fiction virtuelle** : l'utilisateur est projeté dans un monde irréel et futur. *Exemples : Œuvre virtuelle, jeux virtuels ;*
- **IA_i.T₋.(L₀+L_→)** **Passé imaginaire virtuel** : l'utilisateur est projeté dans un monde irréel et passé. *Exemples : Œuvre virtuelle, jeux virtuels ;*

1.2.3 LES DOMAINES ÉCONOMIQUES DES APPLICATIONS ACTUELLES

Sans passer en revue tous les domaines où les techniques de la réalité virtuelle peuvent s'appliquer, nous pouvons indiquer ceux qui ont une certaine importance au niveau économique.

1.2.3.1 Domaines non professionnels

- Les activités ludiques : quelques jeux individuels utilisent les techniques de réalité virtuelle. Mais nous assistons surtout au développement de jeux de réalité virtuelle dans les parcs d'attraction et les salles de jeux (simulation de sport, simulateur de transport,...). Il faut bien différencier les jeux vidéo des jeux basés sur la réalité virtuelle. Nous pouvons classer les jeux en trois catégories, même s'il y a une certaine continuité au niveau de l'immersion et de l'interaction entre ces catégories :
 - le jeu simple avec interaction en temps réel mais sans recherche d'immersion (par exemple, jeux de logique ou d'adresse, comme un flipper sur ordinateur). L'interaction est simple : commande binaire en général par clavier ou manette de jeu ;
 - le jeu avec interaction en temps réel et faible immersion, faible au sens où il y a peu d'immersion sensori-motrice, l'immersion recherchée étant principalement mentale (par exemple, jeux de simulateur de vol ou de course de voiture, sur simple ordinateur) ;
 - le jeu avec interaction et grande immersion sensori-motrice et mentale, le jeu véritablement basé sur les techniques RV.
- Les activités artistiques et culturelles, que l'on peut citer, malgré qu'elles n'induisent pas un secteur économique important. Il y a les œuvres d'art virtuels, par exemple, « Elle et la voix » de Catherine Ikam, [http16] et les spectacles exploitant la réalité virtuelle (par exemple, les marionnettes virtuelles d'Animação [Jessel, 1999]). Mais dans ce dernier cas, ce sont seulement les acteurs qui sont en immersion et en interaction.

1.2.3.2 Domaines professionnels, hors industrie

- L'information et la formation, avec apprentissage physique : par exemple, information sur le déroulement d'opération de secours, formation à la maintenance dans des centrales nucléaires, à la conduite de véhicule, à des tâches sur infrastructures ferroviaires [Lourdeaux, 1999], au sport (par exemple : le golf virtuel [Huriet, 1997]), etc. La formation classique en environnement réel ne répond pas toujours aux besoins pédagogiques, que la réalité virtuelle peut permettre en environnement virtuel :
 - Réaliser des tâches sans danger ;
 - Reconfigurer l'environnement : changement de terrain, de conditions climatiques ;
 - Réaliser des scénarios impossibles à reproduire dans la réalité (accidents, incidents techniques, opération chirurgicale,...) ;
 - Représentation plus explicite du réel (transparence ou surlignement d'objets, ...), [Lourdeaux, 1999].
- La médecine : dans ce domaine, le développement de l'imagerie médicale tridimensionnelle, parallèlement à celui de la réalité virtuelle, permet d'aider le chirurgien à la

préparation d'une opération ou à la formation des étudiants aux gestes chirurgicaux (Figure 1.10).

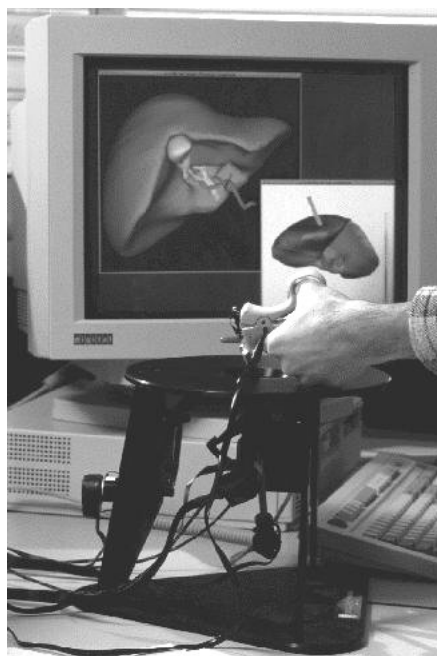


Figure 1.10 : Simulation chirurgicale avec retour d'effort, copyright photo INRIA Sophia Antipolis

Des premiers essais ont été réalisés dans le secteur de la psychothérapie virtuelle : par exemple, le diagnostic et la thérapie de troubles de comportement alimentaire d'un patient immergé dans un environnement virtuel adéquat, [Jullien, 1999]. La télémédecine, concernant l'échange d'informations médicales, dépend avant tout du développement d'Internet et peu des techniques de la réalité virtuelle. Par contre, la téléchirurgie peut exploiter ces techniques : le chirurgien peut insérer ses mini-outils robotisés par une petite ouverture et téléopère ses actions en agissant sur un modèle virtuel. Le système robotique, asservi aux commandes du chirurgien, agit sur le corps à opérer.

- En téléopération avec passage par le virtuel, nous retrouvons cette méthode. Au lieu d'agir à distance par l'intermédiaire d'un robot réel, le téléopérateur voit un modèle du monde réel distant et agit par l'intermédiaire d'un robot virtuel dans son environnement virtuel. Le robot virtuel et l'environnement virtuel doivent être des modèles suffisamment exacts. Les actions sur le robot virtuel sont transmises au robot réel. L'intérêt principal de la téléopération avec passage par le virtuel est de prédire sur les modèles virtuels les futures actions réelles avant d'en être informées au retour (sous réserve de l'exactitude suffisante des modèles). Les techniques de réalité augmentée sont aussi utilisées efficacement dans ce contexte.
- Les sciences cognitives : la réalité virtuelle est utilisée en tant qu'outil d'étude du comportement neurophysiologique ou psychophysique de l'homme. Il est très intéressant pour les spécialistes des sciences cognitives de réaliser des expériences dans

des environnements qu'ils peuvent parfaitement calibrer et contrôler. Ils peuvent aussi tester le comportement humain (et animal) dans des espaces soumis à des lois non naturelles pour mieux analyser et comprendre les réactions humaines. Ils peuvent également analyser le comportement humain dans un environnement virtuel en tant que tel, pour aider au développement de la recherche en réalité virtuelle. Le Laboratoire de Physiologie de la Perception et de l'Action du Collège de France a, par exemple, réalisé des expériences de navigation virtuelle au sol et en vol spatial [http17].

- Les activités militaires : dans ce domaine, la simulation, spécialement sur engins militaires, a été développée depuis plusieurs dizaines d'années. Les techniques de réalité virtuelle permettent maintenant d'autres simulations, comme l'immersion de soldat sur le champ de bataille ou le tir de missile pour l'entraînement [Jolivat, 1994]. Les applications actuelles intéressent cinq domaines, la conception des matériels, les mises au point et les essais dynamiques des systèmes avant ou en parallèle de leur construction, la formation et l'entraînement du personnel, la préparation de missions spécifiques et la réalisation de missions réelles [Papin, 1997]. Pour illustrer ces applications voici un exemple concernant l'utilisation de la réalité virtuelle et de la réalité augmentée pour la conduite en temps réel de véhicules. Cette conduite peut se faire en étant à bord du véhicule ou en situation de télé conduite. Il s'agit de conduire un véhicule sans vue directe de l'extérieur. Aussi, le véhicule dispose d'une série d'écrans dans son habitacle et d'un système de repérage GPS de très haute précision de l'ordre de 10 cm (Figure 1.11). Le circuit routier de l'établissement technique d'Angers a été entièrement numérisé avec une précision de 7 cm pour les routes et les chemins. La précision est de plusieurs dizaines de mètres pour les zones où ne vont pas les véhicules. Le conducteur peut visualiser l'extérieur de trois manières différentes. Dans la première, il n'a devant lui que l'image virtuelle de la route calculée par rapport à la position du véhicule donnée par le GPS. Cette image correspond au paysage qu'il verrait en conduite normale à travers le pare-brise. Dans la deuxième, il voit le paysage comme s'il était derrière le véhicule (conduite en œil de demi-Dieu). Dans la troisième, il conduit à partir d'une image vidéo de la route. Sur cette image, il est possible d'incruster des informations comme le balisage de la route par exemple. Il s'agit alors d'une conduite en environnement augmenté. Les performances actuelles permettent de conduire le véhicule à une vitesse de 60 km/h sur le circuit routier réel.
- Les activités commerciales : les techniques de réalité virtuelle permettent d'étudier l'aménagement de magasin ou le comportement d'achat des consommateurs (exemple au chapitre suivant). Elles peuvent être aussi un complément au développement du commerce électronique, en proposant une interaction et une immersion plus efficaces.
- L'architecture : ce secteur devrait être économiquement important pour la conception et pour la visite de bâtiment par les futurs habitants. Les dispositifs RV vont se développer dans ce secteur car les besoins et les enjeux financiers sont importants. Par exemple, beaucoup de futurs propriétaires de maison individuelle souhaiteraient les visiter virtuellement avant de les faire construire, pour ne pas se fier uniquement à des plans d'architecture qu'ils maîtrisent mal. Il faut noter que la représentation des dimensions d'un bâtiment à l'échelle 1 est un critère important pour se faire une représentation correcte, surtout volumique et esthétique. Le grand champ de vision est lui aussi indispensable. L'outil le plus approprié, mais aussi le plus onéreux, est la salle immersive qui permet à plusieurs participants de déambuler à l'extérieur et à l'intérieur de tout site architectural. Les grands projets architecturaux sont de plus en plus présentés à partir de modèles numériques, souvent exigés par les clients. Ceci doit inciter les architectes à exploiter les potentialités de la réalité virtuelle. Ils

peuvent en tirer trois fonctionnalités pour leur métier. Les «3 D», décrits par Guy Lebras, directeur de la technopôle de Laval : le Design du bâtiment, les Décisions à prendre et la Démonstration aux clients du futur projet.



Figure 1.11 : La conduite réelle avec les techniques de réalité virtuelle et augmentée, photo J.P. Papin

1.2.3.3 Domaines industriels

Avant d'expliquer les liens entre la Conception Assistée par Ordinateur (CAO) et la réalité virtuelle, nous pouvons citer les principaux autres secteurs industriels qui exploitent la réalité virtuelle. Car celle-ci est un outil qui peut être utilisé à différentes phases de la vie d'un produit ou d'une installation :

- Le marketing : ses services ainsi que ceux de ventes et d'après ventes peuvent utiliser la réalité virtuelle pour permettre aux clients d'être mieux informés sur le produit qui est ou sera fabriqué. Par exemple, la présentation virtuelle de la voiture Picasso pour sa vente chez des concessionnaires, avec immersion par visiocasque ou par écran orientable et tactile (Figure 1.12).
- La maintenance et le dépannage : la formation virtuelle sur un produit ou sur une installation, et plus simplement l'information technique de ceux-ci, sont des outils pertinents pour la maintenance et le dépannage. Comme nous l'avons déjà indiqué, une (in)formation en environnement virtuel permet de faire des tâches sans danger, de réaliser des scénarios impossibles à reproduire dans la réalité (accidents, incidents techniques, etc.) et de représenter plus explicitement des mécanismes, des processus ou des structures.
- Les techniques de la réalité augmentée peuvent être aussi exploitées pour la formation et l'information (maintenance, dépannage, documentation technique,...) à partir des vues des objets réels. Elles permettent de même de faire de la téléopération et de la modélisation géométrique sur site d'objets réels [Maman, 1998]. Dans ce dernier cas, il est important de retrouver les modèles «tels que construits» pour, par exemple, la modification ou la rénovation d'une installation. C'est la technique dénommée «Reverse engineering».

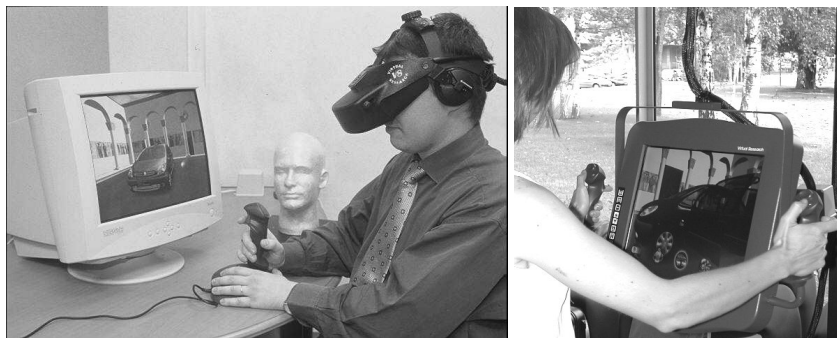


Figure 1.12 : Présentation virtuelle d'une voiture pour la vente, réalisée par PSA, copyright Photo PSA Peugeot Citroën

1.2.3.4 Conception Assistée par Ordinateur et réalité virtuelle

Toute application industrielle, faisant intervenir l'homme, est susceptible d'exploiter la réalité virtuelle. Tout produit, à usage de l'homme, peut être étudié sous l'angle de cette technique innovante, que ce soit pour son aspect technique, fonctionnel ou esthétique. La Réalité Virtuelle complète la Conception Assistée par Ordinateur «classique» permettant à l'ingénieur une meilleure étude de ses produits, [Fuchs, 1998].

Nous pouvons schématiser la «**CAO classique**», très simplifiée, avec la Figure 1.13. Elle présente la modélisation géométrique qui se déroule pendant la phase de conception du produit à partir d'un modèle fonctionnel. Avant de lancer la production, différentes modélisations physiques (mécaniques, thermiques, acoustique, électronique,...) sont réalisées pour valider la conception ou pour modifier le produit en conséquence. Le temps de calculs dans cette étape n'a pas trop d'importance (dans une certaine limite), seules la précision et la validité des calculs sont importantes. Par rebouclage (feedback), la conception est affinée jusqu'à un certain stade, avant de passer à la réalisation matérielle d'un prototype qui doit confirmer la validité de la conception.

Dans cette démarche, seules des caractéristiques techniques et fonctionnelles dépendant de lois physiques peuvent être testées. Les caractéristiques techniques et fonctionnelles dépendant d'un comportement humain comme les aspects ergonomique et esthétique du produit ne sont pas pris en compte ou le sont difficilement. Par exemple, il est difficile pour des ergonomes d'étudier l'ergonomie de l'intérieur d'une voiture en se basant sur des plans ou sur des images de synthèse affichées sur un simple écran d'ordinateur. Une vision immersive est plus adaptée à leur problématique et leur permet de tester et de modifier la conception de l'intérieur de la voiture.

Les contraintes techniques et économiques de la réalité virtuelle ne permettent pas d'envisager tous types de simulation avec l'homme incorporé dans la maquette virtuelle. Mais une bonne connaissance des potentiels et des limites, tant techniques que fonctionnels, des simulations envisageables doit permettre aux industriels d'améliorer leur outil de conception (virtuelle). Le schéma suivant (Figure 1.14) présente les nouvelles possibilités par rapport au schéma de base de la CAO.

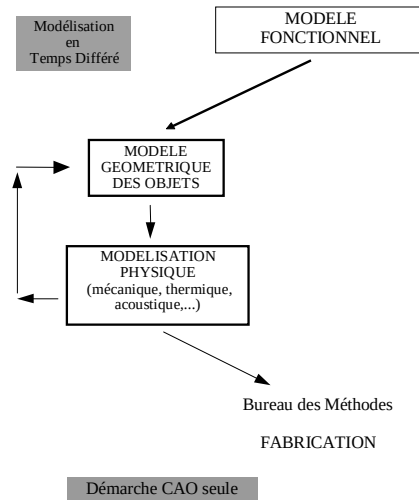


Figure 1.13 : Schéma de la «CAO classique»

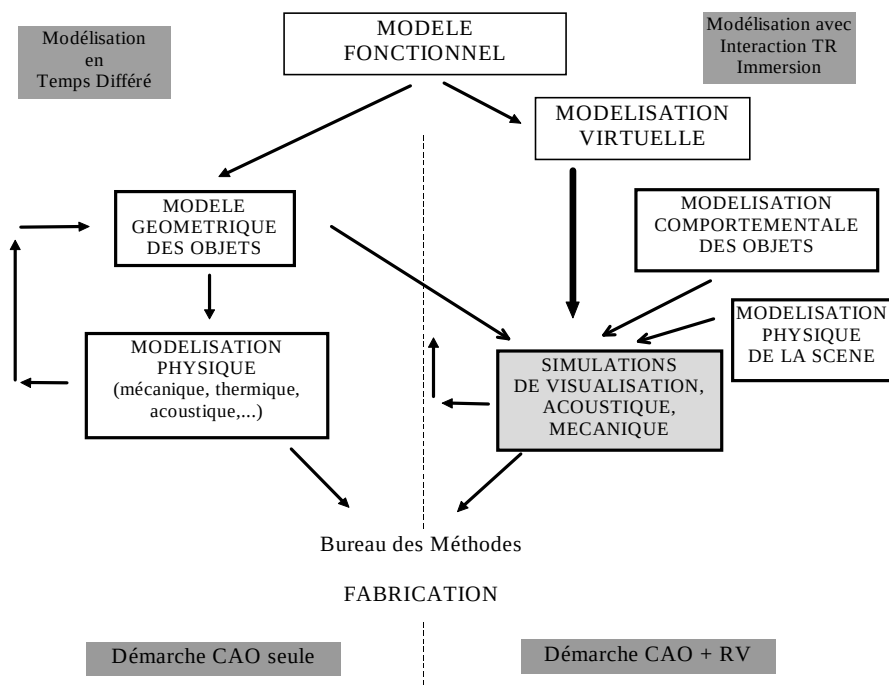


Figure 1.14 : Schéma de la «CARV»

Pour réaliser la modélisation virtuelle, il faut rajouter différentes modélisations à celle géométrique des objets :

- modélisation du comportement des objets (cinématique, dynamique, optique¹², acoustique, fonctionnel,...);
- modélisation de la scène où se situent les objets (par exemple pour une voiture, la scène routière).

Toutes ces modélisations devront être compatibles avec un fonctionnement en temps réel, ce qui est encore un verrou technologique pour le développement de la CAO avec les fonctionnalités de la réalité virtuelle. C'est un nouveau challenge difficile mais prometteur. Nous le nommons par ce nouvel acronyme : **CARV (Conception assistée par Réalité Virtuelle)**.

Une des grandes autres difficultés de la CARV est de pouvoir récupérer efficacement et facilement les données utiles¹³ des modèles géométriques CAO. Et si les concepteurs modifient le modèle en CARV, ils souhaitent pouvoir «remonter» ces modifications dans le modèle CAO. Actuellement, ces deux difficultés ne sont pas encore raisonnablement bien résolues, des travaux de recherche sont à poursuivre. Toute cette problématique, passage des modèles de la CAO à la CARV et réciproquement, est traitée dans la partie «Modélisation des mondes virtuels».

Différents termes pour parler des modèles dans les domaines de la CAO et de la CARV sont employés. Nous ferons la différenciation suivante : dans le cas de la CAO «classique», le concepteur utilise la **maquette numérique** qui contient les informations de conception et de fabrication mécanique. La **maquette virtuelle** est une représentation visuelle «réaliste» du produit conçu mis dans son environnement fonctionnel. L'utilisateur peut ainsi naviguer autour et à l'intérieur de l'objet, du système ou de l'installation. Mais à part cette possibilité de navigation, l'utilisateur est passif face à cette maquette virtuelle. On est entièrement dans le domaine de la CARV avec le **prototype virtuel** qui est une représentation sur laquelle l'utilisateur peut interagir. Ce dernier peut aussi s'immerger dans cette représentation dans le cas d'une installation ou d'un dispositif virtuels. L'utilisateur est donc actif : il peut, par exemple, manipuler l'objet ou un de ses sous-ensembles, tester l'ergonomie ou les fonctionnalités du dispositif. Il y a donc dans ce dernier cas une modélisation physique de la scène et une modélisation comportementale des objets qui permettent une simulation en temps réel.

Quel est l'objectif principal de la CARV : sans aller jusqu'à penser qu'en conception toute maquette physique (réelle) d'un produit sera supprimée, l'objectif est d'en diminuer le nombre et(ou) d'améliorer la qualité de la conception. C'est le principal but de certains industriels, dont les constructeurs d'automobile (chez un constructeur automobile, le nombre de maquettes physiques a été divisé par 3 à 5 pour la conception du style du véhicule). A notre avis (mais sans pouvoir le démontrer!), la maquette virtuelle et le prototype virtuel vont plus facilement montrer la non-validité d'une solution et plus difficilement prouver que la solution testée virtuellement convient. Le passage par la maquette réelle restera dans bien des industries incontournable. En résumé, dans le cycle en V de la vie d'un produit (de sa conception à sa fin d'utilisation), pendant l'étape de conception la CARV permet de court-circuiter une partie du cycle en V (sans la fabrication du produit) pour tester certaines spécifications sur le prototype virtuel, d'où le cycle en « ∇ ».

¹² Si on veut calculer des images «photoréalistes» des objets, il faut connaître les caractéristiques optiques de leur surfaces, qui ne sont pas en général incorporées dans les modèles CAO.

¹³ Toutes les données ne sont pas forcément utiles en CARV : tolérancement, cotation,...

Cette évolution dans le travail de conception grâce à la CARV rentre dans le cadre de la démarche **d'ingénierie concourante** : si on veut associer simultanément des concepteurs de plusieurs métiers, il faut qu'ils puissent dialoguer non seulement grâce une maquette numérique ou une maquette virtuelle, mais aussi grâce à un prototype virtuel. Par exemple, les stylistes et les ergonomes des constructeurs d'automobile vont pouvoir concevoir ensemble et non séquentiellement comme auparavant, avec toutes les contraintes que cela entraînait (perte de temps, aller et retour répété entre les services, réflexion séparée, incompréhension). Mais cela demande une approche différente, un changement culturel dans la façon de travailler. Il y a eu le même problème quand les ingénieurs ont dû passer de la conception 2D à la conception 3D. Le changement ne s'est pas fait sans remise en cause des habitudes de travail et de pensée. D'une façon plus générale, quel que soit le milieu professionnel (industrie, médecine, commercial,...), la technologie propose maintenant des modèles 3D et non 2D (du produit industriel, de l'anatomie du corps humain, du produit commercialisé,...). Les professionnels vont-ils exploiter les modèles 3D aisément ? Et vont-ils ensuite faire appel aux techniques de réalité virtuelle ?

1.2.3.5 Les différentes applications en CARV

Fonctionnalités testées par CARV

La conception d'une pièce, d'un mécanisme ou d'une installation peut exploiter les potentialités de la réalité virtuelle pour différentes fonctionnalités :

Conception ergonomique :

La conception d'un produit peut utiliser la réalité virtuelle pour permettre à un opérateur de tester l'ergonomie du produit, qu'il soit de petite dimension (un appareil électronique portable, par exemple) ou de grande dimension (le poste de commande d'une machine, la signalétique d'une station de métro, par exemple). Les constructeurs de moyens de transport sont très intéressés par cette fonctionnalité. Ils souhaitent, par exemple, tester l'accessibilité et le confort des commandes de l'intérieur d'un véhicule ou la visibilité de la route depuis la place du conducteur (Figure 1.15). Nous avons participé à une conception de dispositif RV avec visiocasque chez PSA Peugeot Citroën pour l'étude de la visibilité de la route à partir de l'intérieur du véhicule [Routier, 2000].

Pour l'aéronautique, un système spécifique d'étude ergonomique avec retour d'effort a été développé pour les cockpits d'avions. Le concepteur, CGSD [http18] a réalisé le dispositif TOPIT (Touched Objects Positioned In Time), qui a été construit pour fournir un retour d'effort et un retour tactile à l'utilisateur d'un cockpit virtuel. L'utilisateur porte un casque qui projette l'intérieur d'un cockpit, incluant le tableau de bord ainsi que la scène extérieure. Une représentation de la main de l'utilisateur est aussi rendue dans la scène. L'utilisateur peut activer différentes sortes de boutons sur le tableau de bord et ressentir précisément les forces et les textures de ces boutons. L'objectif est de fournir un simulateur qui peut être configuré entièrement par logiciel pour représenter différents cockpits.

Conception technique de l'assemblage ou(et) de la maintenance du produit :

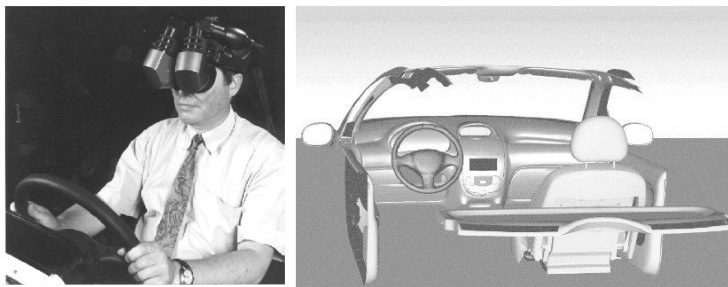


Figure 1.15 : Etude de la visibilité avec un visiocasque, copyright photo PSA Peugeot Citroën

Il s'agit de la conception d'un produit au niveau géométrique : l'étude de l'encombrement des pièces, du montage et du démontage d'un ensemble, testé virtuellement par un utilisateur, pour concevoir la maintenance et le dépannage d'un mécanisme ou d'un processus, qu'il soit simple ou complexe (machine, raffinerie, véhicule,...). Des tests virtuels d'accessibilité et d'assemblage peuvent valider ou remettre en question la conception d'un produit. A noter qu'il y a deux types de problématique aux difficultés bien différentes :

On peut vouloir tester seulement que la pièce (avec éventuellement le porte pièce ou les mains) passe entre les autres pièces et puisse être assemblée. On a dans ce cas des problèmes de détection de collisions et de contrainte de mouvement. Les gestes de l'opérateur ne sont pas le sujet d'étude.

Le deuxième cas est plus difficile à simuler : les gestes de l'opérateur doivent être les mêmes que ceux qui seront exécutés en environnement réel (dans un but d'étude ergonomique ou de formation à la maintenance). Les retours d'efforts, qui doivent dans ce cas impérativement être simulés, sont souvent très difficiles à réaliser.

Conception esthétique :

La qualité esthétique d'un produit ne pouvant être donnée que par l'homme, il est nécessaire d'immerger correctement le concepteur virtuellement si on veut diminuer le nombre de prototypes réels à construire avant fabrication. Cette fonctionnalité est très utilisée par les constructeurs d'automobile, ainsi que par les autres secteurs des transports, ceci dans le but de gagner du temps lors de la conception et de supprimer en partie la fabrication de prototypes coûteux. Exemple : la conception de l'aménagement intérieur de cabine d'avion ou de train, dont la qualité esthétique ne peut être appréciée sur la maquette virtuelle issue du modèle CAO. La perception des espaces et des volumes est difficile à juger sur un écran plat monoscopique. Pour une immersion virtuelle de qualité, on peut exploiter un visiocasque haut de gamme : l'utilisateur a le moyen de circuler virtuellement à l'intérieur de l'avion ou du train pour l'observer et pour apprécier l'aspect du futur aménagement intérieur.

Visualisation explicite de phénomènes physiques :

Si un ingénieur spécialisé en simulation peut aisément interpréter les résultats chiffrés et les courbes d'une simulation numérique, il n'en est pas de même pour une personne

non experte. La visualisation explicite d'un phénomène physique peut être utile pour valider une conception. Par exemple, la société Médialab a fourni à la DCN une installation de réalité virtuelle permettant de visualiser les mouvements d'une frégate, porte hélicoptères, sur la mer à partir d'un code de calcul de simulation de la tenue du bateau en fonction de la houle (Figure 1.16).

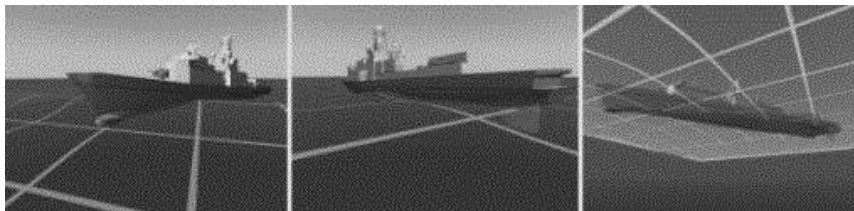


Figure 1.16 : Simulation de la tenue d'un navire sur la mer, photo Médialab

Les méthodes de conception exploitant la CARV

Les techniques de la réalité virtuelle permettent de nouvelles méthodes de travail pendant les phases de conception :

Prototypage virtuel :

Dès l'étude d'un avant-projet, la réalité virtuelle est un outil intéressant dans la démarche de conception. Elle peut ainsi éviter la réalisation de maquette physique pour abaisser les coûts de conception. En partant du modèle fonctionnel du produit, on peut en avoir une représentation virtuelle de celui-ci, sans avoir à spécifier précisément les dimensions et les caractéristiques techniques qui seront définies ensuite dans la phase CAO. En exemple, on peut citer l'étude de maquettage virtuel d'un appareil électronique, portable et utilisé en extérieur : la réalité virtuelle permet de tester l'ergonomie de son boîtier de commande (accessibilité des boutons, qualité de l'affichage sur l'écran, couleurs, facilité de compréhension par un novice,...). En même temps, elle permet de valider le programme informatique dédié à cet appareil.

Travail de conception collaboratif :

La réalité virtuelle distribuée, associant plusieurs personnes distantes sur une représentation virtuelle d'un produit, permet le travail collaboratif très utile pour les entreprises travaillant sur plusieurs sites ou sur des installations distantes. L'essor des réseaux de communication et d'Internet donne une ouverture complémentaire à ces techniques en permettant aux différents utilisateurs de travailler collectivement sur un même environnement virtuel, tout en étant délocalisés. On retrouve le principe de la téléconférence, mais ici au lieu d'échanger ensemble sur un sujet, il s'agit de concevoir ensemble sur un prototype virtuel avec lequel on peut interagir.

Suivi et revue de projet :

Pour suivre l'avancement d'un chantier ou la fabrication d'un produit, la réalité virtuelle permet de créer une maquette virtuelle (Figure 1.17) qui représente l'état d'avancement des travaux à l'instant donné. Elle sert de référence et peut aussi offrir une représentation visuelle différenciée de la maquette selon la catégorie professionnelle

de l'utilisateur : un électricien voudra voir surtout les circuits électriques, un chimiste la partie génie chimique, etc. Cette maquette dynamique peut aussi servir de support de documentation en renseignant sur l'historique des travaux, sur les modifications effectuées ou à réaliser (post-it virtuels), sur les décisions prises, etc. Les entreprises d'ingénierie, comme Technip, utilisent en ce sens la réalité virtuelle comme outil de communication entre leurs différents sites et entre leurs différents corps de métiers, qui ne « parlent » pas nécessairement le même langage.



Figure 1.17 : Revue de projet d'un véhicule, copyright photo PSA Peugeot Citröen

Nous avons cité les principales fonctionnalités et méthodes en CARV, mais la liste n'est pas exhaustive et figée dans le temps.

Pour conclure sur ce chapitre introductif de la réalité virtuelle, nous pouvons dire que les problèmes fondamentaux inhérents à la réalité virtuelle se découpent en deux catégories : ceux concernant la modélisation et la numérisation de l'environnement virtuel et ceux concernant l'interfaçage de l'utilisateur avec ce monde virtuel. C'est cette deuxième catégorie que nous allons aborder dans le prochain chapitre suivant une approche théorique et pragmatique de la réalité virtuelle.

1.3 RÉFÉRENCES BIBLIOGRAPHIQUES

[Boulanger, 1998]	P. Boulanger, J. Taylor, S.F. ElHakim, M. Rioux, « <i>How to virtualize reality : an application to the recreation of world heritage sites</i> » Proceedings of the Conference on Virtual Systems and MultiMedia. Gifu, Japan. November 1820, 1998. pp. 3945. NRC 41599.
[Cadoz, 1994]	C. Cadoz, « <i>Les Réalités Virtuelles</i> », Dominos, Flammarion, 1994.
[Coiffet, 1993]	P. Coiffet et G. Burdea, « <i>La Réalité Virtuelle</i> », éditeur Hermès.
[Ernadote, 1997]	D. Ernadote. « <i>Intégration d'Objets Virtuels dans des Scènes Réelles</i> ». Thèse de docteur de l'Ecole des Mines de Paris, 1997.
[Fuchs, 1996]	P. Fuchs, « <i>Les interfaces de la réalité virtuelle</i> », AJIIMD, 1996, ISBN 2-9509954-0-3, les Presses de l'Ecole des Mines de Paris.

- [Fuchs, 1998] P. Fuchs, F. Nashashibi, «*De la CAO à la Réalité Virtuelle*», Revue internationale de CFAO et d'informatique graphique, Vol. 13, N° 2, P 131-167, 12/1998.
- [Huriet, 1997] C. Huriet, «*Images de synthèse et monde virtuel : techniques et enjeux de société*», Office parlementaire d'évaluation des choix scientifiques et technologiques, N° 169, Sénat, France.
- [Jessel, 1999] J.P. Jessel, C. Jaspard, J.J. Flores, A. Josseau, «*Animation de synthèse et concept de réalité virtuelle pour l'art du spectacle et l'art plastique numérique*», Actes des journées réalité virtuelle et cognition, 14-15 Décembre, ENST, Paris, 41-47.
- [Jolivat, 1994] B. Jolivat, «*La simulation et ses techniques*», PUF, Que sais-je ?, N° 2941.
- [Jullien, 1999] Y. Jullien, C. Defrance, P. Nugues, «*Principes pour une réalité virtuelle à visée thérapeutique, Application aux troubles du comportement alimentaire*», Actes des journées réalité virtuelle et cognition, 14-15 Décembre, ENST, Paris, 87-96.
- [Lourdeaux, 1999] D. Lourdeaux, P. Fuchs, J. M. Burkhardt and F. Bernard, «*Relevance of an intelligent tutorial agent for virtual reality training systems*», à paraître, International Journal of Continuous Engineering and Life-Long Learning.
- [Maman, 1998] D. Maman, «*Recalage de modèles tridimensionnels sur des images réelles*». Thèse de doctorat de l'Ecole des Mines de Paris, 1998.
- [Mantavoni, 1999] G. Mantavoni and G. Riva, «*<Real Presence : How Difference Ontologies Generate Different Criteria for Presence, Telepresence, and Virtual Presence>*», revue Presence, MIT, Vol. 8, N° 5, 540-550, 1999.
- [Milgram, 1996] Milgram, P., Drascic, D. (1996). «*Perceptual issues in augmented reality*». In Proceedings of the SPIE, volume 2653 (Stereoscopic Displays and Virtual Reality Systems III), pages 123-134, San Jose (CA).
- [Papin, 1997] J.P. Papin, P. Hue, 1997, French Applications of virtual reality. In Cdrom RTO meeting MP-O58 March 2001.
- [Pimentel, 1994] K. Pimentel and K. Teixeira, «*La Réalité virtuelle...de l'autre coté du miroir*», Addison-Wesley, 1994.
- [Quéau, 1993] P. Quéau, «*Le virtuel - Vertus et Vertiges*», Champ-Vallon/INA, 1993.
- [Rheingold, 1993] H. Rheingold, «*La Réalité Virtuelle*», édition Dunod.
- [Sheridan, 1999] B. Sheridan, Descartes, Heidegger, Gibson, and God : «*Toward an Electic Ontology or Presence*», revue Presence, MIT, Vol. 8, N° 5, 551-559, 1999.
- [Verna, 1999] D. Verna, «*Définir le virtuel : une approche cognitive*», Actes des journées réalité virtuelle et cognition, 14-15 Décembre, ENST, Paris, 11-22.

Adresses des sites d'entreprises et d'institutions :

- [http1] <http://www.inria.fr/presse/>
- [http2] <http://www-sop.inria.fr/epidaure/GT-RV/gt-rv.html>
- [http3] <http://www.industrie.gouv.fr/rntl/FichesA/Perf-Rv.htm>
- [http4] <http://www.laval-virtual.org>
- [http5] <http://www.enib.fr/gt-rv8>
- [http6] <http://www.ai.cluny.ensam.fr>
- [http7] <http://www.enib.fr>

-
- [http8] <http://www.enst.fr/~revico>
[http9] <http://birp.com/micad>
[http10] <http://www.virtual-worlds.net/vw2000>
[http11] <http://lsc.cemif.univ-evry.fr:8080/dea>
[http12] <http://www-caor.ensmp.fr/>
[http13] <http://www.istia.univ-angers.fr/Innovation>
[http14] <http://www.simteam.com>
[http15] <http://www.immersion.fr>
[http16] <http://www.ubikam.com>
[http17] <http://www.college-de-france.fr/chaire3/space1-4.html>
[http18] <http://www.csgd.com/TOPIT.html>

2 APPROCHE THÉORIQUE ET PRAGMATIQUE DE LA RÉALITÉ VIRTUELLE

Philippe Fuchs

2.1 COMPORTEMENT DE L'HOMME DANS UN ENVIRONNEMENT RÉEL

Suite à la présentation de la finalité et de la définition technique de la réalité virtuelle, nous allons expliquer dans ce chapitre notre méthode pour la conception et l'évaluation de systèmes RV efficaces. Rappelons que la finalité est *de permettre à une ou à plusieurs personnes des activités **sensori-motrices** et donc mentales dans un monde artificiel*. Il est donc souhaitable d'abord de préciser le comportement de l'homme dans un monde réel, avant de penser au virtuel. *Les techniques de la réalité virtuelle sont fondées sur l'interaction en temps réel avec un monde virtuel, à l'aide d'interfaces comportementales permettant l'immersion «pseudo-naturelle» de(s) l'utilisateur(s) dans cet environnement*. Mais comment l'homme est-il immergé et comment interagit-il dans un monde réel ?

Physiquement, l'homme perçoit son environnement par l'intermédiaire de ses sens. La réception d'un stimulus sensoriel se fait par un organe récepteur spécifique (œil, peau, nez, etc.). Cette entrée sensorielle est le point de départ du transfert de l'information dans l'organisme par les voies nerveuses (Figure 2.1). Le message nerveux afférent provenant du récepteur est véhiculé vers les centres nerveux (moelle épinière, cerveau). Après intégration et traitement de l'information dans ces centres, ceux-ci transmettent le message nerveux efférent aux organes effecteurs (muscles squelettiques, muscles oculaires, muscles des cordes vocales, etc.). Ceux-ci réalisent le comportement moteur, se traduisant par des mouvements.

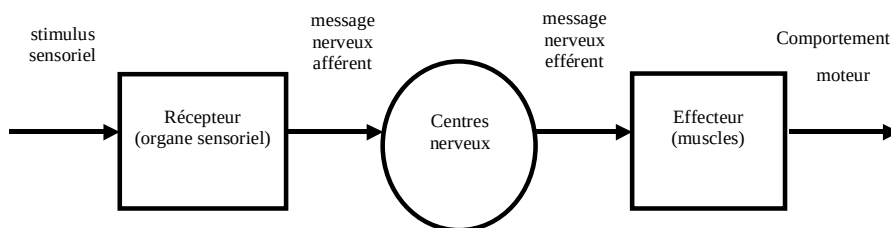


Figure 2.1 : Transfert de l'information dans l'organisme

Nous pouvons présenter le comportement de l'homme dans le monde réel d'un point de vue **anthropocentrique**, que nous utiliserons aussi pour un monde virtuel. L'homme est au centre et perçoit son environnement par ses sens (supérieur à cinq), dont nous avons représenté les principaux. D'où le schéma de la figure 2.2.

Le chapitre suivant «L'être humain dans les environnements réel et virtuel» développera les idées de base du comportement de l'homme dans ces environnements. Mais avant, nous explicitons dans les paragraphes suivants l'interfaçage comportemental et les notions d'immersion et d'interaction.

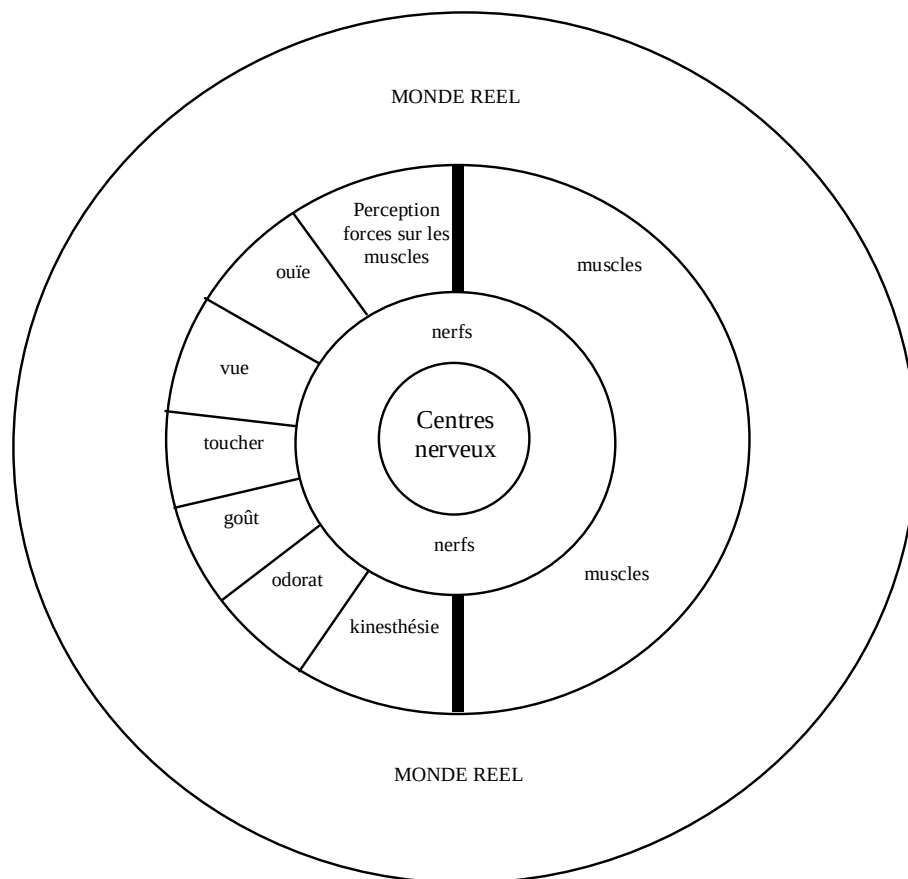


Figure 2.2 : Schéma anthropocentrique de l'homme dans le monde réel

2.2 CONCEPTION DES INTERFACES COMPORTEMENTALES

2.2.1 CONCEPTION MATÉRIELLE

Nous avons déjà expliqué que nous employons le terme d'**interfaces comportementales** plutôt que d'interfaces matérielles ou d'interfaces homme machine (IHM), pour bien différencier celles-ci de celles-là. Pour notre définition, *l'interface comportementale est un dispositif qui exploite un comportement humain*. Au niveau théorique, les «**interfaces sensorielles**» sont conçues pour transmettre des Stimuli Sensoriels (S.S.) de l'ordinateur vers l'homme et inversement les «**interfaces motrices**» pour transmettre des Réponses Motrices (R.M.) de l'homme vers l'ordinateur. Certaines interfaces sensori-motrices transmettent des R.M. et, en réaction, des S.S. sont renvoyés par l'ordinateur.

Comment les interfaces doivent-elles transmettre les stimuli sensoriels et les réponses motrices ? Comment les utiliser pour permettre un interfaçage comportemental efficace ?

La conception matérielle d'une interface comportementale passe en premier lieu par le choix du phénomène physique qui va être exploité. Les critères de choix sont similaires à ceux de tout capteur : faisabilité technico-économique et caractéristiques métrologiques performantes (précision, répétabilité, résolution, bande passante, temps de réponse, etc.). Les caractéristiques métrologiques sont performantes si elles correspondent aux capacités maximales des sens ou des réponses motrices de l'homme. Dans chaque chapitre sur les différentes interfaces comportementales, nous préciserons les correspondances, ou non, entre les capacités humaines et les caractéristiques techniques des interfaces. Nous pouvons déjà indiquer, hélas, que presque toutes les interfaces RV ont des caractéristiques métrologiques bien inférieures aux capacités sensori-motrices maximales humaines. Nous pouvons schématiser la transmission des signaux de l'ordinateur vers le cerveau et réciproquement (Figure 2.3 et Figure 2.4).

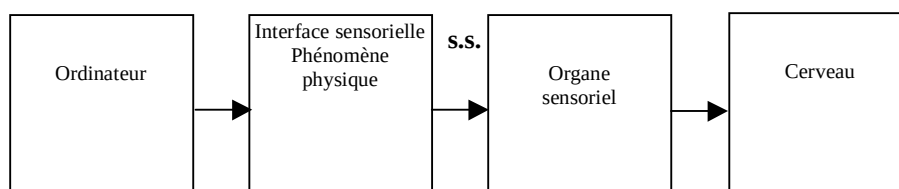


Figure 2.3 : Interfaçage sensoriel

Il faut donc concevoir de telles interfaces en se basant sur un phénomène physique. Il y a le cas particulier des biocapteurs, interfaces enregistrant l'activité nerveuse efférente. Des laboratoires proposent des dispositifs mesurant de faibles différences de potentiel à la surface de la peau. Ces biocapteurs permettent de détecter directement le signal biologique commandant le mouvement des muscles oculaires ou squelettiques. Ces interfaces ne peuvent toutefois être utilisées que par des spécialistes du domaine médical, présentées au chapitre sur les interfaces spécifiques de localisation corporelle.

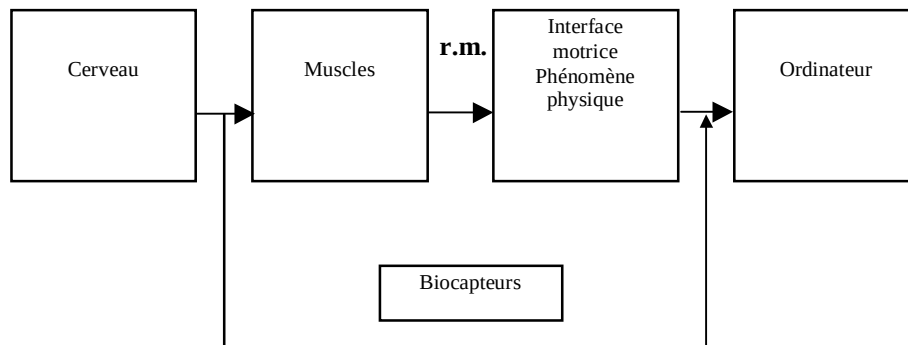


Figure 2.4 : Interfaçage moteur

2.2.2 TRANSPARENCE D'UNE INTERFACE

Un point important pour la conception d'une interface comportementale est sa **transparence** : c'est la capacité de l'interface à devenir transparente (non perçue) par l'utilisateur au cours de son utilisation. Cette dernière notion de transparence, importante en réalité virtuelle, s'étudie sur plusieurs niveaux. Au niveau physique, l'interface est transparente si celle-ci ne gêne pas, ou peu, physiquement l'utilisateur, en particulier dans ses mouvements. Pour atteindre ce but, les concepteurs d'interfaces matérielles exploitent, si possible, des phénomènes physiques qui permettent aux S.S. ou aux R.M. d'être transmises entre l'homme et la machine **sans support matériel**. Cette condition à rechercher oriente souvent la conception des interfaces vers des phénomènes physiques adéquats. Par exemple, on note que la plupart des capteurs de localisation exploitent des principes physiques sans support matériel (lumière, champs électromagnétiques, ondes acoustiques, etc.). Sans support matériel, l'interface peut être plus éloignée de l'utilisateur pour plus de liberté de mouvement : par exemple pour les S.S. de la vue, l'écran peut être placé à quelques centimètres (visiocasque), à quelques mètres (écran d'ordinateur) ou à quelques décimètres (écran en salle). Deux catégories d'interfaces matérielles apparaissent donc :

- les interfaces dont la transmission est **sans support matériel** (sans s.m.) ;
- les interfaces dont la transmission est **avec support matériel** (avec s.m.).

Ces deux catégories dépendent de l'organe sensoriel concerné, par exemple :

- interface visuelle (S.S.) : ondes lumineuses captées par les yeux (sans s.m.) ;
- interface auditive (S.S.) : sons et parole (synthèse vocale) transmis aux oreilles par les ondes acoustiques (sans s.m.) ;
- interface à retour tactile (S.S.) : transmission mécanique sur la peau (avec s.m.) ;
- interface proprioceptive (retour d'effort et simulation de mouvement) (S.S.) : transmission mécanique sur le corps (avec s.m.) ;
- interface de localisation (R.M.) : mouvements corporels détectés par ondes électromagnétiques, acoustiques (sans s.m.) ou mécaniquement (avec s.m.).

Nous en déduisons des difficultés techniques importantes pour réaliser les interfaces à retour tactile, à retour d'effort et à simulation de mouvement du corps qui doivent obligatoirement transmettre les S.S. avec des supports matériels. Mais la notion de transparence doit être aussi analysée au niveau psychologique. Dans ce cas, cela ne dépend pas de la conception physique de l'interface comportementale mais de son exploitation effective dans l'environnement virtuel, ce que nous étudierons dans ce chapitre.

2.2.3 INTERFACES COMMERCIALISÉES ET INTERFACES SPÉCIFIQUES

Depuis l'essor de la réalité virtuelle, des interfaces ont été conçues et commercialisées par des entreprises spécialisées du domaine (visiocasque, gant de données, etc.). Mais, comme nous allons l'expliquer par la suite, il faut parfois réaliser soi-même une interface spécifique en fonction de l'objectif souhaité d'immersion et(ou) d'interaction (par exemples : un chariot instrumenté pour se déplacer dans un magasin, habitacle d'une voiture, voir plus loin). A l'opposé, certaines interfaces sont universelles, non conçues spécifiquement pour la réalité virtuelle (écran, vidéoprojecteur, microphone, haut-parleur). Tout concepteur de dispositif RV ne doit donc pas écarter l'idée de concevoir et de réaliser soi-même les interfaces pour ses besoins propres, ce qui est souvent un gage de réussite si les besoins spécifiques sont bien analysés en terme d'immersion et d'interaction. Bien sûr, le coût de conception et de réalisation peut s'en ressentir, mais si le succès est au bout... A mon avis personnel, certains concepteurs de dispositifs RV auraient dû exploiter cette voie d'interface spécifique (nous ne citerons pas d'exemple!). Une autre problématique dans la réalisation d'une interface est la part du réel et celle du virtuel, puisqu'elle est «à l'interface» de ces deux mondes : dans le cas de l'apprentissage de la conduite d'un avion ou d'une voiture, les concepteurs du simulateur n'ont pas essayé de créer virtuellement les sensations du toucher, vu les difficultés que l'on vient d'explicitier précédemment. Ils préfèrent associer le monde réel (la cabine de l'avion) avec un monde virtuel (images et mouvements). Ceci donne une interface très spécifique et de qualité, même si tout n'est pas virtuel ! Par contre, s'il s'agit d'étudier l'aménagement ou l'esthétique de l'habitacle d'une voiture, il faut travailler plus en virtuel. Mais où sera précisément la frontière ? (le volant peut être réel ou virtuel, par exemple). Comme anecdote à ce propos du positionnement réel-virtuel, nous avons été étonnés d'entendre des personnes nous dire : «c'est dommage que le chariot, pour se déplacer dans un magasin virtuel, soit réel». Pourquoi à tout prix, faire tout en virtuel ? Pourquoi faire compliqué quand on peut faire simple ? Il faut donc bien analyser ce que l'on souhaite réaliser.

A noter aussi, qu'un seul constructeur, IBM, a proposé dans le passé un ensemble complet et figé de réalité virtuelle : ordinateur PC, logiciels et deux interfaces comportementales : un visiocasque et une manette avec capteur de localisation. Nous ne sommes pas surpris du manque de succès de ce produit, qui n'est plus commercialisé à ma connaissance.

Dans l'absolu, on peut schématiser (Figure 2.5) de la façon suivante l'utilisateur totalement immergé dans un monde artificiel, comparativement à l'immersion dans un monde réel.

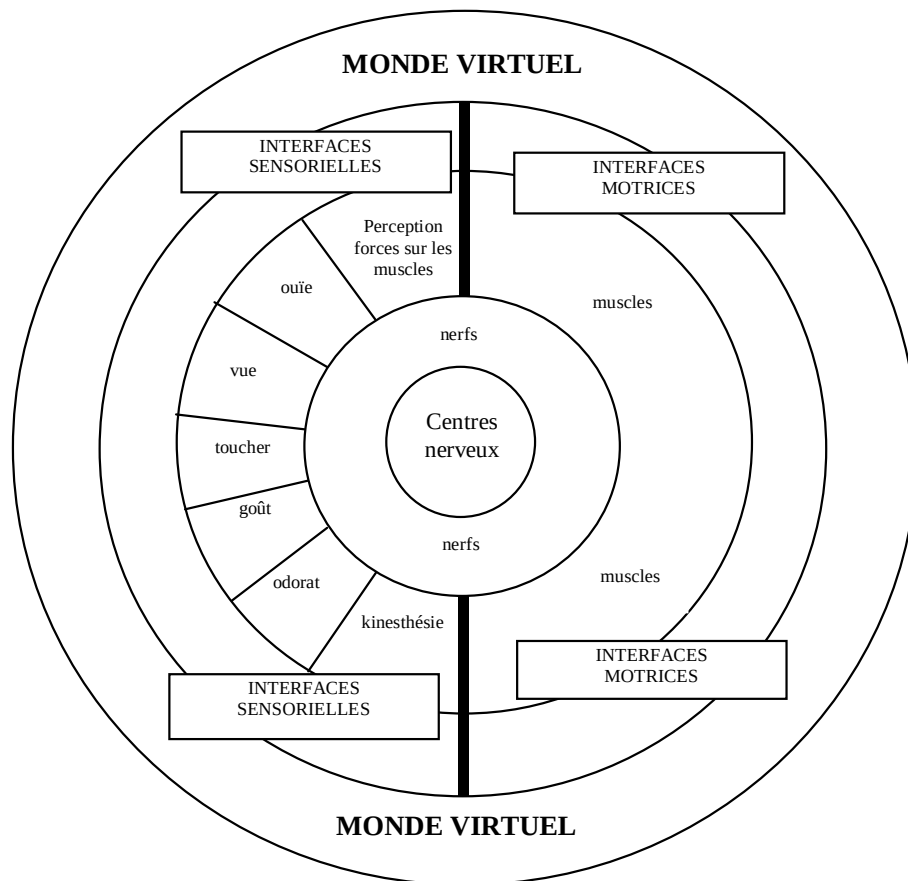


Figure 2.5 : Schéma anthropocentrique de l'homme dans un monde virtuel

2.2.4 COMPARAISON ENTRE INTERFACES EN RÉALITÉ VIRTUELLE ET EN TÉLÉOPÉRATION

La téléopération a pour objectif de permettre à un opérateur d'agir à distance dans une zone inaccessible. Les informations (les commandes) vont dans le sens (opérateur => zone inaccessible) et aussi dans le sens inverse pour informer l'opérateur. Nous présenterons dans ce document les interfaces matérielles de la réalité virtuelle. Mais il faut noter que certaines sont identiques à celles employées en téléopération, lorsque l'on désire que le téléopérateur soit immergé et interagisse «naturellement». Ceci implique un lien étroit entre les évolutions des interfaces de ces deux secteurs. Si certaines interfaces matérielles ont été conçues en premier lieu pour la réalité virtuelle, d'autres l'ont été d'abord pour la téléopération. Le schéma suivant précise la différence d'architecture entre les systèmes de réalité virtuelle et de téléopération (Figure 2.6).

2.3 L'APPROCHE «INSTRUMENTALE» POUR L'IMMERSION ET L'INTERACTION

2.3.1 LES CONCEPTS FONDAMENTAUX POUR L'INTERFAÇAGE COMPORTEMENTAL

Dans tout dispositif en réalité virtuelle, l'homme est au centre du système car l'application virtuelle lui est adressée. Partant de ce constat, nous avons une approche anthropocentrique (dirigée vers l'homme) et non technocentrique (dirigée vers l'ordinateur). L'homme va agir par ses muscles et percevoir le monde virtuel par ses sens qui, rappelons le, sont supérieurs à cinq, si leur dénombrement a une valeur ! Comme l'a écrit le Professeur Alain Berthoz, « *en établir la liste a-t-il encore un sens ?* » [Berthoz, 1998]. Le sens kinesthésique, souvent ignoré, pose bien des problèmes sur les simulateurs de transport statiques et sur des dispositifs RV. Le domaine sensoriel de la proprioception n'est surtout pas à ignorer quand on travaille en réalité virtuelle. Toutes les perceptions sensorielles sont au moins à connaître et nous devons étudier leur impact pour toute application RV, même si elles ne seront pas toutes simulées. Ceci conduit souvent à des incohérences sensorielles ou sensori-motrices qu'il ne faut surtout pas sous-estimer. Nous discuterons plus loin de ces incohérences en réalité virtuelle. Nous n'hésitons pas à représenter le **schéma anthropocentrique** (Figure 2.7) de l'immersion de l'homme dans le monde virtuel, tel que celui-ci doit le percevoir en tant **qu'utilisateur**.

Mais pour le **concepteur**, ce schéma «utilisateur» est un objectif à garder « en tête » qu'il faut éclater et compléter car il est nécessaire d'analyser finement le processus d'interfaçage et les dispositifs à concevoir. Nous avons fait le choix fondamental d'analyser ce processus à trois niveaux, la pertinence de ce choix apparaîtra au lecteur au cours du chapitre. Au premier niveau on peut schématiser l'interfaçage entre l'homme et le monde virtuel *au niveau physique*. Nous parlons dans ce cas **d'immersion et d'interaction sensori-motrices**, puisque *physiquement* l'ordinateur est connecté au corps de l'utilisateur au niveau de ses sens et de ses réponses motrices. Pourquoi employons nous pour les trois niveaux les termes «Immersion ET interaction»? Fondamentalement, comme nous l'avons indiqué dans notre définition de la réalité virtuelle, celle-ci est basée sur l'Immersion d'une personne dans un monde virtuel. Mais nous y associons le terme Interaction, car parler seulement d'immersion serait omettre l'innovation de la

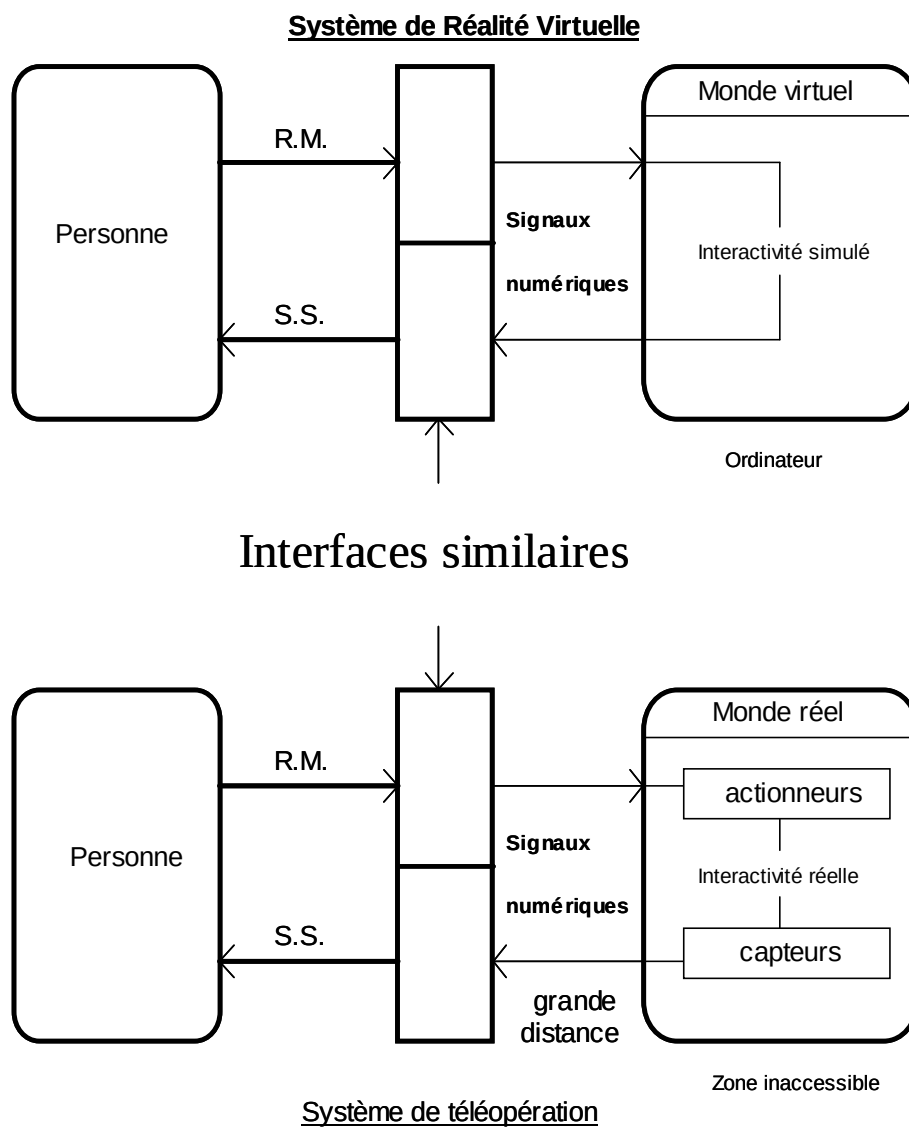


Figure 2.6 : Différence d'architecture entre les systèmes de réalité virtuelle et de téléopération

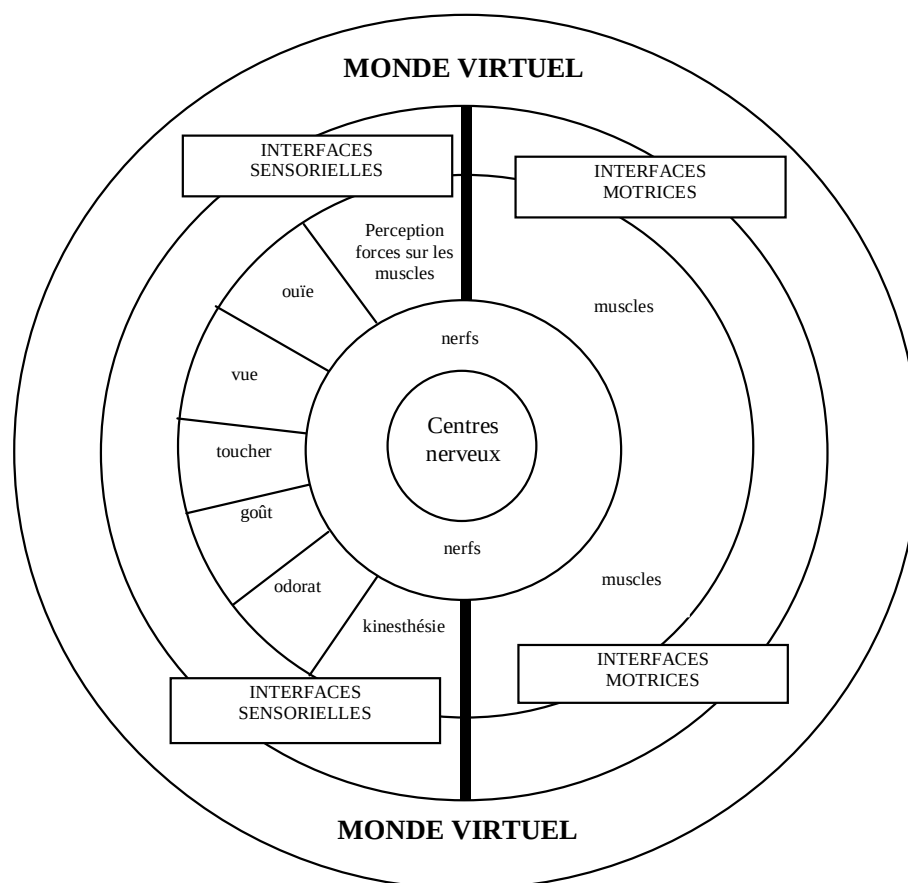


Figure 2.7 : Schéma anthropocentrique de l'homme dans un monde virtuel

réalité virtuelle, qui est de permettre au sujet *d'agir* dans un monde virtuel¹. D'où, au niveau physique d'immersion et d'interaction sensori-motrices, le **schéma technocentrique** (Figure 2.8).

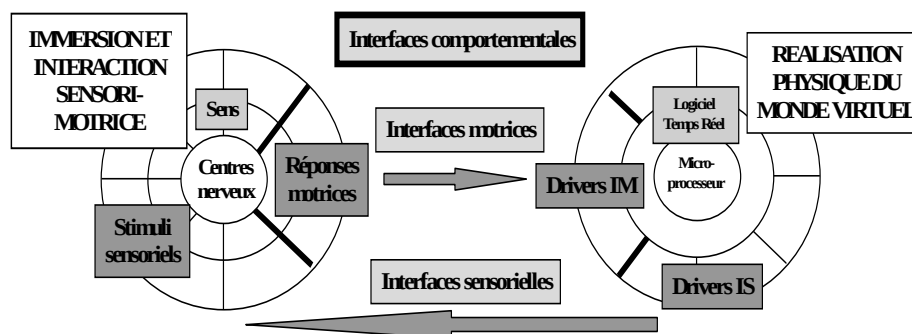


Figure 2.8 : Schéma technocentrique d'Immersion et d'Interaction sensori-motrices

Mais ce schéma est restrictif car il ne représente que partiellement les problèmes et les solutions que doit étudier tout concepteur. C'est une approche limitée, que l'on voit utilisée par certains concepteurs trop orientés sur les aspects techniques de la réalité virtuelle. Si c'est grâce à l'évolution des techniques que la réalité virtuelle peut exister, il ne faut pas que les problèmes techniques soient les seuls abordés et occultent les autres.

L'interfaçage comportemental pose une problématique de l'interfaçage proche de celle d'un opérateur avec sa machine ou avec son outil. Dans ce cas, il ne s'agit pas seulement de penser l'interfaçage au niveau physique (boutons de commande et retours d'informations sur les actions), mais aussi de comprendre à partir de quels modèles mentaux la personne va penser et agir. Comme l'a décrit Rabardel dans son **approche instrumentale** de l'interfaçage [Rabardel, 1995], il est souhaitable de présenter l'interfaçage dans un schéma triadique dont nous proposons l'équivalent en réalité virtuelle pour l'interfaçage comportemental (Figure 2.9) : il est schématisé par deux types de liaisons physiques de la personne à l'interface matérielle et de cette dernière à l'ordinateur. Mais l'utilisateur doit avoir une activité (un comportement) dans le monde virtuel, l'interface devant lui être transparente. La liaison directe du schéma représente cette activité désirée (motricité et perception désirées). Concrètement, cette activité est réalisée par une motricité et une perception effectives entre l'homme et des interfaces comportementales. Celles-ci reposent matériellement sur des artefacts (ou instruments) et sont utilisées suivant des modèles mentaux de l'utilisateur. Nous expliciterons au prochain paragraphe l'intérêt de faire appel à la notion de schème, comme le propose Rabardel dans un cadre similaire.

Pour faciliter l'utilisation matérielle ou mentale des interfaces comportementales, il ne faut pas perdre de vue que, dans notre cas², l'ordinateur peut aider l'utilisateur à employer efficacement ces dispositifs. Il y a bien des possibilités exploitables et program-

¹Et comme nous l'avons déjà signalé, tous les définitions de la réalité virtuelle se rejoignent sur ces deux seules fonctionnalités.

²De même dans le cas des IHM en informatique, mais pas dans le cas d'interfaçage avec un appareil non informatisé, dépourvu d'intelligence artificielle.

mables, dont nous exposerons certaines. Par exemple, on peut rajouter des contraintes spécifiques aux mouvements des objets pour aider l'utilisateur à obtenir la motricité désirée (la face d'appui d'un objet proche d'une table se mettra parallèlement à la table quand elle sera proche de sa surface) [Maman, 1998]. Nous regroupons toutes ces aides d'utilisation des interfaces sous le terme «**Aides Logicielles Comportementales**» (ALC). Il y a par conséquent à concevoir et à programmer le couplage entre l'interface et le monde virtuel programmée, d'où les termes de motricité et de perception programmées dans le schéma suivant (Figure 2.9).

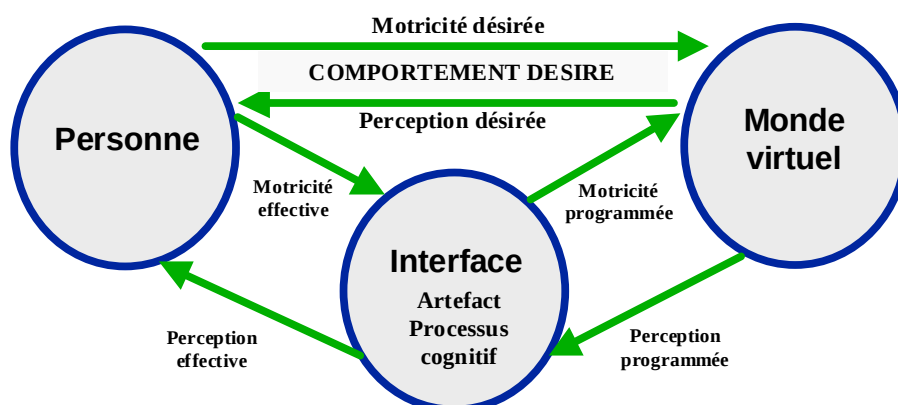


Figure 2.9 : Schéma triadique de l'interfaçage en réalité virtuelle

La perception effective est une perception sensorielle imposée à l'utilisateur, qui est le plus possible naturelle³, mais aussi partielle et plus ou moins cohérente. En exemple, la perception du mouvement fait intervenir en perception naturelle : la vision, la sensibilité aux efforts, l'appareil vestibulaire (canaux semi-circulaires, otolithes) et le toucher. Suivant l'interface comportementale choisie, certains types de récepteurs sensoriels ne seront pas stimulés, d'autres le seront avec des caractéristiques différentes de celles détectables par l'homme. L'ensemble des stimulations effectives pourront donner des incohérences que le cerveau devra gérer sous peine d'échec. Ce problème est similaires pour la motricité effective. Ces difficultés dues aux incohérences sont inhérentes aux techniques de la réalité virtuelle et seront toujours complexes à résoudre dans bien des cas. Le cerveau peut-il s'y adapter ou va-t-il les refuser ? Les expérimentations et les connaissances des spécialistes des sciences cognitives peuvent nous aider à valider ou non nos conceptions.

2.3.2 INTERFACES COMPORTEMENTALES, SCHÈMES ET MÉTAPHORES

2.3.2.1 Notion de schème

Nous avons précisé que nous recherchons à proposer à l'utilisateur une interaction et une immersion le plus possible naturelles, pour que celles-ci soient efficaces par rapport à l'application. Nous parlerons dans ce cas d'immersion **pseudo-naturelle** (idem pour l'interaction). Qu'est ce que cette notion recouvre ? En premier, nous pouvons affirmer

³Nous préciserons au paragraphe suivant cette notion.

que ce qui se fait naturellement se fait inconsciemment et donc sans (ou avec peu) effort mental. Prenons un exemple : dans un magasin virtuel, le consommateur souhaite se déplacer entre les allées. Nous lui proposons comme interface comportementale un chariot réel qu'il va pousser et orienter face à un écran. Dans ce cas, le consommateur va inconsciemment et naturellement exploiter un des automatismes qu'il a acquis lors de ses achats en magasin réel pour les exploiter en virtuel avec quelques différences sensori-motrices. D'où le préfixe « pseudo » est employé pour ce type d'immersion et d'interaction.

Il fait donc appel à un **schème sensori-moteur d'utilisation** qu'il a assimilé. La notion de schème est celle proposée par le psychologue Piaget [Piaget, 1979]. Pour lui, qui a analysé la naissance de l'intelligence de l'enfant, en particulier dans sa dimension sensori-motrice, les schèmes constituent des moyens du sujet à l'aide desquels il peut assimiler les situations et les objets auxquels il est confronté. Nous proposons de faire le parallèle entre la compréhension d'un monde virtuel par un utilisateur et celle de notre monde réel par l'enfant. Pour Piaget, un schème est l'ensemble structuré des caractères généralisables de l'action qui permet de répéter l'action ou de l'appliquer à de nouveaux contenus (comme pour le consommateur manipulant son chariot dans des conditions voisines de celles dans un monde réel). Un schème est donc l'organisation mentale des actions telles qu'elles se transfèrent ou se généralisent lors de la répétition de cette action en des circonstances analogiques. Ils correspondent aux aspects stabilisés des actions pour des classes de situations. Les schèmes d'utilisation ont une dimension privée qui est propre à chaque individu, ils s'inscrivent dans la mémoire personnelle des sujets en tant que ressources mobilisables. Ils ont également une dimension sociale : ils sont communs à l'ensemble ou à de nombreux membres appartenant à un groupe social, une collectivité, un milieu de travail. C'est pourquoi il convient de les considérer aussi en tant que **schèmes sociaux d'utilisation** (SSU), comme ressources inscrites dans la mémoire des collectivités [Rabardel, 1995].

C'est sur ce concept que nous basons notre conception pour obtenir des interfaces vraiment comportementales, proposant une immersion et une interactivité pseudo-naturelles. L'interface comportementale est donc une entité mixte comprenant à la fois un artefact (son dispositif matériel) et un schème d'utilisation. C'est une analyse comparable à celle de l'instrument, décrite par Rabardel. Comme pour lui, il s'agit de rendre compte des usages de l'interface comportementale (ou de l'instrument) par les utilisateurs, du développement de leurs compétences et d'élaborer une conceptualisation des processus de conception de ces dispositifs, pour faciliter sa prise en compte par les concepteurs (voir le paragraphe suivant).

Notons, pour l'instant, comme l'a précisé Piaget, que :

- le schème est reproductible : il contient les conditions de reconnaissance des situations auxquelles il est applicable ;
- le schème est assimilateur : il s'applique à des situations nouvelles ;
- le schème a une finalité ;
- le schème ne constitue pas une connaissance déclarative (il est utilisé et assimilé inconsciemment).

Pour préciser ces notions, prenons l'exemple de la préhension d'un objet. Le schème associé n'est pas le déroulement particulier des mouvements et des perceptions. C'est le

canévas général qui permet de reproduire en des circonstances différentes et de donner lieu à des réalisations variées. On tend plus ou moins le bras ou on ouvre plus ou moins la main selon l'éloignement et la taille de l'objet. Quel que soit l'objet, il s'agit toujours du même schème de préhension. On fait varier ici seulement des paramètres sensori-moteurs sur un même schème.

Le schème d'utilisation ne peut être isolé. Il a une relation d'interdépendance avec l'artefact qui lui est associé dans une application donnée. Mais un même schème peut être exploité avec différents artefacts et réciproquement : pour la manipulation dans l'espace 3D, on peut utiliser avec le même schème un traqueur à six degrés de liberté ou un gant de données. Un schème d'utilisation, exploité en environnement virtuel, est souvent *biaisé* par rapport au schème en environnement réel. Les fonctions sensori-motrices sollicitées peuvent être en partie différentes de celles du monde réel. Ce ne sont pas seulement les paramètres sensori-moteurs qui varient. En exemple, pour la préhension et la manipulation des objets dans le magasin virtuel, les fonctions sensori-motrices sont modifiées : grâce à un traqueur à six degrés de liberté, l'utilisateur peut prendre l'objet virtuellement et le manipuler. Mais son regard est orienté vers l'écran, non vers ses mains, et la prise de l'objet se fait sans la fermeture réelle des mains. Mais si des biais existent, cela n'empêche pas l'utilisateur d'exploiter inconsciemment le même schème.

En conclusion, comme Piaget l'a énoncé, l'intelligence sensori-motrice parvient à résoudre un ensemble de problème d'action (atteindre des objets, etc.) en construisant un système complexe de schèmes d'assimilation et à **organiser le réel** selon un ensemble de règles spatio-temporelles et causales. D'où notre postulat fondamental en réalité virtuelle :

*Dans un environnement virtuel interactif, la personne exploite la même démarche que dans un monde réel, pour **organiser le virtuel** selon un ensemble de règles spatio-temporelles et causales.*

2.3.2.2 Utilisation de schèmes, de métaphores ou de substitution sensori-motrice

Dans le cas de difficultés techniques, économiques ou théoriques ne permettant pas d'exploiter un schème d'utilisation d'une personne, nous pouvons contourner ces difficultés en employant une « **métaphore virtuelle** ». Au lieu d'exploiter un comportement sensori-moteur et acquis de la personne, nous lui proposons, visuellement en général, une *image symbolique* de l'action ou de la perception souhaitée. Par exemple dans le magasin virtuel, nous aurions pu proposer au consommateur de valider l'achat d'un produit en cliquant simplement sur son image et ensuite sur une icône représentant une caisse. Cette action devient symbolique et elle n'est plus représentative de l'action sensori-motrice dans un magasin réel, l'immersion et l'interaction y sont moins pseudo-naturelles.

L'utilisation de métaphore peut demander plus d'efforts cognitifs si la symbolique métaphorique n'est pas connue par les utilisateurs. Ils devront faire l'effort de la comprendre et de l'assimiler, pour qu'elle devienne par l'usage un schème d'utilisation. Mais un schème sensori-moteur peut demander lui aussi certains efforts car il doit être adapté à un monde virtuel avec un certain artefact et sous la contrainte d'incohérences

sensori-motrices. L'exploitation de métaphores nous éloigne t-elle de la finalité de la réalité virtuelle ? Nous le pensons a priori au regard de notre finalité, que nous avons explicitée au premier chapitre (...activités sensori-motrices...) Nous utiliserons une métaphore au lieu d'un schème sensori-moteur en cas de difficultés théorique, économique ou technique. Dans la pratique, nous pouvons avoir quelques fois l'exploitation associée de métaphores et de schèmes, selon le type et l'importance des activités désirées.

Une autre possibilité, en cas difficultés pour exploiter un schème employant les mêmes sens et réponses motrices que dans le monde réel, est d'employer un schème par la « **substitution sensori-motrice** ». Par exemple, il est souvent difficile de transmettre des efforts sur l'utilisateur par une interface à retour d'effort. On peut contourner la difficulté en substituant le sens de perception des efforts par l'ouïe : un son est émis quand l'objet, manipulé par l'utilisateur, rencontre un obstacle. Cette substitution est efficace dans le cas présent s'il ne s'agit que de prévenir l'utilisateur de l'obstacle. Par contre, si la perception des intensités et des directions des efforts est importante pour l'application (assemblage de pièces, par exemple), la substitution n'est pas appropriée. La substitution peut aussi être « motrice » et non sensorielle : au lieu de se déplacer en marchant physiquement, l'utilisateur peut commander à la voix son déplacement virtuel (je désire avancer, reculer, tourner à gauche, à droite, etc.) avec une interface à commande vocale et devant un point de vue dynamique du monde virtuel. Il ne s'agit pas dans ce cas de métaphore virtuelle, car le sujet exploite le schème de déplacement, mais par une substitution motrice.

Comme la métaphore virtuelle, le schème de substitution sensori-motrice éloigne l'action virtuelle de l'action réelle. Mais le lecteur ne doit pas penser qu'automatiquement l'immersion et l'interaction seront plus difficiles à réaliser, avec apprentissage plus long, parce que moins similaires au réel. Si, en n'employant un schème sensori-moteur et non une métaphore virtuelle ou une substitution sensori-motrice, la conséquence est d'utiliser des interfaces trop imparfaites techniquement, cela risque de plus pénaliser les actions de l'utilisateur, même si l'action de ce dernier est plus proche de l'action dans le monde réel.

Un autre critère est important à étudier pour savoir s'il faut ou non exploiter un schème sensori-moteur. Dans le cas où l'action sensori-motrice est elle même sujet d'étude, ce que l'on rencontre souvent pour la conception de produit, l'exploitation d'un schème sensori-moteur dans le monde virtuel est *indispensable* et doit se rapprocher le plus physiquement du schème en environnement réel : par exemple, la manipulation des boutons de commande d'une planche de bord pour une étude ergonomique de ce produit. Par contre, si l'action sensori-motrice ne sert qu'à indiquer ce que souhaite le sujet, l'exploitation d'un schème sensori-moteur n'est pas indispensable et est à comparer à une métaphore virtuelle ou un schème de substitution sensori-motrice. Il faut donc analyser l'application pour savoir s'il s'agit d'une activité étudiée au niveau sensori-motrice (schème indispensable) ou d'une simple indication d'activité.

L'utilisation de schèmes sensori-moteurs, de schèmes avec substitution sensori-motrice ou de métaphores est un débat qui est loin d'être clos dans la communauté scientifique. C'est ce que j'ai pu vérifier lors du débat que j'ai animé pendant le colloque REVICO à l'ENST [Fuchs, 1999a], voir le paragraphe « discussion » à la fin du chapitre.

2.4 MÉTHODE DE CONCEPTION ET D'ÉVALUATION D'UN ENVIRONNEMENT RV

2.4.1 LE SCHÉMA DE RÉFÉRENCE EN RV

Les considérations précédentes sur les processus cognitifs liés à l'interfaçage comportemental ne suffisent pas à établir une méthode de conception d'un dispositif en réalité virtuelle. Dans ce but, nous proposons en premier un schéma général pour la réalité virtuelle qui définit trois niveaux d'immersion et d'interaction avec leurs caractéristiques propres. Nous avons déjà explicité qu'au niveau physique, nous parlons **d'immersion et d'interaction sensori-motrices**, puisque physiquement l'ordinateur est connecté à l'homme par ses sens et ses réponses motrices. Ce niveau d'immersion et d'interaction est quantifiable par rapport aux caractéristiques des sens et des réponses motrices exploités. Comme explicité dans le schéma triadique sur l'interfaçage, l'utilisateur doit s'immerger mentalement dans le monde virtuel, le niveau «inférieur» d'immersion et d'interaction sensori-motrices devant lui être mentalement invisible (transparent). Nous parlons dans ce cas **d'immersion et d'interaction mentales**. C'est à ce niveau que se situent les processus cognitifs de l'interfaçage (schèmes, métaphores, substitutions). A un troisième niveau concernant l'application de réalité virtuelle, l'objectif est de s'attacher à réaliser *une immersion de l'homme pour une tâche donnée (ou une fonctionnalité)* et non pour une simple immersion mentale de l'homme dans ce monde virtuel. Nous parlons dans ce cas **d'immersion et d'interaction fonctionnelles**⁴. Nous pouvons rapprocher ce découpage aux trois niveaux de la représentation voisine de M. Fréjus : modèle sensoriel⁵, modèle cognitif et modèle opératif [Fréjus, 1996]. Ce découpage en trois niveaux nous permet de mieux clarifier les différents problèmes rencontrés lors de l'immersion et de l'interaction d'un sujet. Ils sont étroitement liés et ne s'opposent pas. Pour préciser leur lien, donnons l'exemple de la préhension d'un objet. Nous pouvons utiliser le schème de la préhension (I^2 mentales) avec un traqueur à 6 degrés de liberté et un écran affichant l'objet (I^2 sensori-motrices). Mais si les I^2 sensori-motrices sont de mauvaise qualité (temps de réponse trop grand entre l'action avec le traqueur et la perception du mouvement sur l'écran), le schème de préhension ne pourra être exploité. Inversement, si les interfaces ne font pas appel dans leur usage au schème de la préhension, même avec des interfaces fonctionnant correctement, les I^2 mentales seront en échec.

Le fondement de notre démarche est basé sur ce schéma hiérarchique en trois niveaux et aussi sur un découpage transversal entre le sujet et le monde virtuel : parallèlement aux différents niveaux d' I^2 sensori-motrices et mentales pour la personne, nous avons deux niveaux de fonctionnement logiciel pour le monde virtuel. Symétriquement aux I^2 sensori-motrices, l'ordinateur doit gérer la partie logicielle temps réel (noyau temps réel et drivers pour les interfaces matérielles), permettant une **réalisation physique du monde virtuel**. Face aux I^2 mentales, la partie logicielle spécifique de réalité virtuelle doit gérer la **modélisation comportementale du monde virtuel**. Cette partie logicielle doit fournir des *Aides Logicielles Comportementales* (ALC) de «motricité et de perception programmées», associées aux schèmes exploités, pour faciliter les I^2 mentales de la personne, comme nous l'avons explicité précédemment.

Au niveau I^2 fonctionnelles, par rapport à l'application et ses objectifs, il faut se po-

⁴on utilise pour la suite la notation I^2 pour l'immersion et l'interaction.

⁵qui aurait dû être nommé «sensori-moteur».

ser la question suivante : quelles sont les activités que l'utilisateur doit exécuter ? Si le lecteur veut prendre le temps d'y réfléchir, il verra rapidement que dans toutes les applications RV, les activités du sujet sont toujours décomposables en quelques comportements de base que nous appelons les «Primitives Comportementales Virtuelles» (PCV). Il faut donc au niveau des I² fonctionnelles bien définir les PCV et leurs spécificités. Quelque soit l'application, ces dernières peuvent être regroupées en quatre catégories :

- observer le monde virtuel ;
- se déplacer dans le monde virtuel ;
- agir sur le monde virtuel ;
- communiquer avec autrui⁶.

Les PCV et les ALC sont développées au paragraphe suivant. Nous obtenons en final **le schéma de référence de la réalité virtuelle** (Figure 2.10) qui va nous servir de canevas dans notre démarche de conception (à noter que le schéma triadique a été transformé pour faire ressortir la modélisation comportementale du monde virtuel).

Avec ce schéma pluridisciplinaire⁷, nous avons clarifié les notions d'immersion et d'interaction. Dans l'optique de la conception d'un dispositif RV, au lieu de chercher à reproduire le plus fidèlement possible la réalité, nous nous intéressons à présent à **l'optimum d'I² fonctionnelles** liées à l'application et par analyse et déduction aux optimums des I² mentales et sensori-motrices. C'est cette démarche novatrice que nous allons détailler [Fuchs, 1999b].

Remarques :

- la détermination des I² fonctionnelles correspond à l'étape de *spécifications* dans le cycle en V de la vie d'un produit. L'élaboration des I² sensori-motrices et mentales, interdépendantes, correspond à l'étape de *conception* du cycle en V ;
- à la base, il y a donc à déterminer les I² fonctionnelles. Dans certains cas, elles peuvent être évidentes⁸, dans d'autres cas, elles doivent être étudiées précisément avec le client de l'application. Comme dans d'autres situations de conception, le client formule souvent une demande assez floue de ce qu'il souhaite. Il faut donc faire avec lui une analyse poussée, *indépendamment* des solutions envisagées aux niveaux des I² mentales et sensori-motrices. Nous avons validé cette démarche dans le cadre de plusieurs applications professionnelles.

⁶Le sujet peut aussi avoir à communiquer avec le système informatique, pour modifier sa configuration, pour donner des ordres, etc. Mais dans ce cas, le sujet n'est plus en immersion dans un environnement virtuel. On se situe plus alors dans le domaine classique de dialogue homme-machine des IHM.

⁷A chaque bloc on peut associer : le physicien (pour les I.C.), le neurophysiologiste (sens et réponses motrices), l'informaticien (logiciels), le psychologue (processus mentaux) et l'ergonome (perception et motricité désirées).

⁸Par exemple, lors de la conception d'un produit, si on souhaite *simuler pour tester* une fonctionnalité précise du produit : comme l'étude ergonomique des boutons de commande d'une planche de bord de véhicule.

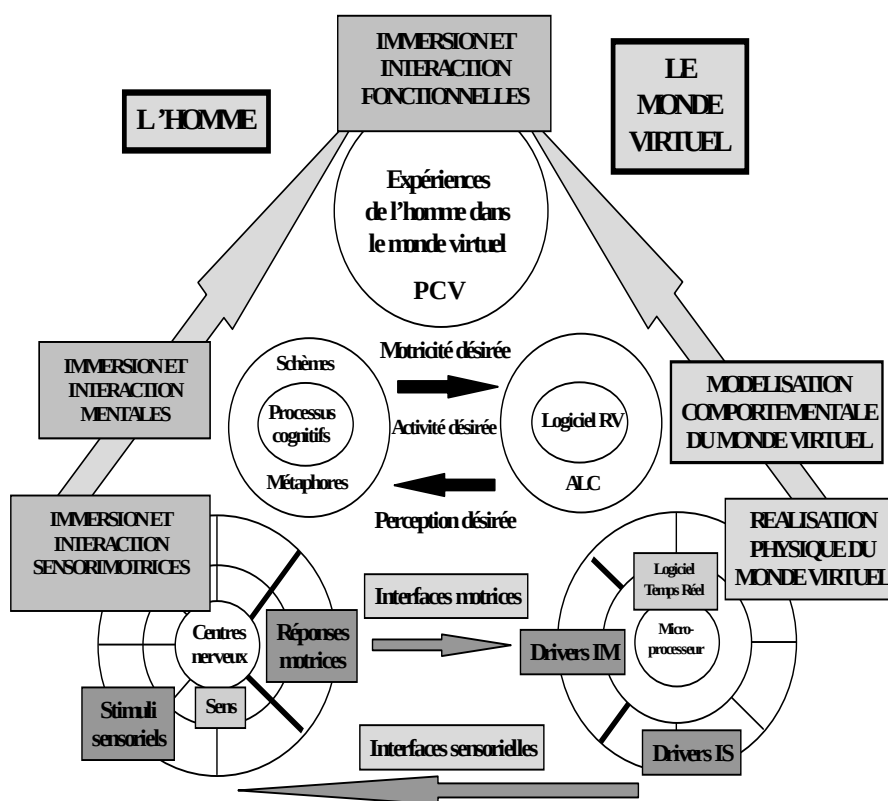


Figure 2.10 : Schéma technocentrique de référence en RV

2.4.2 LES PRIMITIVES COMPORTEMENTALES VIRTUELLES ET LES AIDES LOGICIELLES COMPORTEMENTALES

Ce paragraphe permet de préciser ce que l'on entend par «Primitives Comportementales Virtuelles» (PCV). Quand le sujet se trouve en environnement virtuel, il a une ou plusieurs activités à accomplir. Celles-ci peuvent se décomposer en activités élémentaires, sensori-motrices et mentales, que nous appelons PCV. Après réflexion, ces dernières peuvent être regroupées en quatre catégories en environnement virtuel :

- observer le monde virtuel ;
- se déplacer dans le monde virtuel ;
- agir sur le monde virtuel ;
- communiquer avec autrui.

Dans le cas de l'observation du monde virtuel, nous avons plusieurs sous catégories de PCV :

- selon que l'observation est visuelle, auditive ou tactile ou est réalisée par combinaison de ses sens ;
- selon qu'il s'agit de grands ou de petits espaces ;
- selon qu'il s'agit de comprendre l'environnement ou de s'orienter par rapport à celui-ci.

Dans le cas de déplacement, nous avons aussi plusieurs sous catégories de PCV, selon le type de déplacement : trajectoire 1D (sur une droite ou sur une courbe), déplacement sur une surface (plane ou non), dans l'espace (avec ou non changement d'orientation à 3 ou 6 degrés de liberté (ddl)).

Dans le cas d'action sur le monde virtuel, nous avons plusieurs sous catégories de PCV : manipuler un objet en translation (3ddl), orienter un objet en rotation (3ddl), les deux actions associées, déformer un objet ou assembler des objets.

Dans le cas de la communication, nous avons plusieurs sous catégories de PCV : communiquer avec d'autres utilisateurs, communiquer avec des personnages virtuels (avatars ou clones virtuels) ou communiquer avec le système informatique (mais cette catégorie est hors RV).

Nous pouvons espérer que progressivement toutes les PCV de base seront simulées dans différents dispositifs RV. Avec du recul par rapport à toutes les solutions testées, nous aurons alors plus de facilités pour concevoir le dispositif RV et pronostiquer plus sûrement le succès de la réalisation des I² fonctionnelles.

Les Aides Logicielles Comportementales (ALC) de motricité et de perception programmées, associées aux PCV, apportent des aides dans la tâche de l'utilisateur. Le concepteur doit y réfléchir pour proposer celles qui aident efficacement l'utilisateur, tout en respectant l'objectif des I² fonctionnelles. Par exemple, dans un système de formation en environnement virtuel pour des agents de conduite de TGV (nous en parlerons), l'utilisateur doit se rendre près d'un poteau téléphonique pour décrocher

le téléphone. Vu les interfaces utilisées (grand écran, tapis roulant et guidon), il est difficile au formé de se déplacer aussi facilement qu'à pied en environnement réel (sans ALC, le formé peut tourner désespérément plusieurs fois autour du poteau en manipulant le guidon). Pour faciliter cette tâche, une ALC a été programmée : dès l'approche du poteau et en étant sûr de l'intention du formé, celui-ci se dirige automatiquement face au poteau. Ceci respecte l'objectif des I² fonctionnelles car la formation ne concerne pas le déplacement en environnement virtuel !

Un autre exemple peut être pris dans le domaine des simulateurs fixes de transport. L'INRETS préconise de mettre un retour d'effort sur le volant du véhicule plus important que celui obtenu sur un vrai véhicule. Ainsi, pour des conducteurs novices en simulateur, ceux-ci ont une perception accrue du comportement du véhicule sur la route, malgré l'absence d'information de mouvements sur les systèmes vestibulaires des conducteurs. Dans ce cas, il est évident qu'au niveau des I² fonctionnelles, le réalisme du retour d'effort sur le volant n'est pas exigé, mais qu'une des fonctions est de permettre à tout conducteur novice de maîtriser le véhicule.

Les ALC peuvent être déterminées en fonction des « affordances » des objets ou de l'environnement autour du sujet. D'après la théorie de Gibson [Gibson, 1979], une affordance représente les interactions possibles entre l'objet et le sujet. Ces interactions sont perçues par le sujet par rapport à l'idée qu'il se fait de la fonction de l'objet, plus qu'il ne perçoit les caractéristiques physiques, géométriques, etc., de l'objet ou de l'environnement. Il est donc plus utile de savoir par avance à quoi un élément de l'environnement va servir que d'avoir des notions précises de sa géométrie ou de ses caractéristiques physiques. Des chercheurs se basent sur cette approche pour analyser le comportement du sujet dans un environnement virtuel [Morineau, 2001]. Pour notre part, connaissant les affordances de tel objet, nous pouvons prévoir une ALC qui aide le sujet à interagir efficacement avec cet objet. Si nous reprenons un de nos exemples, une table (affordances : on peut y poser des objets) et des objets manipulables, dès que ceux-ci sont rapprochés de la table par l'utilisateur immergé, ils sont plaqués sur la table et ne gardent que trois degrés de liberté : d'où cette ALC de manipulation.

2.4.3 DÉMARCHE DE CONCEPTION

Pour la conception d'un système de réalité virtuelle, nous faisons l'hypothèse qu'il est préférable d'atteindre le niveau de l'intelligence sensori-motrice de l'utilisateur pour lui faciliter ses I² dans le monde virtuel, plutôt que de faire appel à son niveau d'intelligence sémiotique. Donc pour une PCV donnée, avec certaines caractéristiques exigées, nous privilégions le choix du schème, avec son artefact correspondant, qui soit le plus efficace en terme psychologique, technique et économique, plutôt que d'exploiter une métaphore.

La démarche chronologique est la suivante : on spécifie *très rigoureusement* les I² fonctionnelles pour l'application (étape longue et délicate dans bien des cas). A partir des I² fonctionnelles désirées, on détermine les PCV utiles. Pour chacune des PCV, on recherche au niveau des I² mentales, soit a priori ou soit par rapport aux expériences passées en réalité virtuelle, le ou les schèmes susceptibles de fournir une immersion et une interaction pseudo-naturelles. Dans le cas d'impossibilité de trouver des schèmes efficaces et réalisables techniquement à un prix abordable, on choisit une métaphore appropriée ou une substitution sensori-motrice, compatible avec les I² fonctionnelles. On détermine ensuite simultanément les I² sensori-motrices, les sens et(ou) les réponses

motrices et l'artefact de l'Interface Comportementale, associés au schème choisi. Les caractéristiques métrologiques des interfaces matérielles doivent être compatibles aux caractéristiques psychophysiques des sens et des réponses motrices, en rapport avec les I^2 sensori-motrices recherchées. Des Aides Logicielles Comportementales seront ensuite conçues pour améliorer les I^2 mentales. N'oublions pas qu'il y a toujours des incohérences sensori-motrices et que les ALC peuvent permettre partiellement au cerveau du sujet d'accepter ces incohérences. Il n'y a plus ensuite qu'à programmer le logiciel RV, et pas avant !

Résumé chronologique de la démarche de conception :

- étape 1 : détermination des I^2 fonctionnelles ;
- étape 2 : détermination des PCV requises à partir des I^2 fonctionnelles ;
- étape 3 : conception des Interfaces Comportementales à partir des PCV, (Schèmes, substitution sensori-motrice ou métaphores – artefacts - sens et réponses motrices sollicités) ;
- étape 4 : conception des ALC à partir des IC. ;
- étape 5 : conception du logiciel RV (programmation des ALC, des drivers des IC et de l'environnement virtuel).

Le choix d'un schème, d'une substitution sensori-motrice ou d'une métaphore virtuelle est donc très important dans la phase de conception et doit être étudié avec attention. Comme déjà explicité, les schèmes utilisés en réalité virtuelle sont plus ou moins décalés, «biaisés», par rapport aux schèmes correspondants, exploités dans le réel. En particulier les schèmes sont plus biaisés si le retour sensoriel (le feed-back) est inexistant, partiel ou faux. Dans l'exemple de la manipulation d'un produit dans un magasin virtuel, le biais est dans le retour sensoriel, qui est partiel (il n'y a pas de retour d'effort : poids et inertie du produit inexistant). Les schèmes exploités doivent être assimilés et familiers par la population concernée d'utilisateurs (enfants, professionnels, etc.). Dans tous les cas, les schèmes exploités doivent annihiler les incohérences sensori-motrices, ce qui est possible si elles sont faibles (comme dans l'exemple précédent). Ces incohérences sont un des problèmes majeurs pour les I^2 sensori-motrices et mentales. Les phénomènes perceptifs sont totalement subjectifs et chacun interprète une sensation particulière en référence à ce qu'il a pu apprendre par ses expériences. Celles-ci peuvent être d'une aide précieuse dans certaines applications ou inversement une gêne pour l'utilisateur. Un exemple peut être pris dans le cinéma immersif à champ de vision total (360°)⁹ : les spectateurs immergés visuellement dans une scène vont réagir physiquement selon leurs expériences passées. Si on projette sur les écrans une descente rapide d'une route en lacet, les spectateurs vont incliner leur corps pour réagir à des mouvements bien virtuels ! Ceci malgré les informations transmises par leurs oreilles internes leur indiquant l'absence de toute inclinaison. Certaines personnes peuvent en ressentir une gêne, ce qui va à l'encontre du but recherché : une meilleure immersion. Le concepteur de systèmes de réalité virtuelle doit en conséquence tenir compte de la réaction individuelle de chaque personne qui ne sera pas forcément celle qu'il escompte. Cette discordance entre deux informations sensorielles peut être même une cause de malaise important : «Enfermé» dans un visiocasque, l'utilisateur n'a plus de repère visuel de son espace réel, mais détecte toujours la verticalité de celui-ci par les vestibules de ses oreilles internes. L'espace virtuel observé n'est plus en correspondance avec l'espace réel localisé et ceci peut entraîner un «mal de mer». L'étude

⁹Même s'il ne s'agit pas de procédé de réalité virtuelle car l'interaction y est inexistante.

des incohérences sensori-motrices et de l'adaptation neurophysiologique du cerveau à celles-ci sont des *problèmes ouverts* dans le domaine de la réalité virtuelle. Les spécialistes en neurophysiologie et en sciences cognitives doivent encore approfondir ce domaine contraignant pour définir des règles pratiques et pour aider le concepteur de dispositif RV.

Cette démarche générale de conception doit être validée à différentes étapes par les concepteurs. Il faut avoir à l'esprit qu'un objet technique (donc une interface comportementale) a plusieurs fonctionnements : celui prévu par le concepteur et celui utilisé par l'opérateur. La relation d'usage d'un instrument par l'utilisateur n'est pas forcément celle pensée par le concepteur. Le concepteur a une représentation technique et fonctionnelle de l'interface comportementale (schéma technocentrique) et l'utilisateur en n'aura qu'une représentation externe, par ses effets (c'est, pour lui, une boîte noire, schéma anthropocentrique).

En exemple, le déplacement virtuel de l'utilisateur sur un sol peut se faire par une commande manuelle (gant de données ou traqueur à 6 degrés de liberté) si on ne veut ou ne peut exploiter un tapis roulant. Mais dans ce cas, il y a un risque que l'utilisateur interprète ses commandes sur l'interface différemment de celles prévues par le concepteur. La difficulté pour l'utilisateur est celle de l'auto-représentation de son corps : où suis-je, où vais-je ? Concrètement, quel est le repère que je déplace ? Celui de mon point de vue, celui de la direction de ma trajectoire ou celui au centre de mon corps ? Nous avons testé un mode de déplacement **relatif**. L'utilisateur déplace l'interface vers l'avant pour avancer, vers l'arrière pour reculer, sur le côté pour se déplacer en crabe et tourne l'interface pour pivoter. Sa vitesse de déplacement dans le monde virtuel est fonction de la distance entre la position courante du traqueur et sa position d'origine. L'arrêt du déplacement se fait en ramenant l'interface dans la «zone morte» qui englobe la position d'origine. Pendant tous les déplacements, une main virtuelle représentait à l'écran la main de l'utilisateur. Mais l'image de cette main virtuelle prêtait à confusion. Pour la plupart des utilisateurs ayant testé ce système, le mode de déplacement était **absolu**. Ils cherchaient à aligner leur main virtuelle (et donc l'interface) avec le point qu'ils souhaitaient atteindre. Ils ont utilisé le schème suivant : la main indique la direction où l'on veut aller, ce que n'avaient pas prévu les concepteurs. Certains utilisateurs pensaient que la main représentait tout leur corps et «téléopéraient» leur déplacement.

2.4.4 DÉMARCHE DE L'ÉVALUATION

L'évaluation d'un système en réalité virtuelle doit être basée sur le même canevas de notre schéma général. Elle ne doit pas seulement porter sur les I^2 sensori-motrices, comme trop souvent observé. Il ne faut pas chercher à évaluer le degré de réalisme d'une façon absolue (en espérant avoir le maximum de degré de réalisme), mais on doit se référer aux seules PCV exigées et à leurs schèmes correspondants pour obtenir de correctes I^2 fonctionnelles.

Au niveau des I^2 sensori-motrices, les caractéristiques «métrologiques» de l'artefact utilisé (l'interface matérielle) devraient théoriquement être de valeurs identiques aux caractéristiques maximales psychophysiques des sens et réponses motrices de la personne (sauf exception, on compare aux caractéristiques moyennes humaines). Il n'est bien sûr pas utile d'avoir des caractéristiques métrologiques supérieures à celles psychophysiques. Si elles sont inférieures, ce qui est souvent le cas, hélas, on peut dé-

terminer quantitativement le déficit «sensori-moteur» relatif pour chaque sens et pour chaque réponse motrice sollicités. Par exemple pour toute interface visuelle, il est souhaitable d'avoir un affichage tel que deux pixels adjacents soient distants de moins de 1' d'angle par rapport au point de vue de l'utilisateur, car son acuité visuelle est de 1' d'angle au minimum. Mais on est souvent actuellement à des valeurs plus hautes, surtout dans les visiocasques. En cas de déficit sensori-moteur, on doit étudier si celui-ci est compatible avec les I^2 mentales souhaitées. Les caractéristiques métrologiques les plus importantes pour les interfaces sont la précision, la sensibilité, l'étendue d'utilisation, la fréquence de mesure et le temps de réponse.

Des tests d'évaluation doivent être entrepris au niveau des I^2 sensori-motrices. Le concepteur ne doit pas se satisfaire de vérifier que les caractéristiques de l'interface sont suffisantes par rapport à l'application. Il ne doit pas évaluer *en même temps* au niveau des I^2 sensori-motrices et au niveau des I^2 mentales ! Nous ne pouvons pas bien sûr donner tous les types de tests à faire dans ce chapitre général, nous en parlerons pour chaque type d'interface au cours des chapitres correspondants. Nous ne donnons qu'un exemple ici pour bien attirer l'attention du lecteur. Lors de l'expertise d'un visiocasque haut de gamme, nous avons évalué les champs de vision, horizontal et vertical, des deux yeux. Par déplacement de repère (lignes, croix) dans le champ de vision de plusieurs observateurs, nous avons relevé les valeurs effectives. Elles étaient inférieures à ce qu'indiquait la notice du produit. De plus, le champ de vision maximum horizontal ne pouvait être atteint en même temps que le maximum de recouvrement des champs de vision des deux yeux (le champ stéréoscopique, «overlap»), ce que passait sous silence la notice. Vous devez vérifier toutes les caractéristiques, car même si celles annoncées sur les notices sont justes, rien ne vous garanti que dans votre installation l'interface est exploitée dans les meilleures conditions. Cela peut dépendre de l'ordinateur, de vos logiciels, des données et de l'utilisateur.

Au niveau des I^2 mentales, l'évaluation est plus difficilement quantifiable. Dans le cas d'un choix a priori d'un schème, des tests de validation doivent confirmer le choix effectué. Plus le temps d'apprentissage de l'utilisation de l'I.C. est faible, plus cela indique que le schème est efficace. Pour évaluer concrètement la performance d'un interface comportemental naturel, nous proposons de faire des tests classiques d'ergonomie et des tests «sans explication» : les testeurs essayeront d'utiliser l'interface sans qu'il leur soit donné des explications avant usage. Si l'interface fait appel à un schème, peu biaisé en environnement virtuel, les personnes doivent interagir pseudo-naturellement après un temps très court d'autoapprentissage. La durée du temps d'apprentissage sera considérée courte suivant l'application. Comme nous l'explicitons dans un paragraphe suivant, si une consommatrice doit faire ses achats dans un magasin virtuel pendant 20 minutes, il ne faut pas que le temps total d'apprentissage des I^2 ne dépasse 2 à 3 minutes. Si des agents de la SNCF doivent se former en une journée à des tâches sur une infrastructure ferroviaire, ils ne doivent pas avoir un temps d'apprentissage des I^2 de plus d'un quart d'heure. Durant ce temps d'apprentissage, le sujet va concevoir sa propre utilisation du système et organiser son monde virtuel. Du point de vue psychologique, nous sommes alors dans la phase de *conception par l'usage*.

Au niveau des I^2 fonctionnelles, nous ne pouvons donner qu'une règle générale d'évaluation. La fonction du dispositif RV doit être comparée à la fonction similaire dans un monde réel : la formation des agents de la SNCF est-elle plus efficace dans un environnement virtuel que dans un monde réel ? La consommatrice a-t-elle un comportement similaire dans un magasin virtuel à celui dans un magasin réel ? A chaque application, les concepteurs devront trouver la méthode d'évaluation la plus propice.

2.4.5 OUVERTURE D'UNE PORTE

En premier exemple, nous prenons le cas simple où nous souhaitons que l'utilisateur franchise une porte. Au lieu de faire appel à une métaphore (cliquer sur l'icône d'une porte entrouverte), la personne doit manipuler *naturellement* une poignée. La simple représentation visuelle de la poignée peut indiquer à l'utilisateur le geste qu'il doit faire. La forme géométrique de la poignée fait inconsciemment appel au schème d'ouverture correspondant : mouvement de rotation, de translation ou combiné. Des Aides Logicielles Comportementales, couplés à un traqueur à six degrés de liberté, facilitent la manipulation : un attracteur de proximité et des contraintes géométriques sur le mouvement de la poignée. La représentation visuelle liée à l'interface comportementale est donc importante : elle doit faire appel directement à un schème voulu par les concepteurs. Dans l'exemple, la forme de la poignée indique inconsciemment à l'utilisateur les mouvements à exécuter. Cet exemple rudimentaire avec schème sensori-moteur n'est bien sûr à employer dans une application RV globale que si la simulation de l'ouverture de porte est importante par rapport aux I² fonctionnelles et mentales. Autrement une métaphore ou un schème avec substitution sensori-motrice peut être exploité.

2.4.6 MAGASIN VIRTUEL D'EXPÉRIMENTATION COMMERCIALE

2.4.6.1 Introduction

L'entreprise IN VIVO réalise des études pour des grandes entreprises de la distribution afin de pronostiquer les chances de succès de leurs nouveaux produits. Des supermarchés «factices», mais réels, sont ainsi reconstitués comportant différents rayonnages dont certains contiennent les nouveaux produits à tester. On demande à des consommateurs de faire des courses à partir d'une liste de produits. Leurs choix et leurs comportements sont analysés pendant leurs courses et ensuite par un entretien individuel. Cette entreprise s'est adressée au Centre de Robotique de l'Ecole des Mines de Paris et à l'entreprise SIMTEAM [http1] pour concevoir et réaliser un magasin virtuel d'expérimentation commerciale.

2.4.6.2 L'analyse du problème basée sur notre schéma général de la RV

Au niveau des I² fonctionnelles, le dispositif doit donc permettre à une personne de *faire naturellement ses courses*, comme cela se fait chez IN VIVO, avec la même démarche mais dans un magasin virtuel. Le système doit reproduire de manière réaliste une grande surface. Le consommateur doit pouvoir se déplacer entre les rayonnages, s'arrêter devant le rayon qu'il souhaite et y choisir un ou plusieurs produits. Ceux-ci doivent s'afficher avec une résolution de très grande qualité, car il s'agit de pouvoir y lire les textes imprimés et tester l'aspect esthétique de l'emballage du produit. Les I² fonctionnelles souhaitées peuvent se résumer ainsi : permettre à une personne de choisir un produit parmi d'autres sur un rayonnage en pouvant l'observer librement. Par contre, il ne s'agit pas de tester l'implantation et l'organisation du magasin, la commodité à faire ses courses, la signalétique dans le magasin, etc. En détaillant les différentes PCV dans le magasin virtuel, on peut classer celles-ci de la façon suivante :

1. Les trois PCV fondamentales sont :

- l'observation d'un produit ;
- la manipulation tridimensionnelle et l'orientation spatiale d'un produit (6ddl) ;
- l'observation des produits sur un rayonnage.

Ces trois PCV doivent être réalisées très facilement par le consommateur et l'aspect visuel du produit doit être extrêmement proche de la réalité.

2. Les autres PCV sont secondaires. Elles n'ont pas à être simulées parfaitement car elles ne sont pas le sujet de l'étude (hors des I² fonctionnelles). Elles sont utiles pour mettre les consommateurs en condition d'achat dans un magasin virtuel :

- le client doit pouvoir se déplacer et s'orienter facilement dans le magasin. Il doit pouvoir aller dans l'allée latérale de son choix, où se trouve toute une gamme de produits. La disposition des produits dans les étagères doit être identique à celle d'un magasin réel ;
- le client doit pouvoir prendre un produit ou le remettre dans un rayon et l'acheter en le mettant dans son chariot.

Au niveau des I² mentales, nous en avons déduit :

Pour les trois PCV fondamentales :

- exploitation du schème d'observation visuelle dans une seule direction fixe ;
- exploitation du schème de préhension et de manipulation spatiale d'un produit ;

2 Pour les PCV secondaires :

- exploitation du schème de déplacement en poussant un chariot ;
- exploitation du schème de préhension et de manipulation spatiale d'un produit.

Pour les PCV secondaires, l'exploitation de schèmes n'est pas indispensable, mais les solutions à base de schème étant efficaces et abordables financièrement pour le magasin virtuel, nous les avons utilisées.

Au niveau des I² sensori-motrices, il faut que les schèmes et les artefacts choisis soient d'un emploi aussi naturel que possible pour induire un temps d'adaptation minimal. (Une à trois minutes, temps relativement faible par rapport au temps global de l'expérience : vingt minutes). Dans ces conditions, l'immersion du consommateur sera réussie, il oublie qu'il est face à un système informatique et il peut se concentrer sur ses achats, produisant ainsi des données statistiques non biaisées. Nous avons pu le vérifier dans différents tests, dont un premier en juin 98 : sur 24 personnes testées, toutes ont pu faire leurs courses, après 3 à 5 minutes d'explications. Après ce temps d'apprentissage, toutes les personnes avaient un comportement d'achat et les interfaces comportementales étaient transparentes pour elles. Nous allons détailler au paragraphe suivant les solutions employées pour les différentes PCV.

2.4.6.3 L'observation visuelle des produits

Afin d'obtenir une immersion visuelle la plus naturelle et la plus pertinente, il faut envisager une projection **à l'échelle 1** des rayons. Le consommateur se trouve donc devant un grand écran. Le produit doit être affiché avec une **grande résolution**, correspondant aux capacités visuelles de l'utilisateur (acuité visuelle de 1' d'angle). Celui-ci doit pouvoir lire les petits textes sur les emballages. Il faut que **le champ de vision soit assez grand** (champ de vision de 130 °) pour que le consommateur puisse être suffisamment immergé¹⁰ et puisse s'orienter convenablement dans le magasin. Suite à ces caractéristiques psychophysiques nécessaires pour la vision, la solution consiste à prendre un ensemble de six vidéoprojecteurs LCD (de résolution 1024x768)¹¹. On obtient une image totale de 4m par 2m avec 3072x1536 pixels. Nous utilisons trois ordinateurs PC mis en réseau pour la synchronisation des affichages (Figure 2.11 et Figure 2.12).

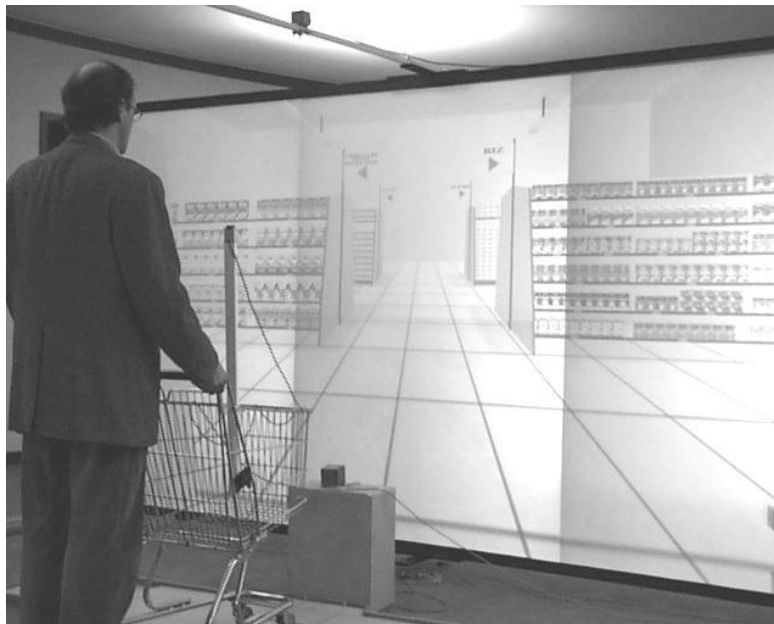


Figure 2.11 : Observation du magasin sur un écran de 4m par 2m, photo Ecole des Mines de Paris

2.4.6.4 La manipulation naturelle des produits 3D avec 6 ddl

Nous nous basons sur le schème de manipulation d'un objet dans le monde réel que nous reproduisons dans l'environnement virtuel. Le seul biais est dans le retour visuel :

¹⁰Face à des étagères réelles, nous avons vérifié qu'un champ de vision de 130° était nécessaire pour voir tous les produits observables.

¹¹Des vidéoprojecteurs à résolution 1280 x 1024 auraient permis de plus se rapprocher de l'acuité visuelle de 1' d'angle, mais leur prix étaient encore élevés (> 100 KF). L'acuité est ici proche de 2' d'angle. Il y a souvent un compromis à trouver !



Figure 2.12 : Réalisation de l'écran avec six vidéoprojecteurs LCD (3072x1536 pixels au total), photo Ecole des Mines de Paris

l'utilisateur regarde un écran et non ses mains, qui manipule un capteur de localisation à six degrés de liberté (traqueur), inclus dans un cube. Les mouvements de manipulation étant identiques à ceux du monde réel, nous avons pu vérifier que les utilisateurs n'ont aucune difficulté à manipuler virtuellement le produit (Figure 2.13). Pour faciliter la prise du produit sur l'étagère et pour la pause dans un chariot, nous avons ajouté des ALC (attracteurs de proximité et contraintes géométriques). Résultats des tests : 22 sur 24 personnes pouvaient observer sans difficulté les six faces d'un paquet et y lire les textes imprimés. Deux personnes avaient des difficultés pour observer les faces inférieures et supérieures d'un produit.

2.4.6.5 La navigation dans le magasin

Pour le déplacement dans les allées, les trajectoires possibles sont *prédéterminées* pour simplifier les commandes de déplacement de la ménagère (ALC). Il n'est pas utile dans notre cas de laisser à la cliente la possibilité de se mouvoir n'importe où, car nous n'étudions pas l'implantation des rayons du magasin. Nous nous basons donc sur le schème du «déplacement en poussant un chariot». La commande de déplacement passe par l'intermédiaire d'un chariot pour se rapprocher des actions dans un magasin réel. L'utilisateur doit pousser le chariot en avant ou en arrière et l'orienter à gauche ou à droite devant l'écran (Figure 2.14 et Figure 2.15). Le seul biais par rapport au schème dans le monde réel est dans l'amplitude des mouvements : devant l'écran, le chariot est déplacé légèrement seulement pour indiquer un déplacement *relatif*. Nous avons mis le chariot sur une plaque en bois ayant de faibles pentes appropriées : l'utilisateur doit

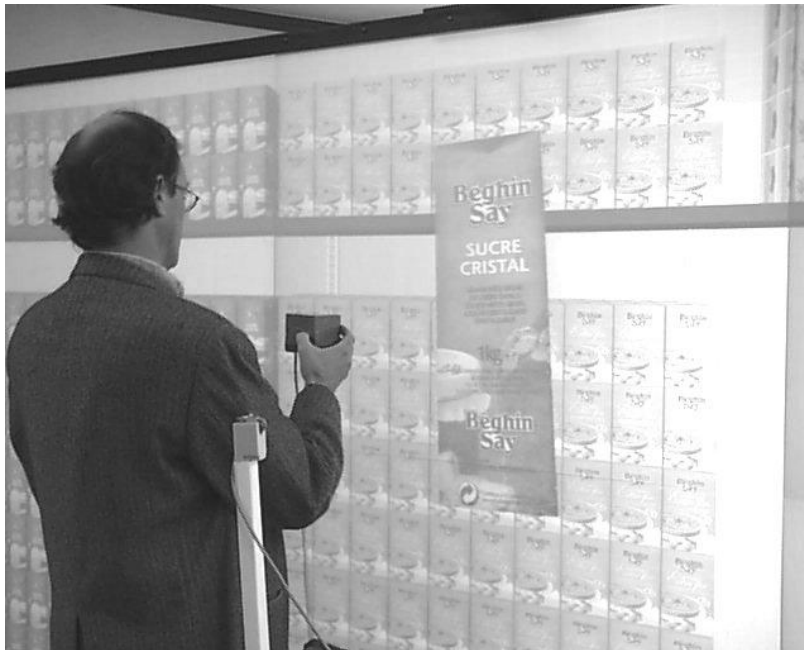


Figure 2.13 : Manipulation d'un produit avec un capteur de localisation (6ddl), inclus dans un cube, photo Ecole des Mines de Paris

alors pousser ou tirer le chariot avec des efforts un peu similaires à ceux accomplis dans le monde réel.

La conception et la réalisation de ce magasin d'expérimentation commerciale ont été basées sur notre schéma général de la réalité virtuelle et sur la recherche de schèmes les plus appropriés :

- faibles biais comportementaux ;
- réalisations des artefacts techniquement et économiquement envisageables ;
- exploitations des sens et des réponses motrices avec des caractéristiques psychophysiques adéquates par rapport à l'application ;
- Aides Logicielles Comportementales facilitant l'utilisation des interfaces comportementales.

Nous avons pu vérifier dans des tests la validité de notre conception brevetée dont la conclusion principale est que le client fait ses achats dans le magasin virtuel et n'est plus devant un écran au bout de trois minutes d'utilisation en moyenne. Une vidéo sur le magasin virtuel est consultable à l'adresse de notre Centre de Robotique [http2].

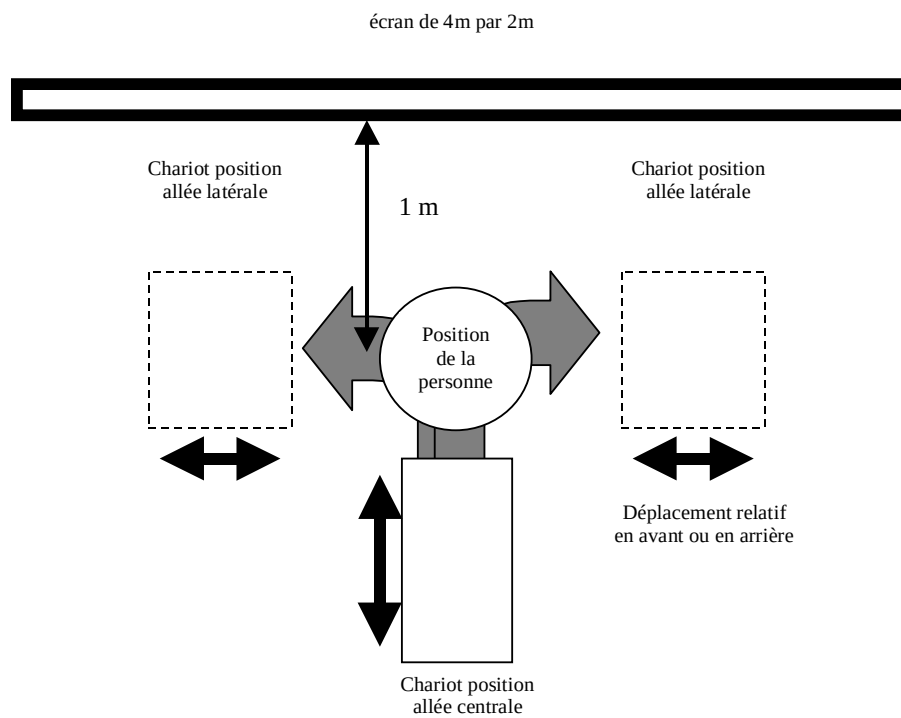


Figure 2.14 : Schéma des déplacements du chariot instrumenté

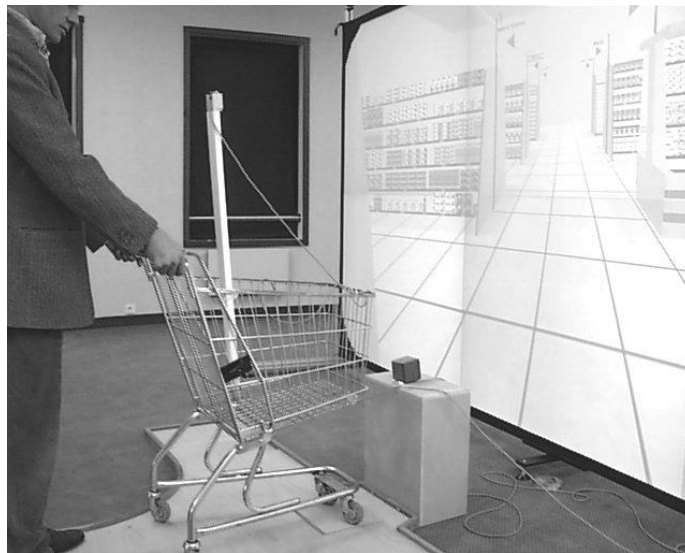


Figure 2.15 : Interfaçage comportemental par chariot instrumenté pour les déplacements dans le magasin, photo Ecole des Mines de Paris

2.4.7 FORMATION EN RÉALITÉ VIRTUELLE SUR DES INFRASTRUCTURES FERROVIAIRES

Les formations sur infrastructures ferroviaires sont souvent complexes à mettre en œuvre, coûteuses financièrement et n'offrent pas toutes les possibilités pédagogiques que permettent les simulations virtuelles. Nous avons collaboré avec la Direction des recherches de la SNCF, intéressée par la possibilité d'une formation en environnement virtuel d'agents de conduite à l'intervention sur les voies ferrées. La formation classique, en environnement réel, ne répond pas toujours aux besoins pédagogiques. La réalité virtuelle présente de nombreux avantages :

- réaliser des tâches sans danger (circulation ferroviaire virtuelle) ;
- reconfigurer l'environnement (changement de terrain, des conditions climatiques,...) ;
- réaliser des scénarios impossibles à reproduire dans la réalité (accidents, incidents techniques) afin de mettre le stagiaire en situation de stress,... [Lourdeaux, 1999].

2.4.7.1 L'analyse du problème basée sur notre schéma général de la RV

Au niveau des I² fonctionnelles, le dispositif RV doit donc permettre à *des agents de se former à une tâche sur des infrastructures ferroviaires, encadrés par un formateur*. Par exemple, les agents de conduite des TGV sont formés au contrôle et à la manœuvre d'aiguillage sur les lignes à grandes vitesses, en cas de dysfonctionnement des commandes des aiguillages. Pour le formateur, le dispositif est *un outil pédagogique* qui doit lui permettre de proposer différentes stratégies pédagogiques aux formés. Ces stratégies permettent d'optimiser le processus d'apprentissage en aidant le formé à se concentrer sur la tâche ou sur une sous-tâche et à choisir les raisonnements adéquats. Des assistances logicielles sont sélectionnées en fonction des comportements de chaque formé (erreurs de tâche, erreurs d'utilisation de l'interface, doutes, etc.) et selon leur niveau de connaissances (l'aide apportée à un novice sera plus importante que pour un expert). Ces assistances peuvent être déclenchées off line ou on line selon des critères définis par le formateur. En détaillant les différentes PCV, on peut classer celles-ci de la façon suivante :

- les trois PCV fondamentales sont l'orientation sur la voie ferrée, le déplacement 2D le long de celle-ci (plusieurs centaines de mètres) et la manipulation tridimensionnelle d'objets (3ddl) : la commande d'aiguillage et le téléphone. Ces PCV doivent être réalisées très facilement par l'agent, sans charge mentale importante pour l'utilisation des interfaces comportementales. Ainsi le formé peut se concentrer sur sa formation ;
- une PCV qui découle des trois précédentes est l'observation visuelle dans toutes les directions, sauf vers le haut. Une autre PCV complémentaire est l'observation sonore, en particulier pour l'écoute du bruit des TGV circulant sur les voies.

Au niveau des I² mentales, on en déduit :

Pour les trois PCV fondamentales :

- exploitation du schème d'orientation visuelle ;
- exploitation du schème de la marche ;
- exploitation du schème de préhension et de manipulation spatiale d'un produit.

Pour la PCV secondaire, exploitation du schème de l'observation visuelle dans toutes les directions.

Au niveau des I² sensori-motrices, il faut que les schèmes et les artefacts choisis soient d'un emploi naturel pour induire un temps d'apprentissage minimal. (Dix à vingt minutes, temps relativement faible par rapport au temps global de formation : quelques heures). L'immersion de l'agent de conduite doit être aussi réussie. Dans ces conditions, il oublie qu'il est face à un système informatique et il peut se concentrer sur sa formation. Nous allons détailler au paragraphe suivant les solutions employées pour les différentes PCV.

2.4.7.2 Le déplacement 2D sur les voies ferrées

A priori, le schème à exploiter permettant une interaction naturelle est évidemment la marche, qui est ici importante pour la formation : le formé doit se rendre compte *physiquement* que ses déplacements peuvent lui demander du temps, surtout s'il se trompe, tout en sachant que son temps d'intervention sur la voie doit être le plus court possible. Le problème technique de base est la difficulté de faire un tapis roulant 2D, autorisant la marche dans les deux directions. Des prototypes ont été réalisés, mais ils posent des difficultés de réalisations mécaniques et d'utilisation. Même un tapis roulant 1D dans une seule direction impose une incohérence sensori-motrice : le système vestibulaire de l'utilisateur indique une absence de déplacement à l'opposé des sensations venant de ses membres inférieurs. On pense que le biais introduit dans le schème de la marche sur un tapis 1D est facilement assimilable par les utilisateurs. En est-il de même pour un tapis 2D ? Dans le cas de l'exploitation d'un tapis 1D, il faut utiliser un autre schème pour permettre de tourner. Nous pensons que nous pouvons exploiter le schème du changement de direction par un guidon (Figure 2.16). Nous avons fait des tests pour validation (le changement de direction se fait dans ce cas par les mains et non par les pieds). Des ALC de motricité sont associées aux artefacts exploités pour faciliter les I² mentales de la personne. Par exemple, quand on détecte l'intention de l'utilisateur de se rapprocher du téléphone ou de la commande de l'aiguillage, le déplacement virtuel de l'utilisateur est dans ce cas contrôlé pour atteindre automatiquement le lieu désiré. Le choix entre un tapis 1D ou 2D a bien sûr une incidence sur l'immersion visuelle : si l'utilisateur tourne réellement dans toutes les directions horizontales, il faut que l'affichage soit aussi dans toutes ses directions.

2.4.7.3 L'orientation sur les voies

L'orientation pose un problème délicat. Des expériences [Chance, 1998] ont montré que la rotation *réelle* du corps de la personne sur elle-même était relativement importante pour son orientation (à l'inverse des translations). L'utilisateur pourra-t-il toujours bien s'orienter si on exploite un tapis 1D ? Des tests sont à réaliser pour analyser les différentes solutions.



Figure 2.16 : *Interfaçage comportemental par un tapis roulant asservi et un «guidon» pour se déplacer, photo Ecole des Mines de Paris*

2.4.7.4 L'immersion visuelle

Le formé devant pouvoir aisément dialoguer avec le formateur à ses côtés, nous ne souhaitons pas utiliser un visiocasque pour l'immersion visuelle. L'utilisateur devant manipuler des objets et se déplacer le plus naturellement, l'environnement virtuel doit être représenté à l'échelle 1. Le formé doit avoir un grand champ de vision pour pouvoir contrôler l'arrivée des trains sur la voie et pour bien s'orienter sur celle-ci. Son champ de vision doit être important vers le bas pour pouvoir manœuvrer la commande d'aiguillage qui est à la hauteur des voies. Pour permettre cela, la vue observée sur l'écran est fonction de la hauteur de tête où est fixé un capteur de localisation. Ainsi, si l'utilisateur se baisse, son point de vue change et il peut voir au niveau du ballast la commande de l'aiguillage. L'affichage est réalisé en rétroprojection par un vidéoprojecteur en employant un tapis 1D (Figure 2.17), (pour un tapis 2D, il faudrait plusieurs vidéoprojecteurs). L'utilisation de quatre grands écrans autour du formé (visiocube ou «Cave») aurait l'avantage de permettre au sujet de mieux observer dans toutes les directions et donc de mieux s'orienter. Le déplacement pourrait être conçu différemment, mais le coût de l'affichage et de la création des images serait approximativement quatre fois plus élevé. Cette solution n'a donc pas été envisagée dans un premier temps.

2.4.7.5 Manipulation naturelle d'objets en 3D avec 3ddl

Par rapport à la manipulation de produits avec 6 ddl dans le magasin virtuel, les objets à manipuler ici ne doivent pas être orientés dans toutes les directions. La manipulation d'un traqueur à **deux mains** n'est donc pas obligatoire dans ce cas. On peut donc utiliser un gant de données pour manipuler **avec une seule main** le téléphone, la commande de l'aiguillage, etc (Figure 2.18). Les avantages du gant sont de transmettre en plus des ordres binaires par le mouvement des doigts (prise et pause d'objet, appui sur des boutons, etc.) et d'exploiter un schème très naturel et facilement assimilable. Le biais principal du schème est dans le manque de retour d'effort et de toucher des objets. Ce manque de retour ne permet pas de reproduire la difficulté d'une manipulation réelle. Sa réalisation sur un aiguillage réel, en différé de la formation RV, peut permettre d'atteindre l'objectif pédagogique. Les manipulations d'objets par joystick ou «spaceball» (souris 3D) sont évidemment à proscrire, car ils imposeraient un temps



Figure 2.17 : Immersion visuelle sur grand écran avec rétroprojection, photo SNCF

d'apprentissage trop long par rapport au temps de formation. Des ALC sont ajoutées à cet artefact pour faciliter la manipulation. On proposera des fonctions d'attraction (dès que le gant est proche de l'objet, ils s'attirent) et des fonctions de contraintes cinématiques (par exemple, la poignée de commande de l'aiguillage a un mouvement prédéfini, même si le gant a un mouvement légèrement différent). Une «poche virtuelle» sera affichée pour permettre de se déplacer avec des objets virtuels sans les tenir. Une vidéo d'un prototype de formation en environnement virtuel réalisée par la SNCF, avec notre collaboration, est consultable à l'adresse de notre Centre de Robotique [http2].

2.5 DISCUSSION SUR L'IMMERSION ET L'INTERACTION DU SUJET

2.5.1 APPROCHE INSTRUMENTALE

Notre méthode de conception et d'évaluation d'un dispositif RV, que nous venons d'exposer, repose sur plusieurs idées fortes :

- le découpage en trois niveaux de l'immersion et de l'interaction du sujet ;
- le parallélisme de fonctionnement entre l'homme et l'ordinateur ,
- privilégier l'emploi des schèmes sensori-moteurs par rapport aux métaphores virtuelles et aux schèmes avec substitutions sensori-motrices ,
- l'approche instrumentale de l'interface développée en sciences cognitives,

Cette approche instrumentale est basée sur une problématique anthropocentrique, spécifique au domaine de la réalité virtuelle, fondée sur le concept d'interface compor-



Figure 2.18 : Interface comportementale : gant de données pour manipuler le téléphone, l'aiguillage, etc., photo SNCF

tementale. L'approche instrumentale permet d'analyser la recomposition de l'activité liée à l'utilisation de cette technologie. Un enseignement essentiel de ces travaux et de ceux de l'école francophone est que l'instrument est un moyen terme entre le sujet et «l'objet» de son activité ou de son action [Rabardel, 1995].

Au lieu de prendre une approche instrumentale, nous aurions pu nous appuyer sur une démarche plus orientée «communication» ou «langage». C'est cette approche qui est souvent exploitée par ceux qui sont issus du domaine des IHM en informatique. Comme nous l'avons déjà signalé dans le premier chapitre, nous ne privilégions par ce point de vue car la réalité virtuelle a pour finalité de proposer fondamentalement des activités sensori-motrices et non d'être un outil de communication. Tout en sachant qu'une des applications de la réalité virtuelle est la communication entre personnes, mais ce n'est pas intrinsèquement la finalité de la réalité virtuelle.

Le début des spécifications d'un dispositif RV concerne les I² fonctionnelles : Il faut analyser quelles actions et quels environnements il est utile de conserver de l'environnement réel dans l'environnement virtuel lorsque la finalité du dispositif RV est de *simuler certains aspects du réel*. Si la finalité du dispositif est de *simuler un monde imaginaire*, les références au monde réel ne sont pas indispensables. Dans ce deuxième cas, des schèmes sensori-moteurs du sujet pourront être détournés de leurs activités en environnement réel : par exemple, l'ouverture d'une porte peut établir symboliquement une communication avec une personne, créant une métaphore virtuelle que l'utilisateur devra assimiler. Pour la simulation d'un monde imaginaire, le choix entre les schèmes, les métaphores ou les substitutions sensori-motrices pour les interfaces comportementales est donc évidemment plus libre : quels schèmes peuvent être importés de l'environnement réel dans un environnement virtuel et quels schèmes seront créés spécifiquement pour les environnements virtuels ? L'imagination des concepteurs est sollicitée pour créer de nouvelles immersions et interactions dans un monde imaginaire.

L'étude des incohérences sensori-motrices et de l'adaptation neurophysiologique à celles-ci sont des problèmes ouverts dans le domaine de la réalité virtuelle et sont une des voies de recherche à entreprendre. De même pour les substitutions sensori-motrices, les travaux sont à poursuivre pour connaître celles qui sont pertinentes : par

exemple, la commande vocale peut être utilisée pour se déplacer. Dans quels cas cette substitution est-elle efficace ?

2.5.2 PRÉSENCE ET IMMERSION

Il existe une grande discussion concernant la notion d'immersion ou de présence dans un environnement virtuel, par exemple dans le rapport «Nouvelles interfaces Homme-Machine» [OFTA, 1996]. Certains auteurs font la distinction entre l'immersion physique et la présence cognitive [Burkhardt, 1999]. Pour nous, le terme «présence» est trop large, recouvrant une immersion cognitive où le sujet est soit passif (spectateur), soit actif (acteur ; ce deuxième cas correspondant à la réalité virtuelle). C'est pour bien spécifier cette distinction que nous employons le terme «Immersion et Interaction mentales», qui correspond à la «présence active du sujet». «I² mentales» est donc plus restrictif que «présence».

Il est à regretter que certains concepteurs développent encore une démarche dont l'objectif est de créer un environnement le plus réaliste possible en sollicitant le plus de capacités sensorielles humaines (auditive, tactile, visuelle, proprioceptive, etc.). Il est nécessaire de se défaire du paradigme de duplication du réel qui revient à faire de l'ancien avec le nouveau alors qu'il s'agit de concevoir de nouvelles réalités.

Comme bien des auteurs, on peut affirmer que tout sujet fait la différence entre une présence dans un monde réel par rapport à une présence dans un monde virtuel, à moins d'être endormi ou de ne prêter aucune attention à ce qui l'entoure. Un de nos postulats est le suivant : quelle que soit l'évolution des techniques, il y aura toujours des différences sensori-motrices entre le monde réel et un monde virtuel, même limité aux I² fonctionnelles que l'on veut simuler. Certaines de ces différences sont évidentes pour l'utilisateur, d'autres sont perçues inconsciemment : par exemple, la vision en espace virtuel ne sera jamais identique à celle en espace réel : les images, même en relief, sont projetées sur des écrans. La différence avec la vision réelle est donc détectée inconsciemment par nos deux cristallins qui accommodent sur ces écrans et non dans un espace tridimensionnel comme en vision réelle (pour plus de détails, voir le chapitre sur les interfaces visuelles). Il y a donc toujours des différences sensori-motrices entre mondes réel et virtuel que l'utilisateur peut toujours percevoir s'il s'en donne la peine. Mais quand peut-il donc se trouver présent dans le monde virtuel qu'on lui propose ? Cette question renvoie à l'interprétation philosophique du monde réel et à la notion de présence dans celui-ci. Comme nous l'avons déjà indiqué, le lecteur intéressé peut consulter l'article de Mantovani et Riva [Mantovani, 1999] sur les différentes présences (physique, subjective et objective) d'un sujet dans un monde réel ou virtuel.

Sans développer toute cette problématique dans ce chapitre, que pouvons nous affirmer concrètement ? Souvent des concepteurs posent la question suivante aux premiers utilisateurs d'un dispositif RV :

«Êtes vous bien (présents) dans le monde que nous simulons (sous-entendu comme dans le monde réel) ?»

Par exemples, on aurait pu poser les questions suivantes : «êtes vous bien (présents) dans un magasin ?», «êtes vous bien (présents) sur des voies ferrées ?». Pour moi, cette question n'a aucun sens et est inappropriée. *Cette question est un piège* que l'on pose à

l'utilisateur. Il se trouve devant un dilemme : «j'ai envie de répondre oui pour faire plaisir au concepteur car j'ai bien perçu la simulation (la copie) d'un monde réel, mais par ailleurs, j'ai envie de répondre non car je perçois bien que je suis dans un monde virtuel, construit artificiellement !». La question n'a pas de sens car le Virtuel est toujours détectable au niveau sensori-moteur, donc aussi automatiquement au niveau mental¹². Après une immersion dans un environnement virtuel, il est inapproprié de poser ce type de question aux utilisateurs du dispositif RV. Pour le lecteur peu convaincu, il peut réfléchir au cas du cinéma, où il n'y a pas d'interaction mais une immersion. Le spectateur va s'immerger dans le scénario du film, va se projeter dans le personnage principal. Le spectateur va alors ressentir divers sentiments, agréables ou effrayants. Et pourtant, si vous interrogez celui-ci, pendant ou après le film, il est bien conscient que l'histoire n'est pas réelle, qu'il ne regarde que des images projetées sur un écran. Mais le but du réalisateur a été atteint : immerger les spectateurs dans un scénario.

Il faut mieux se poser la question suivante :

«Est-ce que le comportement du sujet dans l'environnement virtuel est similaire à celui dans l'environnement réel, par rapport aux I² fonctionnelles que l'on souhaite obtenir ?»

Et la réponse à cette question n'est pas facile à traiter par les concepteurs. Dans la plupart des cas, l'utilisateur ne peut y répondre par lui-même (par exemple : votre comportement d'achat est-il similaire à celui que vous avez dans un magasin réel ?). Lorsque l'on sait le temps qu'il faut prendre pour répondre à ce type de question, souvent à l'aide de campagnes de tests, faites par des experts en ergonomie et en psychologie, il est inapproprié de la poser aux utilisateurs ! Nous pouvons, par contre, poser des questions simples aux utilisateurs sur leurs activités en environnement virtuel, ce qui nous permet ensuite de juger de la pertinence du dispositif : permet-il atteindre les I² fonctionnelles désirées ?

Ces quelques remarques sont présentées ici pour permettre à chacun une réflexion sur les capacités et les problématiques induites par l'immersion et l'interaction d'un sujet dans un monde virtuel. D'autres questions sont à débattre, en particulier la question de la prise en compte du niveau émotionnel du sujet.

2.6 PERSPECTIVES ET CONCLUSIONS DE NOTRE APPROCHE

A partir de nos propres expériences, nous concluons sur la validité de notre démarche de conception et d'évaluation d'un dispositif en réalité virtuelle. Nous proposons à tout concepteur cette démarche, ce qui permettra de la valider et de l'améliorer. L'accumulation des conceptions et des évaluations de dispositif RV doit permettre progressivement d'avoir des connaissances suffisantes pour la réalisation des principales PCV. Sans ambition exhaustive, nous avons en perspective de faire une taxonomie des principales Primitives Comportementales Virtuelles et un catalogue de leurs solutions élaborées dans la communauté scientifique. Ces dernières peuvent être inventoriées en précisant si elles sont basées sur des schèmes, des métaphores ou des substitutions sensori-motrices. Pour chaque solution, des Aides Logicielles Comportementales améliorant les I² mentales sont à concevoir et à évaluer. Toute cette recherche sur les concepts de

¹²Pour être plus précis, le Virtuel est détectable pour les personnes ayant un fonctionnement mental correct, pas pour un sujet sous l'emprise de l'alcool ou de la drogue ou n'ayant pas tous ses esprits. ...Mais ce n'est pas à ceux-ci que le concepteur pose la question, pour se rassurer de la pertinence de son travail !

base en réalité virtuelle ne peut être entreprise que par une collaboration pluridisciplinaire, qui est naissante dans la communauté française.

En parallèle à cette recherche, les spécialistes en réalité virtuelle doivent poursuivre leurs travaux sur toutes les interfaces comportementales et sur leur intégration dans des systèmes RV. Ces mises en œuvre sont plus ou moins délicates, certaines étant assez aisées grâce au recul de ces dix dernières années. Certaines interfaces visuelles sont bien maîtrisées. Par contre, il y a encore beaucoup de recherche à accomplir avec les interfaces à retour tactile et d'effort. Nous allons donc dans les prochains chapitres présenter les différentes interfaces exploitables en réalité virtuelle. Mais avant, deux chapitres indispensables sont consacrés au «comportement de l'homme dans les environnements réel et virtuel» et «aux sens et aux réponses motrices de l'homme», puisque nous souhaitons réaliser une Immersion et une Interaction sensori-motrices.

Avant d'aborder toutes les différentes techniques des interfaces comportementales, nous pouvons affirmer (mais sans le démontrer) : que la réalité virtuelle renferme en elle-même de grandes potentialités, mais que plus on veut en exploiter, plus les difficultés sont importantes, et cela d'une façon exponentielle !

2.7 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Berthoz, 1998] A. Berthoz, «*Le Sens du Mouvement*», éditions Odile Jacob, (345 p.), ISBN 2-7381-0457-6, 1998.
- [Burkhardt, 1999] J.M. Burkhardt, D. Lourdeaux, P. Fuchs, (1999) «*Conception d'un système de RV pour la formation des agents des conduite aux opérations en milieu ferroviaire*», Actes des journées réalité virtuelle et cognition, 14-15 Décembre, ENST, Paris, 123-132.
- [Chance, 1998] S. Chance, F. Gaunet, A. Beall, «*Locomotion Mode Affects the Updating of Objects Encountered During travel*», Presence, vol 7, N° 2, MIT, 168-178, 1998.
- [Fréjus, 1996] M. Frejus, A. Drouin, «*réalité virtuelle et processus cognitifs*», EDFF-DER, Université Paris 5, René Descartes, 1996.
- [Fuchs, 1999a] P. Fuchs, «*Immersion et interaction naturelles dans un environnement virtuel*», Actes des journées réalité virtuelle et cognition, p 71-76, 14-15 Décembre, ENST, Paris.
- [Fuchs, 1999b] P. Fuchs, F. Nashahibi. D. Lourdeaux, «*A theoretical approach of the design and evaluation of a virtual reality device*». (1999) , Actes du colloque International de Laval «*Réalité virtuelle et prototypage* », 3-4 Juin, 11-20, Laval.
- [Gibson, 1979] J. Gibson, «*The ecological approach to visual perception* », London, Lawrence Erlbaum Associates, 1979.
- [Lourdeaux, 1999] D. Lourdeaux, P. Fuchs, J. M. Burkhardt and F. Bernard, «*Relevance of an intelligent tutorial agent for virtual reality training systems*», à paraître, International Journal of Continuous Engineering and Life-Long Learning.
- [Maman, 1998] D. Maman, «*Recalage de modèles tridimensionnels sur des images réelles*». Thèse de doctorat de l'Ecole des Mines de Paris, 1998.
- [Mantavoni1999] G. Mantavoni and G. Riva, «*Real Presence : How Difference Ontologies Generate Different Criteria for Presence*», Telepresence, and Virtual Presence, revue Presence, MIT, Vol. 8, N°5, 540-550, 1999.

- [Morineau, 2001] T. Morineau, P. Chedmail et M. Parenthoën, « *An affordance-based model to support simulation in Virtual Environmenrt* », Actes du congrès VRIC, Laval Virtual 2001, p 19 – 25, 2001.
- [OFTA, 1996] OFTA, «*Nouvelles Interfaces Homme-Machine*», série ARAGO, Observatoire Français des Techniques Avancées, 1996.
- [Piaget, 1979] J. Piaget et N. Chomsky, « Théorie du langage et théories de l'apprentissage », éditeur Seuil, 1979.
- [Rabardel, 1995] P. Rabardel, «Les hommes et les technologies, une approche cognitive des instruments contemporains», Paris, Colin, 1995.

Adresses Web des entreprises et centres de recherches :

- [http1] [http ://www.simteam.com](http://www.simteam.com)
- [http2] [http ://www-caor.ensmp.fr/Fr/Recherche/RV-RA](http://www-caor.ensmp.fr/Fr/Recherche/RV-RA)

Deuxième partie

**Les interfaces
comportementales**

3 L'ÊTRE HUMAIN DANS LES ENVIRONNEMENTS RÉEL ET VIRTUEL

Jean-Paul Papin

3.1 INTRODUCTION

Comme cela a été dit dans le premier chapitre, la réalité virtuelle a pour but de recréer artificiellement un environnement réel ou imaginaire. Dans celui-ci un humain est immergé par l'intermédiaire d'interfaces en contact avec les organes sensoriels et moteurs de son organisme pour mener une activité particulière. Cependant l'utilisation de toute nouvelle technique et cela depuis le premier usage de la pierre jusqu'à la réalité virtuelle en passant par la voiture, l'avion, la micro informatique ou le téléphone mobile, entraîne parfois des troubles (physiques ou psychiques) et toujours des nécessités d'adaptations humaines qu'il faut prendre en considération. En effet, même si l'outil et la tâche accomplies avec celui-ci sont le plus parfaitement adaptés à son utilisateur, il est rare que ce dernier puisse l'utiliser sans un minimum d'apprentissage c'est à dire d'adaptation de l'homme à son système d'activité. De même, il est souvent nécessaire de faire une sélection des utilisateurs pour des systèmes complexes car tout le monde n'a pas forcément toutes les aptitudes requises pour se retrouver dans n'importe quelle situation. Enfin de manière plus récente, l'emploi des outils pose, outre les problèmes de médecine du travail et de sécurité, des problèmes éthiques et juridiques. Dans le cadre du développement et de l'utilisation de systèmes de réalité virtuelle, une approche ergonomique permet de bien connaître les différents environnements réels ou ambiances réelles. Comment sont ils perçus par l'homme ? Comment l'homme peut-il agir sur eux ? Quels sont les risques éventuels encourus si les ambiances dépassent certaines limites ?

3.2 L'ÊTRE HUMAIN, SON ENVIRONNEMENT NATUREL

Naturellement l'être humain, du fait de sa nature de mammifère terrestre, aérobique, homéotherme, soumise à des biorhythmes et sociale, n'est adapté qu'à une biosphère restreinte représentant sa niche écologique :

- terrestre l'homme est soumis à une force attractive (pesanteur ou gravité de $9,8 \text{ m/s}^2$). Il supporte au niveau de la mer une pression de 760 mmHg ou d'une atmosphère ;
- aérobique, il dispose au niveau de la mer d'environ 21 % d' O_2 autour d'une pression partielle de 150 mm Hg. L'oxygénation du sang est possible sans adaptation particulière pour des pressions partielles variant entre 110 et 300 mm Hg ;
- homéotherme, il a une température centrale autour de 37°C . Le maintien de celle-ci est obtenu par une thermogénèse ou une thermolyse qui dépend des possibilités d'échanges thermiques avec le monde environnant. Ces échanges sont minimaux pour des valeurs qui correspondent à la neutralité thermique et qui s'accompagnent d'une sensation de bien être optimale. Pour un sujet nu, l'air doit être aux environs de 28°C , l'humidité à 20 %, la vitesse du vent à 0,2 m/s et la pression de vapeur d'eau supérieure à 1,5 KPa. Tout ceci au repos, il justifie son appellation de singe nu tropical.

L'exposition à des environnements ne répondant pas aux limites décrites ci-dessus montre que l'être humain est limité. Il n'est pas capable, par exemple, de survivre plus de quelques minutes dans des environnements sans oxygène, à des températures très élevées ou très basses, à des accélérations élevées, à l'exposition de produits chimiques particuliers.

De plus, comme tous les animaux, il est doté de réactivité en terme de plaisir et de désagrément. A des sollicitations physiques ou sociales (vibrations, bruit, isolement, confinement) excessives mais non mortelles, il peut répondre par des lésions et des troubles plus ou moins définitifs non seulement sur le plan organique, mais aussi sur le plan psychologique.

Enfin, être pensant, capable, en plus de raisonner, de faire des choix et de juger en terme de bien et de mal, il se caractérise par une soif infinie de dépassement et il cherche comme dans le mythe de Prométhée à rompre les chaînes qui le retiennent à sa niche écologique. Pour cela, il fabrique des outils ou des systèmes de protection lui permettant de s'adapter même à des milieux hostiles. Il a pu ainsi élargir son écoumène, c'est-à-dire l'espace habitable à la surface de la terre. Par exemple, il a fait aussi bien que les animaux marins en se déplaçant et même en vivant sous l'eau, que les animaux aériens en se déplaçant dans les airs ou que les animaux vivant sous terre en descendant dans ses entrailles. Il a fait mieux que toutes les autres espèces en quittant la planète pour aller dans l'espace et même sur d'autres mondes (Figure 3.1). De plus comme le dit Piéron dans la conclusion de son ouvrage sur les sensations et ceci dès les années 50 « l'homme ne s'est pas contenté des informations que nos sens nous fournissent sur les événements extérieurs, il a inventé des systèmes récepteurs variés, transformant des sources d'informations directement irrecevables en modalités stimulatrices efficaces ». Comme par exemple, « l'amplification spatiale multipliant l'acuité discriminatrice (téléscope, microscope)...et même, rendu visible ou audible des phénomènes physiques qui ne sont ni lumineux ni sonores comme l'électricité et les actions magnétiques ».

Maintenant, il veut conquérir des espaces virtuels en s'immergeant dans des mondes qu'il crée lui-même. D'ailleurs ce désir est ancien puisque déjà chez des tribus primitives, il était utilisé des substances hallucinogènes qui permettaient de vivre des situations extraordinaires avec une sensation de présence dans ces mondes rêvés qui faisaient croire que ces situations avaient été réellement vécues. La différence avec les mondes virtuels, c'est que ces hallucinations ne sont pas contrôlées par le sujet, même si elles sont vraisemblablement en partie construite par l'inconscient.

Avant d'aborder cette immersion dans les mondes virtuels, il est bon de s'arrêter un peu sur l'environnement naturel de l'être humain et sur les interactions qu'il entretient avec lui. Si l'environnement naturel pour les Grecs de l'antiquité était constitué de quatre éléments l'air, l'eau, la terre et le feu, cette vue est un peu dépassée, mais il n'en reste pas moins vrai que l'environnement entourant l'humain peut être décomposé en plusieurs ambiances. Chacune correspond à des sensations particulières qui sont captées par des organes spécifiques à chacune de celles-ci. Comme le dit Piéron [Piéron, 1964], ces ambiances ont été réifiées à partir des propriétés de chaque organe sensoriel. C'est ainsi, par exemple, que l'ambiance lumineuse correspond à la partie des radiations électromagnétiques sensibles pour l'œil. Il en va de même du son qui correspond à ce que l'on entend ou des odeurs qui correspondent à ce que l'on peut sentir. D'autres ambiances sont en interaction avec le corps de manière plus globale. C'est le cas de l'ambiance thermique (correspondant au feu des anciens), fortement liée à l'ambiance atmosphérique (correspondant à l'air des anciens) ou de milieux que l'homme a pé-

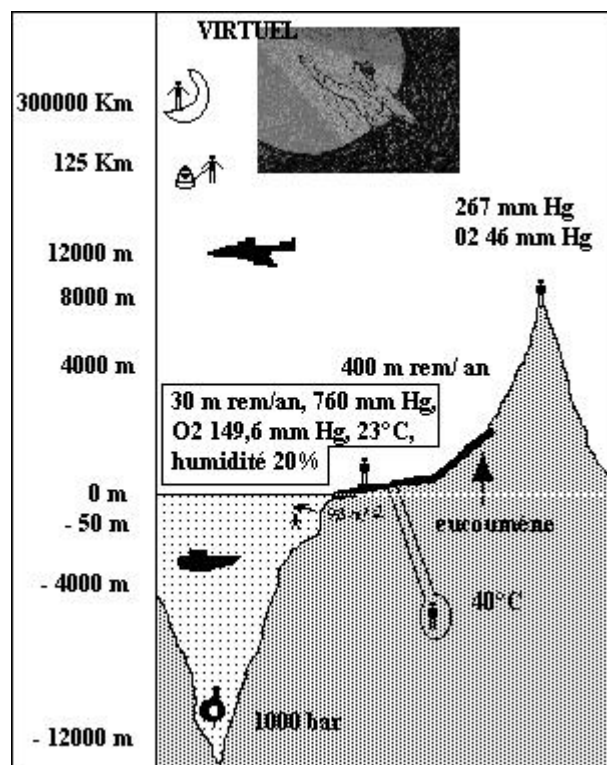


Figure 3.1 : L'expansion de l'homme

nétrés comme l'eau (quatrième élément des anciens) et l'espace (cinquième élément pris parfois en compte chez les anciens) qui pose tout le problème des ambiances en apesanteur. Pour décrire les différentes ambiances et les capteurs physiologiques, il est possible de prendre l'exemple d'un joueur de tennis [Papin, 2000b], c'est-à-dire un être humain en action dans un environnement qui est en interaction avec lui (Figure 3.2).

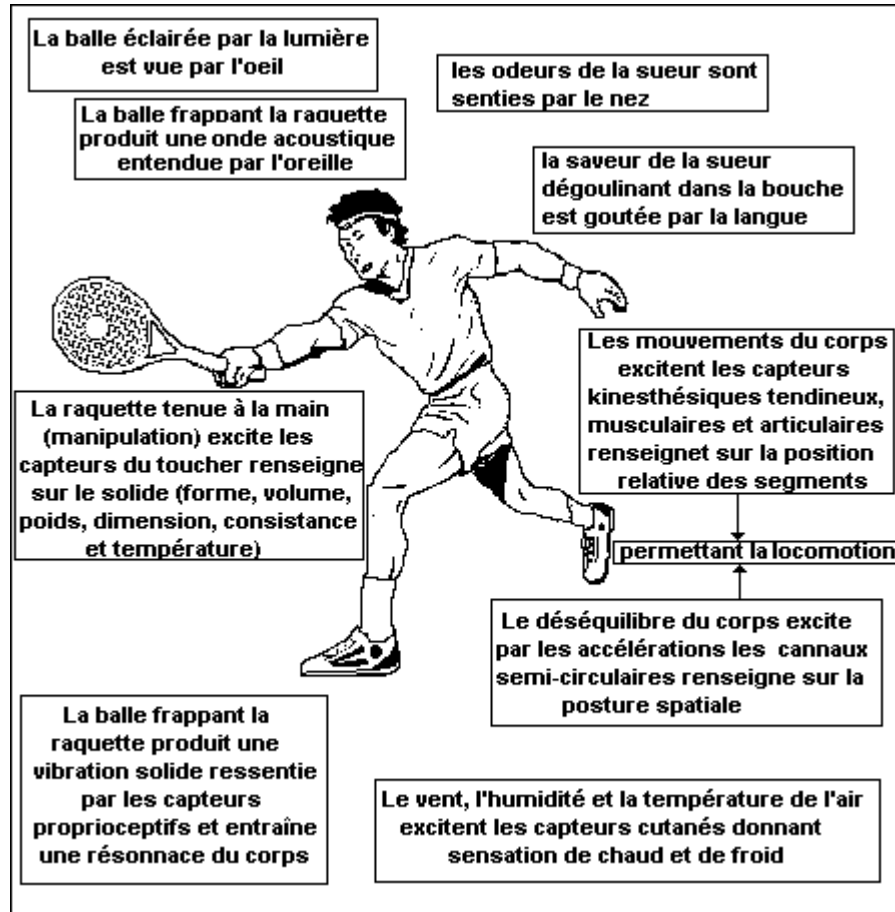


Figure 3.2 : Les différentes ambiances réelles et les organes des sens

- la lumière (radiations électromagnétiques situées entre 380 et 750 nanomètres éclaire (ambiance lumineuse) la balle qui est vue (vision) par un capteur sensoriel : l'œil ;
- la balle frappant la raquette provoque un certain nombre de phénomènes mécaniques correspondants aux ambiances mécaniques :
- une vibration qui se propage dans l'air sous la forme d'ondes acoustiques situées entre 20 et 20000 Hz (ambiance acoustique) qui est entendue (audition) par un capteur sensoriel, l'oreille ;
- cette vibration produite par le solide balle sur le solide raquette entraîne un choc et des vibrations solides (ambiance vibratoire) captées par la peau et les senseurs kinesthésiques profonds avec une résonance osseuse ;

- la raquette tenue à la main est sentie dans sa forme, son poids, la texture de sa surface, sa température (ambiance de contact) par les capteurs sensoriels cutanés du sens du toucher ;
- les déplacements du joueur pour rattraper la balle produisent des accélérations (ambiance gravifique) captées par le système vestibulaire et les organes proprioceptifs ;
- les actions menées en retour des informations kinesthésiques et vestibulaires permettent la manipulation de la raquette et la locomotion du joueur ;
- la chaleur, l'humidité, la vitesse du vent (ambiance climatique) sont perçues par la peau mais aussi par les régulateurs thermiques profonds. La sueur du joueur dégoulinant sur la peau donne la sensation de liquide ;
- la sueur produisant des molécules volatiles (ambiance chimique odorante) est sentie (odorat) par le nez ;
- la sueur dégoulinant sur le visage est avalée, les molécules en solution (ambiance chimique gustative) sont goûtées par la langue ;
- les réactions des spectateurs (ambiance sociale), le score obtenu (performance de la tâche), les stratégies mises en jeu (intellect), la motivation, l'implication, les réactions comportementales (affect) constituent l'ambiance psychologique dans laquelle s'inscrit l'activité de prise d'informations, de traitements intellectuels et de réponses motrices, c'est-à-dire la tâche.

A chacune des ambiances et des situations, les réponses de chacun peuvent être différentes d'un sujet à l'autre mais aussi pour un même sujet d'une fois sur l'autre. Enfin, si les sollicitations sont trop intenses, elles peuvent avoir des effets secondaires organiques plus ou moins définitifs sur chaque organe sensoriel (éblouissement par une lumière trop vive, traumatisme sonore par un bruit intense), mais aussi sur d'autres (la lumière peut brûler la peau ou entraîner une déshydratation par son pouvoir calorifique). Les effets secondaires peuvent être aussi d'ordre psychologique (névrose post traumatique). Tout cela pose des problèmes de santé mais aussi moraux et juridiques. Face à ces problèmes l'ergonomie, la médecine, la psychologie, le droit et l'éthique répondent par des réglementations et des recommandations basées sur la connaissance des réactions, des comportements et des capacités physiques, physiologiques et psychiques de l'être humain. Ceci est développé dans le chapitre «Les effets sur l'homme des environnements» et le chapitre «Ethique, droit et réalité virtuelle». Maintenant nous allons aborder quelques notions concernant les facteurs humains.

3.3 LES FACTEURS HUMAINS ET LE FACTEUR HUMAIN

3.3.1 DÉFINITIONS

«Les facteurs humains recouvrent l'ensemble des faits scientifiques concernant l'homme sain, sa diversité et sa variabilité. Ils incluent les aspects biologiques, physiologiques, psychologiques et psychosociaux. Ils recouvrent les comportements en situation et environnements naturels ou artificiels. Ils comportent les principes et les applications dans les domaines de l'ergonomie, de l'aptitude et de la sélection du personnel, de la tolérance, de l'adaptation comme de la performance à l'emploi, et de l'évaluation de la performance humaine, dans les conditions de fonctionnement normal ou dégradé du système. Ils s'intègrent enfin dans une logique de systèmes» [Curé, 1996]. Quant au facteur humain, au singulier, il s'attache à l'individu et à sa composante psychologique.

Pour aborder les facteurs humains, il est possible d'adopter la démarche de l'ergonome qui s'intéresse à l'homme au travail, plongé dans une situation particulière. Les facteurs d'ambiances mais aussi les exigences de la tâche, voire l'organisation du travail peuvent elles-mêmes représenter une ambiance nuisante et même extrême. Le spécialiste de réalité virtuelle doit s'intéresser à l'homme, plongé dans un monde qu'il a fabriqué, en vue de le faire agir et réagir comme dans le monde réel. Pour cela il se doit d'avoir la même démarche systémique qui lui permettra de prendre en compte le plus grand nombre de facteurs qui peuvent interagir.

3.3.2 DÉMARCHE SYSTÉMIQUE

Un système, en plus de pouvoir être nommé, se caractérise par sa faculté de fonctionner de manière autonome pendant une durée déterminée, dans un environnement défini et avec un objectif précis. Il est possible de décrire toute activité humaine en terme de système. Ce dernier comprend des hommes, des matériels, des environnements et des réglementations. Dans ce système l'homme agit (directement ou par l'intermédiaire de matériels), selon des règles plus ou moins formelles dans un but donné et pendant une durée déterminée dans un environnement physique et social donné avec lequel il interagit. Cette approche permet de voir la similitude qui existe entre un système de travail et un système de réalité virtuelle qui peut d'ailleurs être utilisé à des fins professionnelles. Dans cette démarche systémique, il apparaît que l'homme en action présente de nombreuses interactions au niveau de ses différents organes sensoriels et moteurs mais aussi, à travers eux, avec les différents éléments de son environnement. Ces éléments présentent, eux aussi, de nombreuses interactions entre eux. C'est pourquoi, il semble intéressant de s'attarder dans un premier temps sur l'homme en action.

3.3.3 L'HOMME EN ACTION

Dans la majorité des cas, l'activité humaine est mise en jeu pour aboutir à un résultat escompté (gagner de l'argent, augmenter ses connaissances, se détendre). Dans tous les cas, au départ, il y a donc une opération intellectuelle de construction d'un projet. L'origine de ce projet peut être personnelle ou émanée d'un supérieur hiérarchique. Dans ce dernier cas il est nécessaire, pour le sujet qui doit agir, qu'il fasse sien ce projet. Pour mener à bien ce projet l'homme doit prendre de l'information sur le monde extérieur, traiter celle-ci intellectuellement, prendre une décision et agir de manière motrice par des gestes ou des ordres verbaux. Il faut signaler que la prise d'information s'accompagne très souvent d'une action motrice. Par exemple la prise d'information visuelle s'accompagne d'une stratégie d'exploration qui met en jeux non seulement les muscles des globes oculaires mais aussi ceux de la tête voire ceux de tout le corps. Nous allons aborder maintenant la problématique de la prise d'information.

3.3.3.1 La prise d'information

La prise d'informations ne correspond pas à une réaction automatique mais à une action élaborée même si parfois elle reste inconsciente. Ce qui arrive au cerveau est une image du monde environnant au sens «*vision intérieure plus ou moins exacte d'un être ou d'une chose*» (Petit ROBERT). Cette représentation du monde que chacun se fait est

propre à chacun, elle dépend de son histoire personnelle. Elle est réalité pour chacun d'entre nous. Cependant elle peut être faussée, parce que les organes sensoriels sont imparfaits ou parce que, par exemple au niveau visuel, on ne voit que ce que l'on veut voir et cela en toute bonne foi. A titre d'exemple, voici une phrase :

«les arbres de la
la forêt».

Il est classique de ne pas percevoir la répétition de «la».

Sur le plan auditif il en va de même, un message verbal peut avoir été entendu et même répété par un sujet et pourtant il peut ne pas être pris en compte. Ce type de situation se rencontre lorsque la charge physique est importante. Cas d'un élève pilote qui, au cours d'un décollage de Djibouti, le moniteur ayant dit : «il y a un camion en bout de piste, arrête-toi», répète : «oui il y a un camion en bout de piste, il faut s'arrêter» et continue. Après reprise en catastrophe par le moniteur et décollage avec évitement, celui-ci demande à l'élève ce qui vient de se passer. Ce dernier ne se rappelle de rien et accepte son erreur qu'après avoir vu et entendu la bande d'enregistrement ! Dans ce cas plusieurs facteurs ont pu intervenir. Il s'agit d'un élève, c'est son premier décollage de Djibouti. La charge thermique est importante (plus de 45 °C avec une humidité supérieure à 50 %). Il existe une réverbération sur la piste. Le bruit du moteur en pleine puissance est malgré le port du casque non négligeable.

Ces exemples montrent la difficulté qu'il peut y avoir à saisir la réalité d'une prise d'informations. Ce n'est pas parce qu'une information est présente, utile et stimulant l'organe sensoriel concerné qu'elle a été forcément prise en compte de manière opérante (c'est-à-dire s'il y a eu perception). Les travaux sur l'analyse des stratégies visuelles de prises d'informations visuelles mettent en évidence que nous avons la certitude qu'une consultation du regard sur un objet est bien une prise d'information que lorsqu'il est possible de constater qu'une action logique lui succède. Cela pose une interrogation dans le cadre des immersions par l'intermédiaire de casques dans des espaces visuels virtuels. En effet, l'enregistrement des consultations du regard sur les objets du monde virtuel est techniquement très difficile, voire aujourd'hui impossible avec les matériels disponibles. Classiquement la prise d'information visuelle est analysée à partir d'enregistrements de la direction du regard basés sur les propriétés du reflet cornéen. Une caméra de télévision miniaturisée, solidaire de la tête du sujet enregistre en permanence le paysage situé en avant du sujet dans son plan frontal (il s'agit de l'ergorama). Par un jeu de systèmes optiques on superpose sur les images ainsi obtenues un index indiquant la projection de l'axe du regard sur l'ergorama. Pour déterminer cet axe, on éclaire un œil avec une lumière infrarouge dont le support est solidaire de la tête, grâce à un miroir semi-réfléchissant situé devant l'œil. On renvoie l'image de la lumière qui se forme sur l'œil dans le système optique. Un tel dispositif est difficilement compatible avec le port d'un casque de réalité virtuelle. Il serait éventuellement possible d'utiliser l'enregistrement des mouvements oculaires par des électrodes et par calcul de déterminer l'axe du regard et sa projection sur l'image virtuelle, celle-ci étant asservie au mouvement de la tête. Il ne semble pas à ce jour qu'une telle démarche ait été entreprise. Pourtant cela pourrait être riche d'enseignements en référence aux résultats obtenus dans l'étude des stratégies de prises d'informations avec champ visuel réduit de pilotes d'hélicoptère (Figure3.3). Dans cette étude la réduction du champ visuel, à l'aide de caches ne permettant une vision que dans un champ de 40° ou de 60°, soit en vision monoculaire, soit en vision binoculaire, a permis de déduire la nature des informations perçues en vision centrale ou en vision périphérique. De plus ces enregistrements ont permis de vérifier que la vision périphérique sert dans l'appréciation de la vitesse (grâce au défilement), de la hauteur. Les pilotes vont chercher en vision centrale (sur la planche

de bord), des informations qu'habituellement ils prennent en vision périphérique. De même dans une autre étude sur des pilotes de chasse, il a été possible de montrer que le temps de consultation d'un objet permettant une perception et donc une prise d'information opérante joue un rôle important. C'est ainsi qu'il faut, à un pilote d'avion consultant sa planche de bord, 100 ms pour dire qu'une aiguille sur un cadran n'a pas bougé, 200 ms pour dire qu'il y a eu une variation angulaire, 250 ms pour en donner la valeur chiffrée et 450 ms pour lire un chiffre.

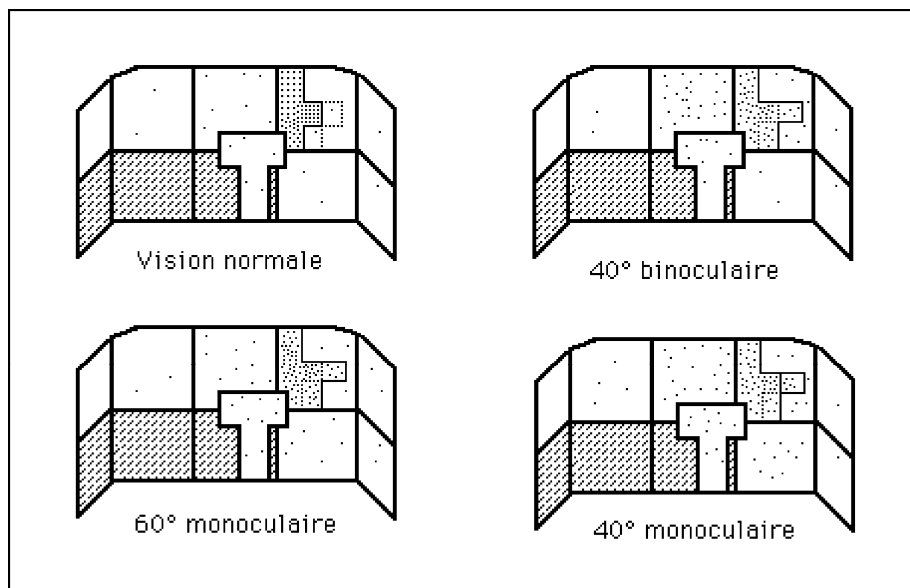


Figure 3.3 : Importance relative des consultations visuelles, au cours d'un vol à vue en hélicoptère, en fonction du champ visuel. Chaque point représente 1 % des consultations

Une telle démarche est donc très riche. Elle l'est d'autant plus que l'expérience nous a montré que l'interrogatoire d'opérateurs sur la nature des informations réellement utilisées au cours d'une tâche était pauvre, voire erroné. Au cours de telles enquêtes, le dépouillement des entretiens fait ressortir parfois que, sur l'ensemble d'une population donnée, toutes les informations disponibles servent à tout moment. De plus, très souvent, les sujets récitent ce qu'ils auraient dû faire (tâche formelle) et pas forcément ce qu'ils ont fait (tâche informelle). Cette remarque est valable pour la prise d'informations mais aussi pour les opérations intellectuelles ou pour les actions gestuelles.

Un exemple peut illustrer ce propos. En aéronautique, dans le cas d'une attaque au sol avec pénétration à basse altitude, il est recommandé au pilote de conserver une vitesse constante de 500 nœuds, d'où la proposition dans cette phase de privilégier la présentation de la vitesse. Une analyse du pilotage dans cette situation fait ressortir que le pilote ne regarde jamais son indicateur de vitesse et que pourtant il maintient sa vitesse constante. L'entretien avec le pilote permet de s'apercevoir que ce dernier utilise d'autres sources pour apprécier la vitesse. S'il entend un sifflement il est au-dessus de 500 nœuds ; si ça vibre, il est en dessous. S'il n'y a ni sifflement ni vibration, il est à une vitesse correcte. Il est remarquable d'ailleurs de constater que les deux éléments utilisés sont considérés habituellement comme des nuisances que le concepteur va s'efforcer

de supprimer sans se rendre compte que par-là même il augmente le travail visuel qui est pourtant déjà très important. Ce premier exemple montre bien que l'on puisse se tromper sur la pertinence d'un indice et que le traitement de l'information est une opération capitale.

3.3.3.2 *Traitement de l'information*

Nous venons de voir que la prise d'information est déjà un traitement intellectuel. Il s'agit en fait d'une mise en forme, de la création de percepts qui vont être utilisés pour mener à bien une action projetée. Ces percepts permettent d'autre part de construire le projet, de chercher les moyens de le réaliser et de décider de mettre en œuvre ces moyens. Ce sont toutes ces opérations qui sont à considérer. Il s'agit des fonctions intellectuelles qui vont faire appel à la cognition, c'est-à-dire à l'utilisation des connaissances acquises dans l'immédiat ou dans le passé. Le «cognitivisme» à la mode en psychologie actuellement correspondant à l'étude de ce concept.

A partir des images ainsi formées le cerveau humain peut imaginer, c'est-à-dire former de nouvelles images qui ne sont pas représentation du réel. On peut ainsi développer des concepts, des idées, des projets. Ces représentations mentales peuvent être symbolisées, c'est-à-dire qu'un objet réel peut être remplacé par un autre objet, le plus souvent plus simple et identifiant de façon indéniable l'objet réel. L'assemblage de ces codes permet de créer des objets imaginaires ou chimères qu'il est possible de coder pour les identifier en tant que concepts. Donc l'imagination permet de conceptualiser non seulement un objet possible, une situation future qui pourront être perçus (il s'agit alors de l'élaboration d'un projet avec anticipation du résultat tangible escompté) mais aussi d'exprimer une idée, une idéologie, c'est-à-dire un objet purement intellectuel tel que la liberté, la justice...

Il devient possible alors de poser des problèmes, de chercher des solutions et de les résoudre. L'homme est capable d'élaborer et de programmer.

L'analyse de ces capacités n'est pas simple. Le cerveau analyse le cerveau avec seulement le cerveau. Nous sommes dans la situation d'un outil qui doit s'analyser lui-même. S'il existe des instruments rendant compte de l'activité du cerveau, ils ne sont pas encore assez élaborés pour déterminer le contenu de la nature des opérations en cours. Cependant actuellement il devient déjà possible d'appréhender la nature de certaines opérations en cours. Ainsi l'électroencéphalogramme permet de déterminer l'état d'activation du cerveau et de juger s'il est dans un état de veille suffisant pour effectuer tel ou tel type d'activité. Les techniques d'analyse par thermographie des différentes zones cérébrales permettent d'aller encore plus loin. Cependant la technique de base de l'analyse des opérations intellectuelles passant par l'analyse de contenu du discours d'opérateurs reste un bon outil. Dans certains cas il peut s'agir du discours écrit ou du dépouillement d'un film de la suite des graphismes effectués au cours de la résolution d'un problème. Dans l'activité de peinture, l'analyse du film pris pendant la réalisation d'un tableau permet de mettre à jour les processus de création de l'artiste.

Cette démarche permet d'avoir une idée du degré de complexité des opérations nécessaires à la mise en œuvre d'une réponse, c'est-à-dire de l'acte qui va entraîner une modification de la situation extérieure. On peut maintenant aborder le problème des actions ou plus exactement celui des sorties du système homme puisque les éléments

de perception et d'opérations intellectuelles sont eux-mêmes des actions.

3.3.3.3 Actions

Sous ce vocable comme nous venons de le dire, il s'agit de s'intéresser aux sorties et nous allons donc détailler celles-ci. Nous citons qu'une des sorties possible est la prise d'information active, pour l'éliminer de ce chapitre puisque déjà largement abordée. Deux autres sorties sont possibles : la parole et les gestes. Une remarque s'impose, du fait qu'il s'agit de sorties elles sont destinées à exciter les entrées d'un système. Elles peuvent être captées et effectivement de nombreux capteurs existent : enregistrement sur magnétophone pour la parole, enregistrement sur film pour les gestes mais aussi enregistrement des forces déployées ou même réponses des organes matériels d'interface. Plusieurs méthodes sont utilisées séparément ou conjointement en fonction des situations étudiées. A titre d'exemple nous avons utilisé des caméras de télévisions miniaturisées à très bas niveau de luminance pour enregistrer le comportement des équipages de char. Ces caméras font 75 g et ont le volume d'une petite boîte d'allumettes.

La parole

L'analyse peut se faire à différents niveaux, en particulier au niveau du contenu mais aussi à celui de la forme. Au niveau du contenu, on peut distinguer l'émission d'ordre donné, de demande d'information, d'une réponse à une question ou d'un commentaire. En effet la parole ne peut pas se ramener à un simple message nécessaire à l'action. Elle a d'autres fonctions. L'une est liée à la nature d'être social de l'homme, elle permet d'instaurer un climat de convivialité garant d'une cohésion du groupe par l'échange d'informations concernant la personne elle-même, ses idées, ses goûts ou sa famille. Une autre est de maintenir un niveau de vigilance donné. Au niveau de la forme, le discours peut être bref ou délayé, le ton cassant ou enjoué. Ces éléments ont un rôle renforçant de l'importance que le locuteur accorde à l'émission.

Pour illustrer ceci, voici des résultats concernant les échanges verbaux au cours d'un vol en avion de transport de longue durée. Première constatation il n'existe aucun silence d'une durée supérieure à 3 minutes. Au cours de 8 H de vol on ne parle pas que travail et on n'échange pas par ce canal que de l'information nécessaire au vol. Ce type d'activité est en fait le plus court. De plus elle est d'autant plus brève que la situation est critique. Type de dialogue :

Altitude (sur un ton cassant par le pilote)

- 150 (réponse immédiate du mécanicien sans préciser l'unité de mesure).

A l'inverse dans une phase sans problème ce dialogue devient poussé à l'extrême :

- Monsieur le mécanicien pouvez-vous me dire à quelle altitude nous volons actuellement ?

- Mais bien sûr mon commandant, nous sommes actuellement d'après l'altimètre à 460 pieds et d'après la radiosonde à 153 mètres. Je peux vous dire aussi que nous volons à 450 nœuds et que nous nous trouvons au-dessus de Cherbourg.

- Merci, au fait comment va votre fils ?

Donc dans les phases de vol paisible, on utilise la parole pour augmenter la convivialité de l'équipage. Et quand le tour des histoires familiales, politiques ou artistiques a été fait et qu'il n'y a pas d'informations concernant le vol, un des membres de l'équipage chante ou siffle. Cette émission n'est destinée à personne en particulier. Son rôle est de rompre la monotonie du bruit des moteurs et d'activer le système nerveux central afin

de conserver un niveau de vigilance suffisant pour faire face à un événement imprévu.

Les gestes

Ces derniers servent à actionner des commandes ou des outils mais ils peuvent, comme la parole, être un langage de communication. Il en va de même pour le regard ou le visage. Ces systèmes de communications extra verbaux sont très importants à analyser dans l'optique de la conception d'un matériel. Car s'ils sont omis à l'analyse et impossibles à réaliser avec le nouveau matériel, il y aura un manque et une baisse de l'efficacité. En avion le mécanicien donne le top de larguer au pilote par une tape sur l'épaule. De même un chef d'équipe apprécie si ses subordonnés sont prêts à l'action à leurs mimiques du visage. Ce type de communication peut devenir impossible dans des environnements virtuels si les avatars ne sont pas assez représentatifs.

La fonction gestuelle correspondant à l'action sur des commandes ne peut être dissociée des commandes elles-mêmes. Dans le cadre de la réalité virtuelle, les interfaces manuelles devront répondre aux mêmes exigences qu'en ergonomie. Tout se résume en facilité et en sécurité d'emploi pour les «actionneurs» humains que sont les mains et les pieds avec lesquels il faut agir, sans possibilité de renvoyer ces «pièces» au Constructeur. Ainsi les mains sont plus adaptées aux actions de vitesse et surtout de précision, les pieds aux actions en force. L'acte précis sera meilleur s'il est effectué par la main dominante. Les éléments à prendre en compte correspondent à des données de l'anthropométrie dynamique : la force du geste, son amplitude, son sens, sa vitesse, sa durée et son nombre de répétition. Toutes ces activités sont influencées par des facteurs internes et externes qui vont être présentés dans le paragraphe suivant.

3.3.4 FACTEURS INFLUENTS

Le facteur influent premier est lié à la diversité humaine, les facteurs seconds concernent la nature et les variations de l'environnement physique et sociologique, le dernier concerne la nature de la tâche à effectuée.

3.3.4.1 La diversité humaine

Cette diversité se situe aux niveaux morphologique, physiologique, intellectuel et affectif. Elle se caractérise le plus souvent par une distribution gaussienne [Papin, 2000a]. C'est ainsi que sur le plan morphologique il existe certes, une disparité importante entre chaque individu, mais aussi une disparité entre les différents segments corporels. A titre d'exemple il est possible de rappeler que la distance entre les deux pupilles varie de 5 cm entre les sujets qui ont les yeux les plus rapprochés et ceux qui ont les yeux les plus écartés. Cela pose des problèmes pour réaliser par exemple des systèmes optiques standards utilisables par tous. De même deux sujets de même taille n'ont pas forcément les mêmes dimensions au niveau de leurs membres inférieurs. En position assise n'auront pas le regard au même endroit. Cela représente aussi des contraintes à prendre en compte au cours de projections d'images virtuelles destinées à plusieurs individus. Sur le plan physiologique, la sensibilité au contraste peut être différente d'un individu à l'autre, de même la sensibilité à la chaleur ou au bruit. Sur le plan intellectuel la manière d'appréhender l'environnement dépend parfois du QI. C'est ainsi par exemple qu'un test dit d'ajustement sensori-moteur est en réalité un test de compréhension pour

des individus à faible QI, un test de résolution de problèmes pour les très hauts QI et d'ajustement que pour les individus ayant un QI moyen. Ce test consiste, à l'aide de deux manivelles à maintenir une lumière dans un trou situé à la périphérie d'un disque opaque qui tourne. Les deux manivelles sont reliées à la lampe et permettent de lui faire décrire une trajectoire rectiligne, lorsque la lampe passe dans la région opaque le système s'arrête. Les sujets à QI moyen réalisent cette tâche en terme d'ajustement, Ils analysent dans quel sens la lumière est en train de se cacher dans la partie opaque et tournent les manivelles en conséquence. Les sujets à très haut QI comprennent très vite qu'il s'agit de commander un déplacement circulaire ($X^2 + Y^2 = \text{constante}$) et anticipent la manœuvre des manivelles. Les sujets à très faible QI ne comprennent pas la consigne. Sur le plan affectif rappelons simplement que la personnalité de chacun si elle est marquée par des traits intellectuels non négligeables (la raison) dépend aussi des traits affectifs qui jouent un rôle très important dans la manière de se comporter dans le monde physique et social qui nous entoure. Ce qui fait dire à Freud : « les déterminants décisifs du comportement sont inconscients » et « tout comportement est déterminé en dernier ressort par des pulsions ». Les manifestations du système conscient, « le moi », seraient guidées par le principe de réalité et obéissent aux lois physiques, aux conventions sociales et à la logique. Celles du « ça » par celui du plaisir dominé par les besoins subjectifs d'ordre biologique ou affectif.

3.3.4.2 Les facteurs extérieurs

Pour montrer l'importance de ces facteurs externes prenons comme exemple l'évolution des performances physiques et intellectuelles en fonction de la température extérieure (Figure 3.4). C'est ainsi que pour une tâche intellectuelle l'optimal de performance se situe aux environs de 21° effectifs (la température effective est la résultante de la température sèche, de l'humidité et du vent). A partir de 33° effectifs, il y a risque de coup de chaleur. Pour une tâche physique, le risque se situe à 30 ° effectifs. Le risque s'accroît avec la durée d'exposition. Toute activité est impossible au-delà de 72 heures. Pour le froid des gelures peuvent apparaître, s'il y a beaucoup d'humidité et de vent autour de 5° effectifs.

3.3.4.3 La tâche, notion de charge de travail

La charge de travail peut se définir comme le pourcentage d'activités demandées par rapport à la capacité instantanée d'activités possibles. Cette charge augmente soit par une activité demandée accrue et/ou soit par une diminution des activités possibles. En terme d'accroissement de la demande celle-ci se fait sous l'influence de déterminants externes qui sont de quatre ordres.

- premièrement en relation avec des facteurs d'environnement. Par exemple en conduite automobile, la présence de brume demande plus d'attention au conducteur et par-là même son activité générale est augmentée. Dans le cadre de la réalité virtuelle, cela peut être le bruit généré par les ordinateurs dans la salle où se déroule l'immersion ;
- deuxièmement en relation avec les conditions de réalisation de la tâche. Par exemple la conduite en convoi est plus difficile que la conduite en solitaire. Dans le cadre de la réalité virtuelle, cela peut être le partage par deux sujets du même environnement virtuel ;

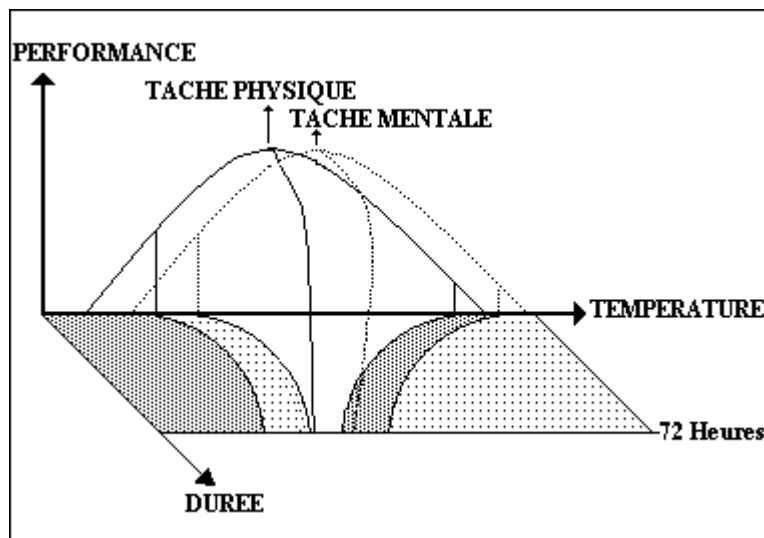


Figure 3.4 : Performances humaines et ambiance thermique

- troisièmement en relation avec l'organisation professionnelle. Un contrôle permanent de la conduite par relevé GPS d'un chef pointilleux accroît la tension des conducteurs. Dans le cadre de la réalité virtuelle, cela peut être un manque de précision dans les consignes données au sujet immergé ;
- quatrième en rapport avec des facteurs techniques ou organiques. Par exemple un mauvais éclairage de la planche de bord. Dans le cadre de la réalité virtuelle, cela peut être une mauvaise définition des images présentées dans les casques.

En terme de diminution des possibilités d'activités instantanées, celle-ci se fait sous l'influence de déterminants internes de quatre ordres.

- premièrement en relation avec des facteurs physiques ou organiques : une fatigue ou un rhume diminue les possibilités intellectuelles d'un sujet ;
- deuxièmement en relation avec des facteurs psychologique : des problèmes familiaux (maladie d'un enfant, dispute avec un proche) mais aussi un désir excessif de réussite rendent le sujet moins attentif à la tâche en cours ;
- troisièmement en relation avec des facteurs techniques : par exemple une expérience limitée de la tâche, un entraînement insuffisant et dans le cadre de la réalité virtuelle une accoutumance au port des interfaces (surtout visuelles) insuffisante entraîne une moins bonne mobilisation des possibilités d'actions instantanées des sujets ;
- quatrième en fonction de la durée : les possibilités instantanées diminuent progressivement dans la journée. Elles seront récupérées après repos et sommeil. Tout se passe comme si au fur et à mesure d'une période de veille, il y avait accumulation en « mémoire vive » d'informations qui saturent progressivement les neurones, en laissant de moins en moins pour agir. Le sommeil, en particulier paradoxal, permet de classer toutes ces informations et de les stocker en « mémoire morte ». A signaler qu'en informatique lorsqu'une sauvegarde s'effectue, il n'est plus possible d'entrer

de données. L'ordinateur est « en sommeil », comme l'organisme humain lorsque ce dernier stocke l'information d'une période de veille.

Ces remarques montrent qu'une même activité demandée peut représenter pour un sujet une charge de travail très différente en fonction de l'heure où elle est accomplie, des conditions extérieures dans laquelle elle se déroule et de l'état général du sujet à un moment donné. Cet état général correspond à la capacité instantanée du sujet ou à sa forme. Celle-ci correspond à la résultante de toutes les influences extérieures (climat, conditions de vie, formation, etc...) et intérieures (motivation par exemple) sur les caractéristiques morphologiques et les aptitudes (possibilités humaines) physiques, intellectuelles et affectives (Figure 3.5).

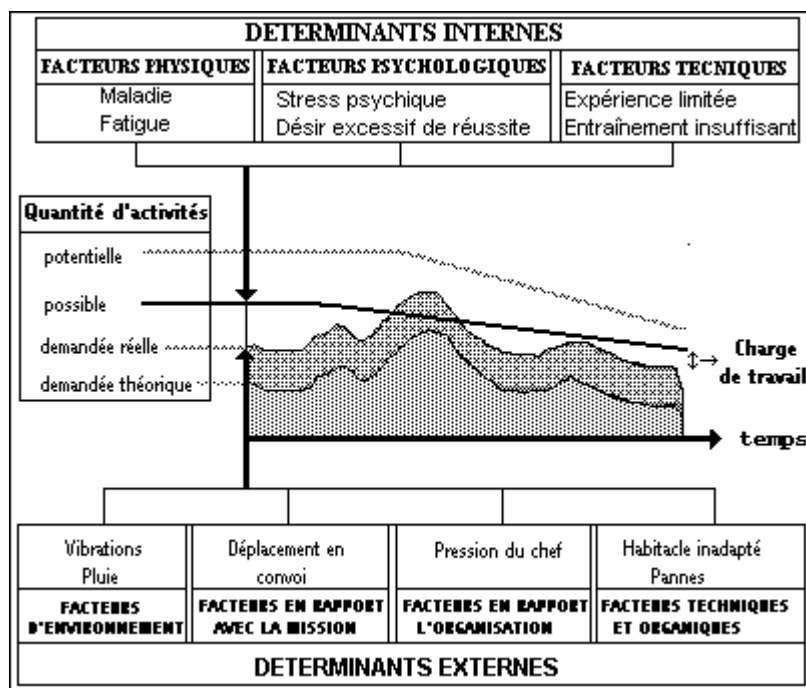


Figure 3.5 : Facteurs intervenant dans la réussite d'une tâche

La charge de travail physique peut se déduire des activités physiques mises en jeu, par le taux d'utilisation des commandes, par mesures des efforts exercés sur des commandes, par la consommation d'oxygène etc... La charge de travail mental s'évalue par des techniques indirectes comme [Santucci, 1987] :

- l'analyse des erreurs d'après la technique des incidents critiques de Flanagan ;
- l'analyse par double tâche de Brainbrige ;
- l'utilisation de SWAT (Subjective Workload Assessment Technics) ;
- la mesure de concomitants physiologiques comme la fréquence cardiaque, les clignements palpébraux.

3.4 CONCLUSION

Tout ce qui vient d'être dit montre que l'homme est en perpétuelle interaction avec son environnement et que son comportement dépend fortement du type de relation qu'il établit avec cet environnement. Son mode de réaction dans les environnements virtuels ne va pas être très différent, si ce n'est qu'il se fait par l'intermédiaire d'interfaces sur un mode créé de toutes pièces par l'homme.

3.5 RÉFÉRENCES BIBLIOGRAPHIQUES

[Curé, 1996]	M. Curé, «L'ergonomie», in L'Armement n°53 Juillet août (1996).
[Papin, 1992]	J-P Papin, «Cours d'ergonomie» EMSST, Paris (1992).
[Papin, 2000a]	J-P Papin, «Facteurs humains et autorité». Cours aux Ecoles Militaires de Coëquidan. (2000).
[Papin, 2000b]	J-P Papin, «Facteurs humains et réalité virtuelle». Cours à l'Ecole Navale de Brest (2000).
[Piéron, 1964]	H. Piéron, «La sensation», P.U.F (4 ^e édition, 1 ^{er} en 1952) Paris (1964).
[Santucci, 1987]	G. Santucci, «Evaluation de la charge de travail en aéronautique» Médecine Aéronautique et spatiale Tome XXVI n° 102 et 103 (1987).

4 LES SENS ET LES RÉPONSES MOTRICES DE L'HOMME

Jean-Paul Papin et Philippe Fuchs (partie vision)

Nous avons vu précédemment que la conception et l'évaluation d'un dispositif RV se font à partir de trois niveaux d'Immersion et d'Interaction, dont le premier est le niveau des I² sensori-motrices. Les techniques de la réalité virtuelle font intervenir des interfaces comportementales (sensorielles ou(et) motrices) pour coupler physiquement l'ordinateur à l'homme. La conception de telles interfaces a pour objectif de réaliser un dispositif matériel ayant des caractéristiques métrologiques performantes qui doivent, si possible, correspondre aux capacités maximums des sens ou(et) des réponses motrices de l'homme. Dans ces conditions, il est indispensable d'avoir une vision claire et précise du comportement sensori-moteur de l'homme dans le monde réel, aussi bien pour analyser efficacement les interfaces comportementales que les I² sensori-motrices. C'est pourquoi cette deuxième partie est consacrée d'abord au fonctionnement de l'homme sur le plan sensori-moteur : après avoir abordé les organes sensoriels spécifiques (vision, ouïe, toucher, odorat et goût) et les organes proprioceptifs permettant la localisation spatiale, l'équilibre, les mouvements et les déplacements. Nous traiterons de manière spécifique la manipulation d'objets. Mais en premier lieu, faisons un rappel général sur le fonctionnement des sens.

4.1 LE FONCTIONNEMENT DES SENS

Nous avons vu que physiquement, l'homme perçoit son environnement par l'intermédiaire de ses sens. La réception d'un stimulus sensoriel est le point de départ du transfert de l'information dans l'organisme par les voies nerveuses. Le message nerveux afférent provenant du récepteur sensoriel est véhiculé vers les centres nerveux (moelle épinière, cerveau). Après intégration et traitement de l'information dans ces centres, ceux-ci transmettent le message nerveux efférent aux organes effecteurs (muscles squelettiques, muscles oculaires, muscles des cordes vocales, etc.). Ceux-ci réalisent le comportement moteur, se traduisant par des mouvements, sauf pour le cas particulier des muscles des cordes vocales permettant la parole (Figure 4.1).

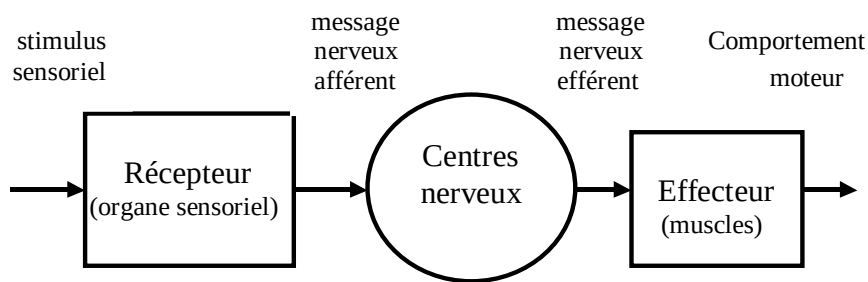


Figure 4.1 : Transfert de l'information dans l'organisme

Tout influx sensitif naît au sein de récepteurs spécialisés. Ils peuvent être classés en deux groupes. Les uns, répartis dans tout notre organisme, renseignent sur l'état du

corps et sur sa sensibilité générale (somesthésie). D'autres, spécifiques à un sens, sont situés dans l'organe sensoriel correspondant. Ce sont par exemple les photorécepteurs de la rétine pour la vue. Les récepteurs extéroceptifs, sensibles à l'extérieur, nous renseignent sur notre environnement. Les récepteurs proprioceptifs sont sensibles aux actions sur l'organisme : le mouvement du corps, sa position dans l'espace et les forces exercées sur les muscles.

Tous les récepteurs réalisent le codage de la stimulation sensorielle d'une façon identique : le stimulus appliqué crée une différence de potentiel au niveau de la membrane du site transducteur. Cette dépolarisation membranaire constitue le potentiel de récepteur dont l'intensité dépend de la valeur de la stimulation. Le potentiel de récepteur entraîne alors la formation de potentiels d'action qui vont être véhiculés par la fibre nerveuse sous forme d'une succession de potentiels, à la vitesse de 1 à 100 m/s. Leur amplitude et leur durée sont constantes (quelques dizaines de millivolts de 1 à 2 millisecondes) et leur nombre est fonction de la valeur du potentiel récepteur. Le message est donc codé en fréquence et transmis vers les centres nerveux. Le lecteur peut noter les similitudes (signaux binaires) et les différences (codage numérique ou fréquentiel) pour la transmission des signaux dans l'ordinateur et ceux dans le corps.

Il existe un seuil minimum en dessous duquel toute stimulation sur un organe sensoriel ne provoque aucun effet. Au-dessus du seuil, la variation minimum perceptible d'un stimulus est proportionnelle à la valeur absolue de l'intensité du stimulus (loi de Weber). Qualitativement pour cette loi, on peut dire pour la vision que dans le cas d'une faible intensité lumineuse perçue par l'œil, une petite variation d'intensité peut être détectée tandis que pour une forte intensité lumineuse perçue par l'œil, il faut une grande variation d'intensité pour quelle soit détectée. Par ailleurs, un stimulus doit avoir une durée minimale pour être perçu. A l'opposé, si une stimulation est maintenue en permanence, il y a généralement disparition ou atténuation de la sensation (phénomène d'adaptation), sauf pour la douleur et certaines fibres nerveuses. Quand nous parlerons de la résolution des interfaces, nous verrons l'importance de la densité des récepteurs et de la variation locale de cette dernière dans les organes sensoriels. La précision des sens humains est importante à connaître pour réaliser des interfaces sensorielles adéquates. En général, la précision absolue des sens est faible comparativement à la précision relative, l'homme ayant une grande capacité de comparaison entre deux stimuli, par exemples :

- il est difficile de définir une couleur d'un objet seul, mais il est plus aisé de percevoir une légère différence de deux couleurs observées simultanément ;
- la perception absolue de profondeur est difficile à estimer comparativement à la détection d'une faible différence de profondeur entre deux objets voisins ;
- il est délicat de quantifier la température absolue de l'air ou de l'eau en contact avec notre peau. Par contre, nous percevons facilement une légère différence de température entre deux fluides observés simultanément.

La fréquence maximum F_m de variation d'un stimuli sensoriel qui soit perceptible par l'homme est aussi une caractéristique importante à connaître. Les interfaces sensorielles doivent avoir, quand cela est techniquement possible, une bande passante de 0 Hz à F_m . Par exemple, il est souhaitable que les images soient affichées sur les écrans à plus de 25 (30) images par seconde¹, ce qui permet de voir un monde animé virtuel, fluide visuellement.

¹Mais si les écrans permettent cette fréquence, il reste le problème d'avoir un ordinateur capable de les

4.2 LA VISION

Une interface visuelle de qualité est plus délicate à concevoir qu'il n'y paraît, car la vision humaine est un sens complexe et souvent très méconnu. Mais celle-ci est si naturelle qu'on en oublie sa complexité et sa puissance informative. Il est important de retenir que l'étude de ce sens, comme des autres, ne peut être parfaitement comprise qu'à partir de ses fonctionnements physiques, physiologiques et cognitifs.

Le fonctionnement du système visuel ne se limite pas à une analyse statique des images qui sont projetées sur la rétine mais celui-ci construit dynamiquement un monde visuel. A la réception sur la rétine de la lumière, celle-ci est transformée en signaux chimiques puis électriques et ensuite des traitements spécialisés travaillent en parallèle et décomposent l'information. Les images formées sur les deux rétines sont décomposées selon des critères de profondeur, de forme, de couleur et de mouvement. Mais avant l'extraction des formes ou de la profondeur, des informations de bas niveau sont traitées et exploitées : des canaux fréquentiels² travaillant en parallèle sont chargés d'établir des représentations de l'information visuelle à différentes échelles de l'espace observé.

Après avoir brièvement décrit l'organisation du système visuel, nous nous concentrons sur le fonctionnement de ces canaux et présenterons une modélisation mathématique liée à la notion de fréquence dans le but d'analyser le fonctionnement du système visuel. Nous utiliserons en pratique cette analyse pour la caractérisation des performances en vision stéréoscopique et pour l'amélioration de ce type de vision. Mais pour cela, après l'étude des traitements monoculaires de l'information, nous aborderons l'étude de la perception de la profondeur par vision monoculaire et par vision binoculaire.

4.2.1 LE SYSTÈME VISUEL HUMAIN

4.2.1.1 L'ensemble du système visuel

Les deux yeux, récepteurs sensoriels du système visuel, participent à l'observation de l'espace. Les nerfs optiques sortant des yeux se regroupent au niveau du chiasma optique où les fibres temporales gauches rejoignent les fibres nasales droites et inversement. Les fibres du nerf optique sont alors dirigées vers le corps genouillé latéral, sauf une petite partie des fibres qui est dirigée vers le colliculus supérieur. Un des rôles supposés du corps genouillé latéral est de séparer en différentes couches cellulaires les neurones recevant les afférences rétiniennes. Du corps genouillé latéral, l'information est envoyée au cortex occipital, encore appelé cortex strié de par sa structure lamellaire. La Figure 4.2 montre le trajet de l'information visuelle depuis l'œil jusqu'au cortex visuel.

Le cortex strié, ou cortex visuel primaire ou aire V1, est organisé de manière à produire une représentation de l'information incidente sur la rétine. Comme dans le corps genouillé latéral, toutes les régions de la rétine y sont projetées et la représentation spatiale y conserve ses contiguïtés [DeValois e.a., 1988] : l'organisation du cortex strié est dite rétinotopique. La topologie est toutefois modifiée, la rétine centrale se projetant

calculer et de les transmettre aux écrans à cette fréquence ; voir le chapitre 10.

²Fréquences dans le domaine spatial des images, voir plus loin.

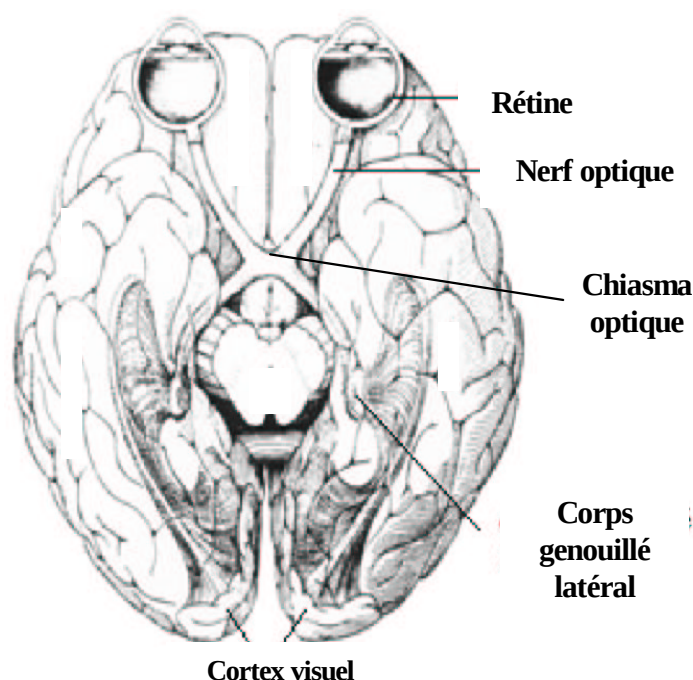


Figure 4.2 : Projections rétiniennes vers le corps genouillé latéral, puis le cortex strié

avec un facteur de dilatation, plus important que celui de la rétine périphérique, conséquence notamment de la densité de l'échantillonnage rétinien. La représentation de l'information dans le cortex strié est cependant beaucoup plus complexe que la simple organisation géométrique. Le traitement cortical de l'information ne se limite pas au cortex strié. Au-delà, une trentaine d'aires corticales a été décrite. Ces aires corticales diffèrent notamment par les propriétés de sélectivité de leurs neurones.

4.2.1.2 L'œil

La fonction des yeux est de canaliser la lumière, de longueurs d'ondes entre 400 et 700 nm, émise ou réfléchie par un objet pour créer une image nette qui s'imprime sur la partie de l'œil couverte de récepteurs sensoriels : la rétine. L'œil est constitué d'une succession de milieux transparents qui jouent le rôle d'une lentille convergente, dont la focale globale peut varier par modification de la courbure du cristallin (Figure 4.3).

Schématiquement, l'œil est une chambre noire de géométrie sphérique à l'entrée de laquelle se situent plusieurs dioptries et à l'arrière de laquelle se trouve la structure réceptrice :

- la pupille représente le diaphragme du système. Les rayons lumineux sont ensuite projetés sur la rétine qui fait office d'écran sphérique ;
- la cornée est une membrane fibreuse, transparente, qui constitue la lentille principale

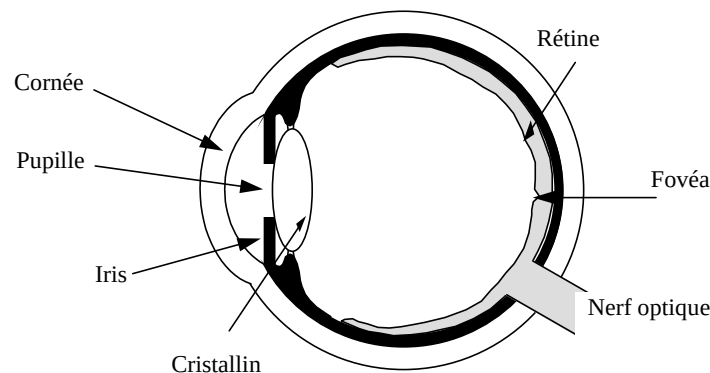


Figure 4.3 : Structure anatomique de l'œil humain

du système optique. Elle est le pôle antérieur de l'œil et l'axe optique passe en son centre. Son diamètre horizontal est de 11,5 à 12 mm ;

- le cristallin est une lentille biconvexe transparente placée entre l'iris et le corps vitré.

4.2.1.3 Accommodation et convergence

En regardant un objet à moins de 65 mètres environ, l'image réfléchie deviendrait floue sur la rétine si l'œil ne possédait le pouvoir d'accommoder automatiquement. Sous l'action des muscles ciliaires, la puissance du cristallin varie et permet de mettre au point sur la rétine pour voir des objets proches ou éloignés (Figure 4.4) : c'est le phénomène d'accommodation. Le cristallin est le seul dioptré de la chaîne visuelle dont la puissance soit variable. Tous les rayons lumineux sont déviés vers la fovéa qui est le point central de la rétine. Le niveau d'accommodation est réglé pour obtenir une image nette.

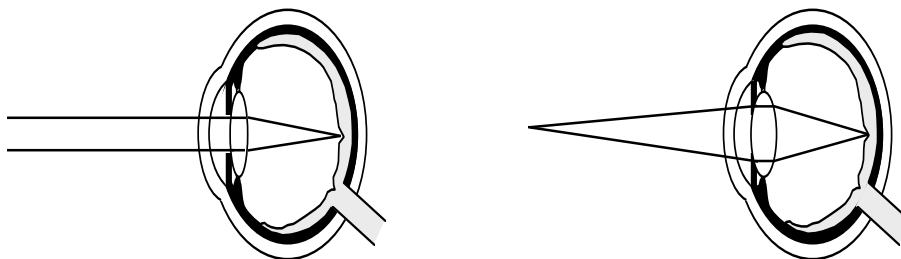


Figure 4.4 : Accommodation à l'infini et accommodation à courte distance (schéma optique simplifié)

Les muscles des globes orbitaux permettent d'orienter les deux yeux en les faisant converger vers le point de l'espace observé. C'est le phénomène de convergence. Comme

l'accommodation, la convergence des yeux se fait inconsciemment, sauf si la personne veut loucher. Le mouvement des yeux dans les globes orbitaux, pour changer le point de convergence ou(et) pour suivre un objet en déplacement, peut être très rapide avec une vitesse maximum de 600 degrés par seconde environ.

L'accommodation et la convergence des yeux étant toutes deux liées à la profondeur de l'objet que l'on observe, il en découle une relation entre celles-ci. Cette correspondance n'est pas innée chez l'homme mais est apprise expérimentalement et inconsciemment par l'enfant. Cette relation naturelle peut être mise en défaut artificiellement lorsque l'on regarde des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait converger les axes optiques sur lui tandis que les yeux s'accommodent sur l'écran. Cette modification de la relation convergence-accommodation est susceptible de créer des difficultés visuelles et de la fatigue pour l'utilisateur. Ce problème sera étudié dans le paragraphe sur la création d'images stéréoscopiques dans le chapitre des interfaces visuelles.

4.2.1.4 La rétine

La rétine est le lieu de la transformation de l'énergie lumineuse en signal électrochimique discret. Ce signal sort de l'œil par le nerf optique. Après avoir décrit son organisation générale, nous détaillons chacun des trois niveaux rétiniens : les photorécepteurs, les cellules bipolaires et les cellules ganglionnaires (Figure 4.5).

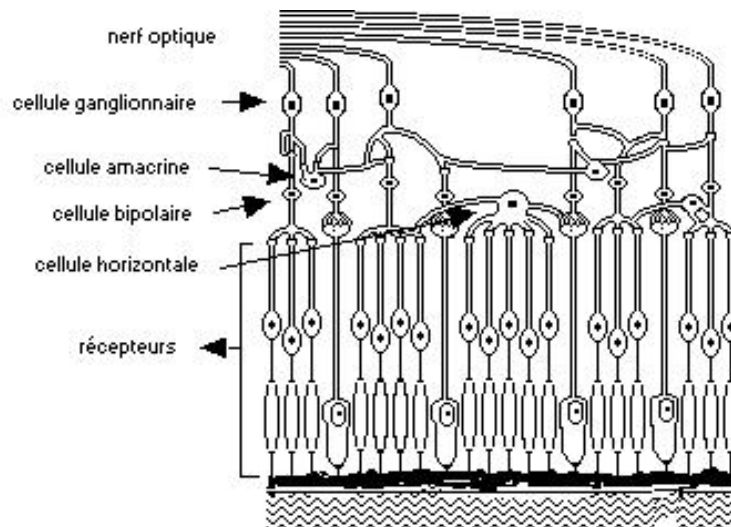


Figure 4.5 : La structure anatomique de la rétine

Organisation générale

La rétine est composée de plusieurs couches de neurones qui recouvrent le fond de l'œil. Ces couches forment un tapis de cellules connectées horizontalement au sein d'une même couche, et verticalement d'une couche à l'autre créant une architecture

neuronale fortement pyramidale. En effet, si le nombre de photorécepteurs est de l'ordre de 100 millions, le nombre des cellules ganglionnaires, dont les axones constituent le nerf optique, est de l'ordre du million. Entre les photorécepteurs et les cellules ganglionnaires, nous trouvons les cellules horizontales, les cellules bipolaires et les cellules amacrines. Paradoxalement, la première couche de cellules exposée à la lumière n'est pas celle des photorécepteurs, mais celle des cellules ganglionnaires, sauf au niveau de la fovéa. Les axones des cellules ganglionnaires sortent toutes de la rétine dans la même région dépourvue de photorécepteurs appelée tâche aveugle ou papille optique. Les photorécepteurs sont reliés entre eux et aux cellules bipolaires par les cellules horizontales, qui forment la couche plexiforme externe. Les cellules bipolaires créent des liens verticaux en reliant les photorécepteurs aux cellules de la couche suivante.

Les photorécepteurs : premier niveau rétinien

Il existe deux types de photorécepteurs, les cônes et les bâtonnets, de répartition très différente sur la rétine [Buser e.a., 1987] :

- les cônes, concentrés à la fovéa, sont sensibles à la longueur d'onde : ils sont de trois types car respectivement sensibles à des longueurs d'onde centrées sur 560, 530 et 420 nm. L'information de couleur est produite au niveau cortical par la comparaison des informations de plusieurs récepteurs contenant différents photopigments ;
- les bâtonnets sont, à l'inverse, très peu sensibles à la couleur : présents presque partout dans la rétine, ils sont absents au centre de la fovéa, et très denses à environ 20° d'angle visuel.

La densité des photorécepteurs et la bande passante en fréquences spatiales sont liées. Dans les conditions optimales, chez l'homme, la fréquence de coupure est de l'ordre de 60 cpd (cycle par degré) et la densité linéaire maximale proche de 120 unités par degré, ce qui est cohérent avec le théorème d'échantillonnage de Shannon. Même si les bâtonnets sont beaucoup plus nombreux que les cônes ($120 \cdot 10^6$ contre $6 \cdot 10^6$), ce sont les cônes qui contribuent en majorité à l'information transmise aux structures profondes du système visuel. Grâce à ses cônes et ses bâtonnets, le système visuel peut s'adapter à l'intensité lumineuse ambiante et couvrir une gamme d'intensités de l'ordre de 7 unités logarithmiques. Sous très faible éclairage, seuls les bâtonnets sont sensibles. Sous faible éclairage, les cônes sont couplés positivement à leurs voisins pour augmenter la surface du détecteur de photons. L'augmentation de sensibilité se fait au détriment de la résolution du système. En revanche, quand l'intensité lumineuse croît, les photorécepteurs sont découplés pour obtenir une résolution maximale [DeValois e.a., 1988].

Les cellules bipolaires : deuxième niveau rétinien

Il existe deux types de cellules bipolaires. Les unes répondent de la même façon que les photorécepteurs auxquels elles sont liées (non-inverseuses), les autres inversent le sens de la polarisation (inverseuses). La cellule bipolaire est reliée aux photorécepteurs soit directement, soit par l'intermédiaire des cellules horizontales.

Les cellules ganglionnaires : troisième niveau rétinien

Les cellules ganglionnaires sont classées en trois catégories en fonction de leurs propriétés fonctionnelles. Les deux classes principales, les cellules X et Y, diffèrent par

leur propriété de sommation sur leur champ récepteur. Les cellules X donnent des réponses très voisines de celles des cellules bipolaires. Leur champ récepteur possède des régions activatrices et inhibitrices. Leur réponse est qualifiée de linéaire. Sur le plan temporel, leur réponse est dite tonique ou soutenue car elle commence avec l'apparition de la stimulation, dure tant que la stimulation est appliquée, pour s'arrêter à l'extinction. A l'inverse, la réponse temporelle des cellules Y est dite phasique ou transitoire : elles ne répondent pas par une activité continue à l'application prolongée d'une stimulation donnée, mais ont une activité importante à l'allumage ou à l'extinction de la stimulation.

Importance des structures transversales

Grâce aux structures transversales qui la composent, la rétine permet de mettre en correspondance des structures cellulaires distantes, qui peuvent être activées de manière homogène ou contrastée. Par une organisation spatiale de leur champ récepteur en régions de même activation (régions excitatrices d'une part, régions inhibitrices d'autre part), les cellules ganglionnaires se définissent une affinité pour des motifs spécifiques en taille (et éventuellement en orientation), c'est-à-dire des motifs définis par leur fréquence spatiale.

4.2.2 TRAITEMENT DE L'INFORMATION DANS LE SYSTÈME VISUEL

4.2.2.1 Spécialisation du système visuel

Jusqu'aux années 1970, le concept de vision était fondé sur une dualité entre perception et interprétation :

- l'image s'imprime sur la rétine comme sur une plaque photographique ;
- l'impression laissée sur la rétine est comparée à d'autres impressions similaires perçues auparavant et interprétée en fonction de celles-ci.

Cette dualité, quasiment philosophique, se retrouve chez Kant qui sépare la sensation (passive) de la compréhension (active). Cette hypothèse semblait confirmée par les connexions entre la rétine et le cortex strié, plus particulièrement par la rétinitopie. L'étude des aires corticales proches de l'aire V1 (le cortex préstrié) a infirmé la conception dualiste de la vision, notamment lorsque Zeki a montré, par une étude chez le singe rhésus, que ces aires étaient fonctionnellement spécialisées [Zeki e.a., 1978]. Ces travaux ont pu être confirmés chez l'homme par la neurobiologie clinique.

Le corps genouillé latéral contient six couches de cellules : les quatre couches supérieures, contenant des cellules à petit corps, sont les couches parvocellulaires, les deux autres sont les couches magnocellulaires. Il ressort de différents travaux que le parvosystème et le magnosystème diffèrent non seulement par leur sélectivité aux couleurs mais également par leur sensibilité aux contrastes, leur résolution temporelle et leur résolution spatiale [Wong-Riley e.a., 1979], [Livingstone e.a., 1988]. D'après Livingstone et Hubel, la séparation des informations se ferait de la manière suivante :

- les couches parvocellulaires traitent l'information haute résolution sur les contours formés par les contrastes de couleurs ;

- les neurones K des couches parvocellulaires traitent l'information sur les couleurs et niveau de gris basse résolution ;
- le mouvement et la profondeur stéréoscopique sont traités par le canal magnocellulaire.

L'aire V1 est divisée en couches parallèles à la surface du cortex. C'est la porte d'entrée du système visuel, recevant presque la totalité de l'information du corps genouillé latéral. L'organisation de V1 est basée sur le regroupement des neurones participant à la même fonction. Afin de réaliser un échantillonnage fin et complet de l'image rétinienne, les neurones sont groupés selon au moins six paramètres : la cible de projection (autre aire corticale), le type d'afférent (parvo- ou magnocellulaire), la sélectivité à l'orientation, la sélectivité en fréquences spatiales, la dominance oculaire et la représentation du champ visuel [Bullier, 1998].

Autour de l'aire V1 se trouve l'aire V2 avec qui elle est fortement interconnectée, organisée elle aussi de façon modulaire. Les connexions de V1 à V2 se font par plusieurs canaux traitant différents types d'information en parallèle : le mouvement par une connexion avec l'aire V5, la couleur et la texture (surface et transition) par une connexion avec l'aire V4. A partir de V2, le système visuel s'organise selon deux voies anatomiques :

- une voie ventrale, appelée la voie du « what », responsable de la reconnaissance visuelle des formes ;
- une voie dorsale, la voie du « where », spécialisée dans la vision spatiale et la coordination visuo-motrice.

Remarques : ce découpage n'est pas sans rappeler les psychologues du Gestaltisme, qui suggèrent qu'une telle organisation soit possible sur la base des propriétés visuelles utilisées par le cerveau pour regrouper ou séparer des éléments d'une image. Ces propriétés sont typiquement la direction et la vitesse de déplacement, la coplanarité, la profondeur, la luminosité, la texture des surfaces, etc.

4.2.2.2 La notion de fréquence spatiale

Un son est caractérisé par sa fréquence, traduisant par-là la rapidité de ses oscillations. De même, nous pouvons définir pour un stimulus visuel une mesure de la finesse de ses formes. C'est la notion de fréquence spatiale. La définition de cette fréquence est la suivante : c'est le nombre de cycles d'une oscillation périodique de luminance ou de couleur dans un degré d'angle visuel. Sa valeur est donc donnée en cycles par degré (cpd), ce que l'on peut représenter schématiquement sur la Figure 4.6.

La fréquence spatiale, intuitivement décrite ici, peut être définie dans un cadre mathématique rigoureux [Perrin, 1998]. Campbell et Robson ont suggéré en 1968 que le système visuel pourrait être constitué de groupes de filtres passe-bande quasi-linéaires et indépendants, chacun centré assez étroitement sur une bande de fréquences spatiales [Campbell e.a., 1968] : on parle de canaux fréquentiels. La fonction de sensibilité au contraste ne refléterait donc pas la sensibilité d'un type de cellules mais serait l'enveloppe des sensibilités de tous ces filtres. Plusieurs aspects de cette hypothèse sont extrêmement attirants et motivent un grand nombre de recherches sur le sujet :

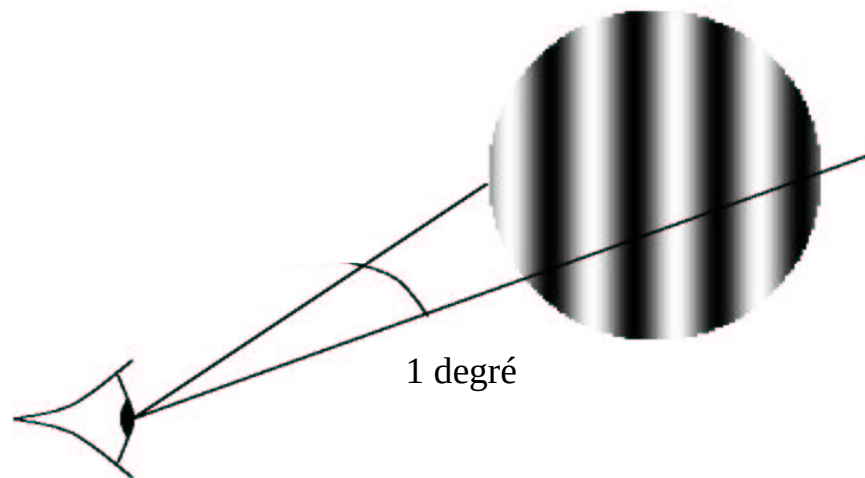


Figure 4.6 : Définition de la fréquence spatiale : sur l'exemple, la fréquence spatiale est de 4 cpd

- la similarité avec le système auditif, ajoutée à l'idée que le système nerveux traite probablement des problèmes analogues par des méthodes semblables ;
- la compression de l'information : en effet, l'information visuelle possède un grand nombre de périodicités (au moins localement) qui doivent être codées économiquement puisque le nerf optique ne possède qu'un million de fibres pour cent vingt-six millions de récepteurs.

Dans l'article qui a initialement émis la possibilité du filtrage de l'information visuelle en fonction de son contenu fréquentiel, Campbell et Robson ont mesuré la détection de divers motifs complexes [Campbell e.a., 1968]. Ils ont relié la détectabilité de ces motifs à leur amplitude spectrale et à la fonction de sensibilité au contraste d'un observateur pour des réseaux sinusoïdaux.

4.2.2.3 Les champs récepteurs des cellules rétinienne

L'information véhiculée par une cellule rétinienne ne concerne pas l'intégralité du champ visuel. En effet, le système visuel opère un échantillonnage par un pavage du champ visuel définissant la zone spatiale d'intégration de l'information pour une cellule.

Champ récepteur d'un neurone visuel :

Il est composé des récepteurs desquels il reçoit son information : le lieu des points du champ visuel qui se projettent sur les récepteurs liés à ce neurone et les activent. Nous nous intéresserons ici principalement aux champs récepteurs des cellules ganglionnaires de type X, plus précisément à la courbe de réponse de ces cellules à travers leur champ récepteur. L'existence de champs récepteurs possédant une structure à antagonisme concentrique a été mise en évidence à partir des cellules ganglionnaires du

chat et du lapin. On a découvert qu'un stimulus couvrant l'intégralité du champ récepteur produisait une réponse plus faible qu'un stimulus ne couvrant que le centre. Après avoir établi une carte des régions excitatrices et inhibitrices des champs récepteurs, on a montré que les cellules ganglionnaires X pouvaient être de type ON-OFF (ON-center, région excitatrice au centre) ou OFF-ON (OFF-center, région inhibitrice au centre). Pour une région du champ récepteur stimulée en intensité lumineuse, nous avons plusieurs réponses possibles de la cellule ganglionnaire (Figure 4.7).

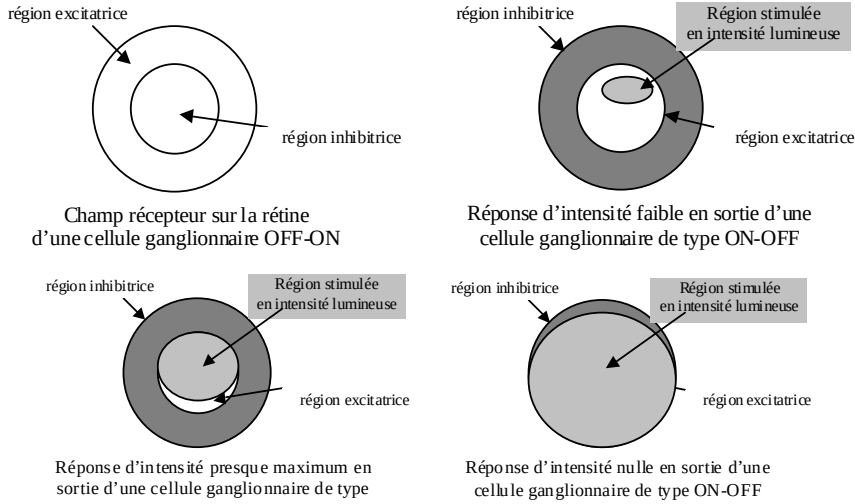


Figure 4.7: Les réponses d'une cellule ganglionnaire à différents stimuli lumineux

Les cellules ganglionnaires X sont donc assimilables à des filtres passe-bande en fréquence spatiale, c'est-à-dire que seule une bande de fréquences donnée produit des réponses non négligeables alors que les réponses à des stimulus de fréquence plus haute ou plus basse sont faibles, voire nulles. Elles sont donc spécifiques en fréquence.

Ceci explique, par exemple, certains phénomènes perceptifs comme le phénomène de contraste de la grille d'Hermann-Hering où les tâches grises aux intersections seraient le résultat d'un antagonisme entre des champs récepteurs circulaires (Figure 4.8).

Le champ récepteur de ces cellules a été modélisé par Rodieck et Stone comme étant la différence de deux dômes [Rodieck e.a., 1965], que Enroth-Cugel et Robson ont modélisée ensuite par la différence de deux gaussiennes ou DOG (Difference Of Gaussians) [Enroth-Cugel e.a., 1966]. La DOG correspond à une équation du type :

$$CR(x, y) = a \exp\left(-\frac{x^2 + y^2}{2\pi\sigma^2}\right) - \exp\left(-\frac{x^2 + y^2}{2\pi a^2\sigma^2}\right) \quad (4.1)$$

Cette équation définit un filtre passe-bande centré sur la fréquence $1/\sigma$ dont la largeur de bande est définie par σ et a . Un exemple de DOGs est représenté sur la Figure 4.9.

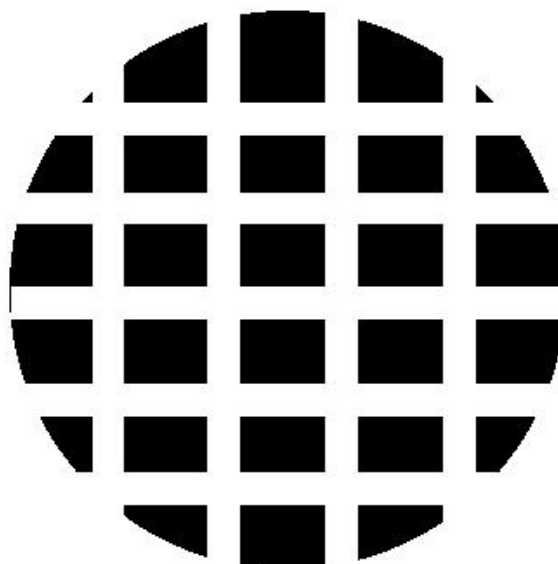


Figure 4.8 : Le phénomène de contraste de la grille d'Hermann-Hering

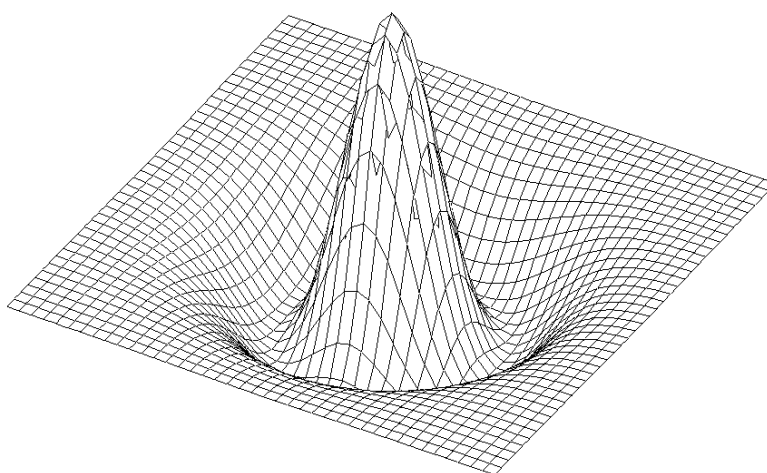


Figure 4.9 : Modélisation du profil des champs récepteurs d'une cellule ON-OFF par une différence de gaussiennes

D'un point de vue mathématique, la DOG peut être rapprochée du laplacien de la Gaussienne (LOG, Laplacian Of Gaussian), bien connu en traitement d'image. Nous pouvons ainsi montrer que les profils d'une DOG et d'un LOG sont très peu différents pour une valeur de a proche de 1,6. Marr et al. ont montré que l'opérateur LOG est optimal pour la détection des changements d'intensité dans les images [Marr e.a., 1980].

4.2.2.4 Les canaux fréquentiels

Un modèle simple a été décrit par Wilson et Bergen, composé de quatre filtres de tailles différentes pour l'analyse des images rétiniennes [Wilson e.a., 1979]. Ces canaux baptisés N, S, T et U ont pour largeurs centrales 3,1 min arc, 6,2 min arc, 11,7 min arc et 21 min arc au niveau de la fovéa, leur largeur croissant avec l'excentricité [Marr e.a., 1982]. Tous les auteurs ne sont pas d'accord sur le nombre de canaux fréquentiels et encore moins sur les fréquences centrales. Ainsi, Watson identifie en chaque emplacement de la fovéa sept détecteurs élémentaires qui couvrent les fréquences depuis 0,25 cpd jusqu'à 32 cpd [Watson e.a., 1982]. Wilson et al. ont mis en évidence une structure avec des largeurs de bande relative décroissant de 2,5 octaves (à 0,5 cpd) à 1,5 octaves (à 4 cpd) pour rester constante ensuite [Wilson e.a., 1983]. Les largeurs de bande sont proportionnelles aux fréquences centrales. Cela suggère une représentation mathématique espace-échelle³ plus qu'espace-fréquence de la scène visuelle et explique l'intérêt d'utiliser les analyses par **ondelettes** dans le domaine de la vision [Perrin, 1998]. L'analyse en ondelettes offre une bonne modélisation du traitement fréquentiel opéré par les cellules ganglionnaires.

En conclusion, l'étude détaillée des mécanismes du traitement de l'information par le système visuel peut sembler, pour le lecteur, éloigné des préoccupations du domaine de la réalité virtuelle. Nous verrons qu'il n'en est rien. Quel intérêt y aurait-il à vouloir faire une immersion par des images en relief si après quelques instants l'utilisateur a du mal à les observer ? Nous présenterons, pour exemple, dans le chapitre interfaces visuelles une application du traitement des images par ondelettes pour l'amélioration de la perception des images en relief. Car l'analyse en ondelettes est aussi intéressante pour la perception de la profondeur par le système visuel humain.. Mais avant nous allons préciser de quelles façons l'homme perçoit la profondeur car encore beaucoup de monde confonde la perception visuelle de la profondeur et la vision en relief ou stéréoscopique (voir le petit bêtisier de la réalité virtuelle).

4.2.3 PRISE VISUELLE D'INFORMATIONS

La prise visuelle d'information est traitée dans le chapitre «L'être humain, ses caractéristiques, ses relations avec ses environnements». Nous rappellerons seulement que la prise d'informations, qu'elle soit visuelle ou non, ne correspond pas à une réaction automatique mais à une action élaborée même si parfois elle reste inconsciente. L'exemple déjà présenté, concernant une phase de vol en hélicoptère, montre que les pilotes se repèrent à l'aide des verticales et non à l'aide de l'horizon et que la vision périphérique sert dans l'appréciation de la vitesse (grâce au défilement) et de la hauteur. Les études du comportement visuel ont permis de mettre en évidence que le temps

³Dans ce cas, la dimension de la région sur laquelle est analysé le signal est fonction de la valeur de la fréquence analysée.

d'arrêt du regard sur un objet est variable en fonction de la nature de l'information qui doit être perçue. Par exemple, il faut à un pilote d'avion consultant sa planche de bord 100 ms pour dire qu'une aiguille sur un cadran n'a pas bougé, 200 ms pour dire qu'il y a eu une variation angulaire, 250 ms pour en donner la valeur chiffrée et 450 ms pour lire un chiffre [Papin, 1992].

4.2.4 LA PERCEPTION VISUELLE DE LA PROFONDEUR

4.2.4.1 Perception cognitive par indices monoculaires

Il faut d'abord bien comprendre que la perception de la profondeur d'un monde tridimensionnel ne découle pas seulement de la vision binoculaire. Avec seulement un œil, l'homme interprète l'image reçue et en déduit des notions de profondeur. Cette interprétation se fait inconsciemment par traitement cognitif grâce à l'expérience apprise dès le plus jeune âge. Au stade précoce du traitement visuel, la décomposition et le traitement de l'information visuelle par différents canaux permettent d'utiliser une série variée d'indices pour la perception de l'espace tridimensionnel, donc de la profondeur. Ces indices se décomposent en deux grandes catégories :

- les indices proprioceptifs (dus aux actions des muscles orbitaux et ciliaires), composés de l'**accommodation** et de la **convergence** : ces indices sont fiables à faible distance (quelques mètres) ;
- les indices visuels, composés des **indices binoculaires** et des **indices monoculaires**.

Les indices proprioceptifs, accommodation et convergence, sont ajustés par le système visuel. C'est par la valeur de commande des muscles orbitaux et ciliaires que le cerveau a une perception proprioceptive de la profondeur de l'objet observé. Rappelons que la vision naturelle humaine est une vision *active*. Les yeux sont presque toujours en mouvement pour observer l'environnement et les cristallins se déforment pour permettre l'accommodation. Par contre, un appareil photographique observe en vision *passive* l'environnement réel. La profondeur de champ de toute image photographique est un indice de profondeur qui correspond partiellement au phénomène d'accommodation. On peut donc utiliser des effets de profondeur de champ sur des images de synthèse pour donner plus de « relief » à l'espace observé.

En vision monoculaire, les indices monoculaires, appris inconsciemment, permettent d'appréhender la 3ème dimension avec un œil, même si la perception de la profondeur est nettement mieux quantifiée avec la vision binoculaire. Les indices monoculaires peuvent être classés de la façon suivante :

-1- Ombres et lumière :

Les variations de lumière et les ombres sur les objets permettent d'avoir une meilleure perception de la forme tridimensionnelle de ceux ci (Figure 4.10).

-2- Dimensions relatives :

Chaque objet envoie au fond de l'œil son image dont les dimensions sont proportion-

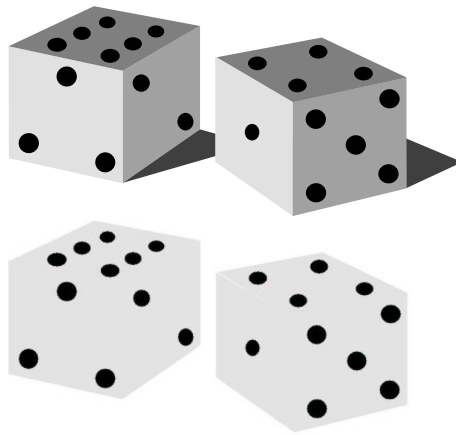


Figure 4.10 : Les variations de lumière et les ombres augmentent le « relief » des deux cubes

nelles à celle de l'objet et décroissante en fonction de la distance par rapport à l'œil. Le cerveau connaît les dimensions « normales » des objets réels. Il peut donc en déduire leurs distances approximativement. De même l'estimation de profondeur est facilitée pour une série d'objets, réels ou imaginaires, de formes identiques : ils sont *a priori* perçus à taille identique et positionnés à différentes distances dans un espace 3D (Figure 4.11).

-3- Interposition ou occultation :

Un objet peut cacher partiellement un autre objet placé derrière lui, ce qui permet de positionner relativement ces objets en profondeur (Figure 4.12).

C'est par un raisonnement cognitif que le cerveau interprète cette image : il perçoit spontanément une ellipse devant un rectangle et non deux formes adjacentes sur un même plan.

-4- Gradient de la texture d'une surface :

La texture d'une surface est perçue plus nettement si la surface est positionnée à faible profondeur. Ce gradient de la texture donne une information complémentaire de profondeur.

-5- Variation de visibilité dans une scène extérieure :

La visibilité d'une scène extérieure est décroissante en fonction de l'épaisseur de la couche atmosphérique, donc de la profondeur (Figure 4.13).

-6- Parallaxe due aux mouvements :

Quand un observateur se déplace, les images rétinienne d'objets immobiles ont des mouvements relatifs dépendant des distances de ceux-ci. Il en est de même quand des objets se déplacent relativement entre eux. En pratique, on peut utiliser cette règle pour rendre compte du relief par rotation d'un produit ou de toute une scène 3D visualisés sur un écran monoscopique. On peut aussi modifier les images sur l'écran, monoscopique ou stéréoscopique, en fonction du déplacement de l'observateur devant l'écran. Mais cette solution est plus difficile à mettre en pratique car il faut détecter en temps



Figure 4.11 : Voyez-vous 4 roses de même taille plus ou moins éloignés ou voyez-vous 4 roses sur un même plan ?

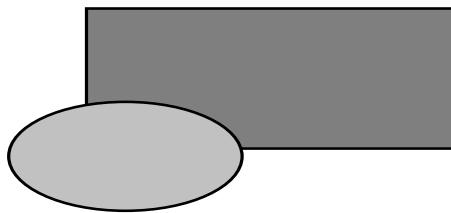


Figure 4.12 : Voyez-vous une ellipse devant un rectangle ou une ellipse à même distance d'un rectangle découpé sur un côté ?



Figure 4.13 : Variation de la visibilité et de la perception de la texture, photo A. Fuchs

réel la position de la tête de l'observateur.

-7- Perspective :

C'est bien sûr la règle la plus efficace pour percevoir le monde en relief sur un écran monoscopique. Celle-ci est employée en peinture depuis l'époque de la Renaissance pour visualiser sur une surface plane l'espace tridimensionnel. A noter qu'en passant d'images planes (sans perspectives) dans les jeux vidéo aux images avec perspectives sur écran monoscopique, les concepteurs de jeux vidéo parlent de jeux vidéo « 3D ». Il existe plusieurs types de perspectives (cavalière, géométrique, photographique, artistiques). La perspective cavalière est utilisée en dessin industriel. Elle est plus éloignée de la vision réelle que la perspective photographique mais tellement plus facile à réaliser à la main. Il n'y a pas de point de fuite, mais des « fuyantes » parallèles, comme sur la Figure 4.14.

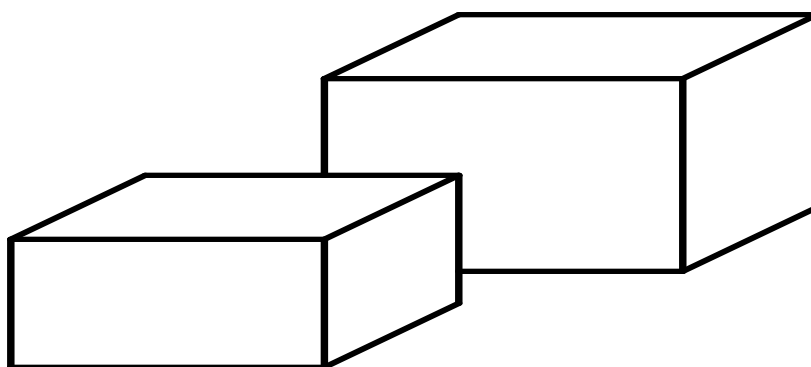


Figure 4.14 : Deux parallélépipèdes et non un ensemble de lignes sur une feuille plane !

Remarque : Le lecteur peut remarquer que nous venons d'utiliser ces règles sur ces pages (Figure 4.15). Vous pouvez fermer un œil, la perception de la profondeur est inchangée ! Ces règles sont maintenant couramment employées en images de synthèse pour donner une représentation tridimensionnelle sur un écran d'ordinateur. Dans Open GL par exemple, vous pouvez calculer vos images en perspective cavalière (orthographic projection) avec la commande `gluOrtho2D` ou en perspective photographique⁴ (perspective projection) avec la commande `gluPerspective`.

4.2.4.2 Convergence et disparité rétinienne

Précisons le phénomène de convergence et de son incidence sur les images rétinienne : en regardant fixement un de vos doigts placé à 20 cm devant vous, vous voyez une image dédoublée de tout objet situé en arrière plan. Inversement si vous fixez du regard un objet placé derrière votre doigt, vous voyez ce dernier se dédoubler (Figure 4.16) : Le cerveau n'arrive pas à fusionner les deux images différentes de votre doigt, la disparité rétinienne étant trop grande.

⁴Plus précisément, avec un modèle sténopé, donc sans faire intervenir les distorsions des optiques.

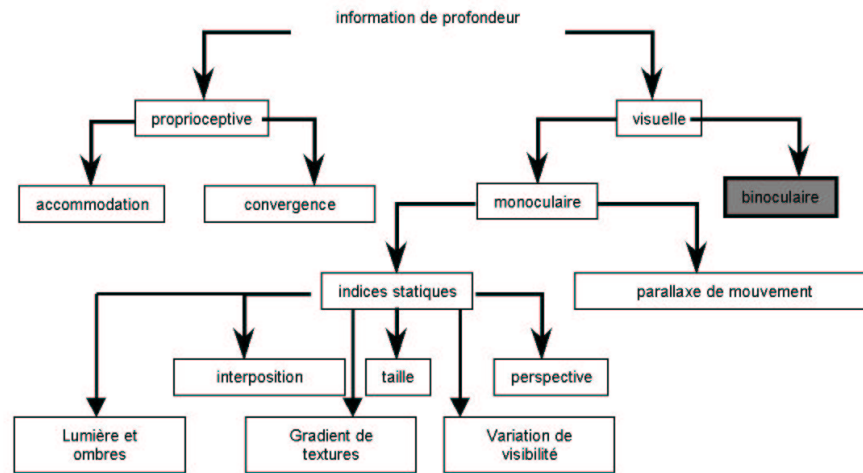


Figure 4.15 : Les indices permettant la perception de la profondeur de l'espace tridimensionnel

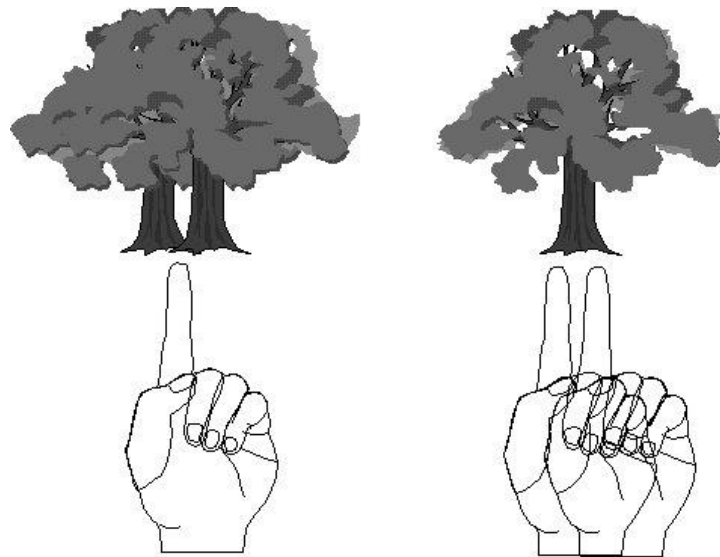


Figure 4.16 : Fixation sur le doigt ou fixation sur l'arbre

Du fait du recouvrement (partiel) des champs visuels des deux yeux et de la convergence des axes optiques, nous possédons deux vues légèrement différentes de la même scène permettant d'extraire une information de profondeur.

Les axes optiques convergeant sur un objet, le cerveau obtient donc des images dédoublées des autres objets éloignés de ce dernier. Que fait le cerveau de ces images dédoublées ? En pratique, il n'en tient pas compte mentalement. En particulier, quand vous regardez, même au loin, vous voyez toujours deux images floues de votre nez mais vous n'y prêtez pas attention... sauf maintenant... cela va peut-être vous perturber si vous y faites attention ! Dans le cas de la vision naturelle, les yeux n'accommodent pas sur les images dédoublées (les objets correspondants sont trop éloignés ou trop proches). Mais cela n'est pas identique sur un écran où toutes les images des objets sont nettes et à même distance. Nous détaillerons ce problème dans le paragraphe sur la création des images en relief. Normalement, la convergence des axes visuels est commandée par le cerveau, mais on peut physiquement faire diverger ces axes et voir les images se dédoubler : Regardez ce texte et appuyez avec un doigt sur un œil pour le faire tourner, les images se dédoublent, la fusion n'est plus possible.

Mais analysons maintenant plus géométriquement et optiquement ce qui se produit en vision naturelle à partir de la Figure 4.17 où un observateur regarde le point F.

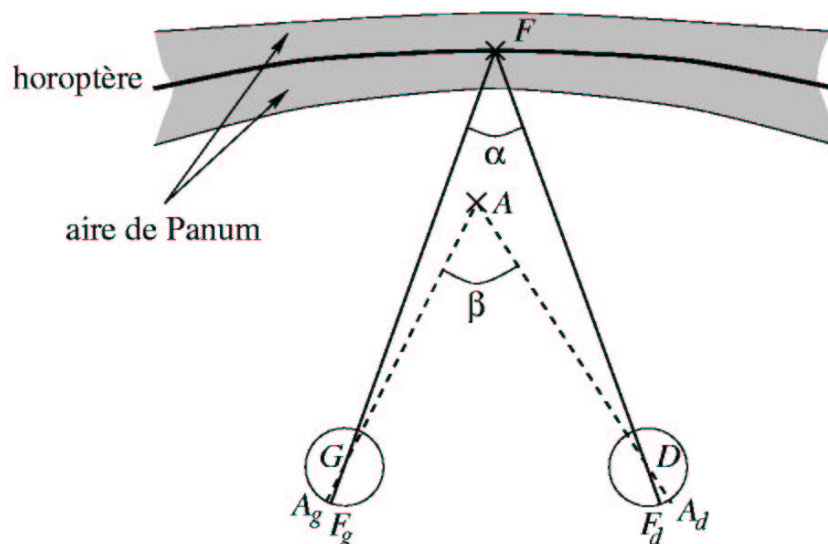


Figure 4.17 : Perception de la différence de profondeur entre le point de fixation F et le point A, à partir de la disparité

La disparité (horizontale) quantifie la notion de différence entre les images en mesurant des écarts. La figure donne une illustration de la disparité rétinienne : pour le point de fixation F nécessitant un angle de convergence α , la disparité d au point A sous tendant un angle β est définie par : $d = \beta - \alpha$, qui correspond au double de l'angle DF_dA_d quand A et F sont sur la normale à la droite DG. Géométriquement, les disparités sont suffisamment différentes que si les objets observés sont proches des yeux. La perception binoculaire de la profondeur n'est donc efficace qu'à petites distances (quelques mètres).

4.2.4.3 Stéréopsie et diplopie

Pour un point de fixation donné, l'ensemble des points de l'espace se projetant en des positions rétinienne correspondantes forme l'horoptère (sur la figure précédente, les points B perçus à même profondeur que le point de fixation F). L'horoptère théorique est appelé cercle de Vieth-Müller : c'est le cercle qui passe par le point de fixation et le point nodal de chacun des yeux. La détermination de l'horoptère empirique est basée sur différents critères selon la définition effectivement choisie [Tyler, 1991]. Les points situés en avant de l'horoptère sont dits à disparité croisée ou convergente, alors que ceux situés en arrière sont à disparité directe ou divergente.

Si les images rétinienne sont proches, avec de petites disparités horizontales, les stimulus binoculaires sont perçus en une seule image par fusion des points qui se situent dans l'aire de Panum (la zone grisée sur la figure précédente) : c'est la stéréopsie. Les objets sont perçus en profondeur en avant ou en arrière du point de fixation par stéréopsie. Cependant, la stéréopsie et l'apparence de profondeur sont dégradées par la présence de disparité verticale [Tyler, 1991]. Au-delà de l'aire de Panum, les stimulus de chaque œil sont perçus séparément. Ils sont dédoublés : c'est la diplopie.

4.2.4.4 Mécanismes neurophysiologiques de la perception de la profondeur

Après la mise en correspondance, la profondeur peut être estimée. Ce problème est loin d'être trivial puisque la disparité est un indice ambigu. En effet, la même valeur de disparité peut être associée à des distances différentes, selon la convergence du regard. Les informations de disparité doivent donc être calibrées en fonction de la convergence pour estimer la profondeur véritable : ce processus est nommé constance de profondeur. L'activité des cellules corticales traitant la disparité doit donc être extrêmement spécialisée et modulée par la distance de fixation. Ainsi, Trotter note que lorsqu'on modifie la distance réelle de profondeur, l'activité des neurones change de façon remarquable. Les neurones préfèrent toujours la même disparité rétinienne, par exemple 'plus près' ou 'plus loin' que le point de fixation. Mais leur niveau d'activité est en grande partie sous la dépendance de la distance de l'objet. Trotter démontre aussi que la perception cohérente de l'espace en trois dimensions est le résultat de la combinaison, au niveau d'une population de neurones dans le cortex visuel primaire, des messages issus de la rétine et d'informations extra-rétiniennes sur la position des yeux. L'information la plus probable est le signal proprioceptif de la musculature extrinsèque des yeux [Trotter, 1995].

Pour terminer ce long paragraphe sur la perception visuelle de la profondeur, nous pouvons faire la remarque globale sur le phénomène de la vision «3D» : d'un objet réel 3D, l'observateur recueille deux images 2D sur le fond de ses deux rétines. Son système visuel en extrait une perception 3D. En vision artificielle en relief sur écran, nous affichons deux images 2D calculées à partir d'un modèle informatique 3D de l'objet. Les rétines de yeux reçoivent encore deux images 2D qui sont perçues par le système visuel en 3D !

4.2.5 LES CARACTÉRISTIQUES PSYCHOPHYSIQUES DE LA VISION

Nous précisons dans ce paragraphe les caractéristiques utiles pour la compréhension d'une interface visuelle : Celles correspondant aux efforts entrepris actuellement pour améliorer les interfaces visuelles :

- amélioration de la définition de l'image ;
- augmentation du champ de vision ;
- vision stéréoscopique ;
- immersion du regard.

Mais pour commencer, nous pouvons parler de la sensibilité de l'œil, c'est-à-dire de sa capacité à réagir à des niveaux très faibles de la stimulation.

4.2.5.1 Sensibilité à la lumière

La quantité de lumière reçue par la rétine dépend de la quantité de lumière émise par une source et du diamètre pupillaire. C'est pourquoi on exprime l'intensité lumineuse, en terme d'éclairement rétinien, en Trolands (Td) : $Td = S \times L$ où S est la surface de la pupille exprimée en mm² et L la luminance exprimée en cd/m². Cependant, à un instant donné, cette sensibilité dépend du niveau d'adaptation lumineuse de la rétine à cet instant. Lorsque l'éclairement est élevé, seuls les cônes sont fonctionnels (les bâtonnets sont saturés). Il s'agit de vision photopique. Lorsque cônes et bâtonnets sont actifs, la vision est mésopique. Enfin lorsque seuls les bâtonnets sont actifs, la vision est scotopique (l'éclairement est insuffisant pour activer les cônes). Pour activer les bâtonnets après un séjour en forte lumière, il est nécessaire que les pigments se régénèrent. Cette durée correspond à l'adaptation à l'obscurité. Si en vision photopique une trop forte intensité lumineuse se produit, il y a éblouissement (saturation des cônes) et là aussi il faudra un temps de récupération pour que les cônes récupèrent leur sensibilité. Après éblouissement, il est possible de tracer la courbe d'adaptation à l'obscurité (Figure 4.18).

En plus de la quantité de lumière, la sensibilité est fonction de la durée et de la surface des stimuli qui expliquent, à côté de cette sensibilité absolue, l'existence des sensibilités fréquentielles.

4.2.5.2 Sensibilités fréquentielles

Elles correspondent à des stimuli dont la luminance varie sinusoïdalement dans le temps et l'espace. On distingue trois sensibilités, une temporelle, une spatiale et une spatio-temporelle. La fréquence temporelle est une variation périodique de l'intensité, caractérisée aussi par l'amplitude de la modulation d'intensité. Elle est décrite par le contraste de Michelson (Cm) qui dépend de la luminance maximale (L_{max}) et de la luminance minimale (L_{min}) :

$$Cm = (L_{max} - L_{min}) / (L_{max} + L_{min})$$

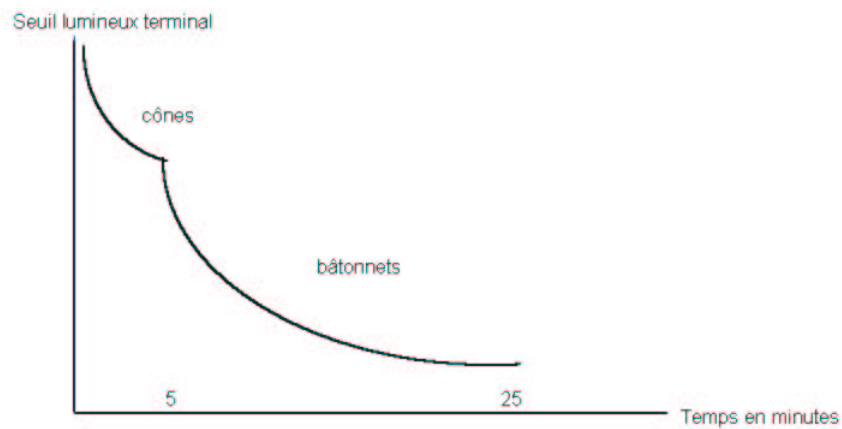


Figure 4.18 : Courbe d'adaptation à l'obscurité

Pour une source lumineuse donnée (en taille, luminance moyenne et spectre donné) il existe une fréquence temporelle de son papillotement qui conduit à une perception stable de l'éclairement. C'est la fréquence critique de fusion (CFF en anglais Critical flicker frequency). Cette fréquence a tendance à augmenter avec l'excentricité rétinienne. C'est pourquoi un écran de télévision par exemple peut donner en vision périphérique une sensation de papillotement.

Les variations des seuils de contraste avec la fréquence temporelle sont décrites par une fonction de sensibilité au contraste temporel (FSCt) selon le niveau moyen de luminance exprimé en trolands. Sur la courbe, on a en abscisse (unités logarithmiques), la fréquence temporelle des modulations et en ordonnée (unités logarithmiques), la sensibilité au contraste exprimée en pourcentage de contraste (origine en haut). Les quatre courbes correspondent à des intensités différentes de droite à gauche 10000 td, 100 td, 10 td et 1td (Figure 4.19).

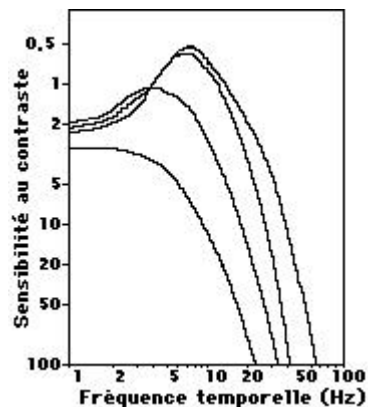


Figure 4.19 : Sensibilité au contraste temporel (d'après de Lange 1958)

4.2.5.3 L'acuité visuelle

L'acuité visuelle n'est pas homogène dans tout le champ de vision. Pour un œil emmétrope (normal), l'acuité monoculaire est très grande pour un cône centré de 2° d'angle, conséquence de la répartition des cônes rétinien limitée à la fovéa de l'œil. La valeur minimum de l'angle sous lequel deux points sont vus séparément est fonction du stimulus observé :

- une fente éclairée sur fond noir : $30''$ d'angle ;
- deux points clairs sur fond noir : $1'$ d'angle ;
- deux points noirs sur fond clair : $2'$ d'angle.

Nous voyons que la valeur moyenne ($1'$) est en accord avec la densité des photorécepteurs au niveau de la fovéa (voir le paragraphe sur la rétine). Un angle de 1 minute correspond à la vision de deux points séparés de 0,1 mm pour une distance entre l'image et les yeux de 35 cm. Pour respecter ces caractéristiques, un écran situé à cette distance et large de 25 cm doit afficher horizontalement 2500 pixels. On observe donc que les bons écrans d'ordinateur (1280 pixels) ont une définition d'image qui doit être améliorée. Ces valeurs numériques sont établies pour un écran donnant un champ de vision étroit (40 degrés). Si on désire augmenter ce champ en gardant une résolution d'écran suffisante, on doit afficher un nombre impressionnant de pixels incompatibles avec les technologies actuelles. A moins de détecter le mouvement des yeux car l'acuité n'est élevée que proche des axes de visée (ce qui n'est pas exploité sur les visiocasques commercialisés). La valeur moyenne (1 minute d'angle) de l'acuité visuelle doit être retenue pour juger de la qualité de la résolution d'un écran en situation dans un dispositif RV.

La vision humaine permet de détecter une différence de profondeur entre deux plans (ordre de grandeur : à un mètre, on perçoit une différence de profondeur de un millimètre). On exprime cette **acuité stéréoscopique** en angle. La Figure 4.20 permet de définir l'acuité stéréoscopique.

Les yeux, séparés de la $DIO = b_0$ et en position L et R, convergent vers le point B. Comment seront perçus les points A et C en avant et en arrière du point B et quelles seront les différences minimums Δr_A et Δr_C détectables ? On suppose que la vue est directe et que l'objet est suffisamment lointain. On pose $w_L = w_R = w$, cela implique que l'angle $LBR = \theta_B = \text{angle } LDR = \theta_D = \theta$. En considérant les triangles LAR et LDR, on peut écrire que :

$\theta_A = \theta + w$ et de même $\theta_C = \theta - w$. L'angle w est petit, donc $\text{tg}(w) = w = BD / r_B$ et en considérant les triangles LAR et BAD, on a avec la distance $BD = \Delta l$:

$$\frac{b_0}{\Delta l} = \frac{r_A}{\Delta r_A} \implies \Delta r_A = \Delta l \frac{r_B - \Delta r_A}{b_0} \quad (4.2)$$

$$\Delta r_A = \frac{r_B^2 w}{b_0 + r_B w} \quad (4.3)$$

D'une manière analogue :

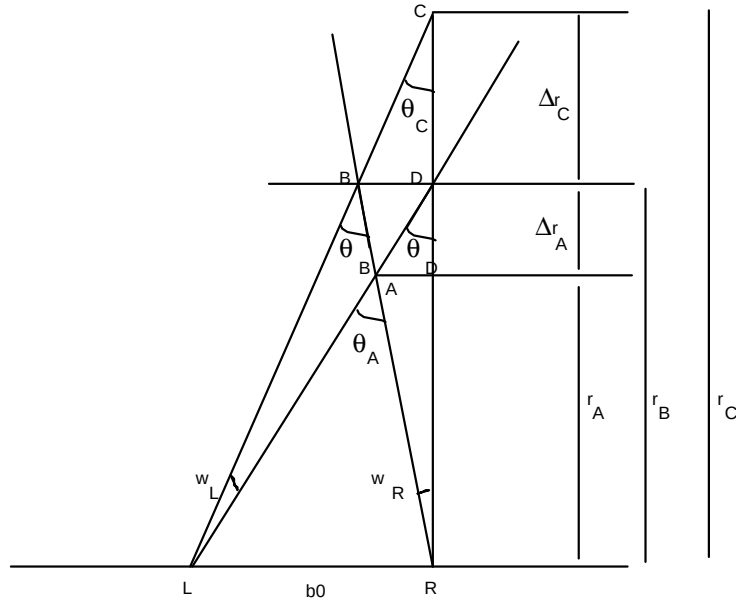


Figure 4.20 : Schéma pour calculer l'acuité stéréoscopique Δr

$$\frac{b_0}{\Delta l} = \frac{r_C}{\Delta r_C} \Rightarrow \Delta r_C = \frac{r_B^2 w}{b_0 - r_B w} \quad (4.4)$$

Les formules (1) et (2) définissent la distance limite de discrimination d'un objet situé à une distance Δr_A avant ou Δr_C après le plan situé à une distance r_B . Les inverses $1/\Delta r_A$ et $1/\Delta r_C$ des distances de discrimination représentent le seuil de discrimination stéréoscopique dans le cas d'un objet situé à une distance r_A avant et une distance r_C derrière le plan de référence qui est lui-même situé à une distance r_B de l'observateur. Le seuil de discrimination est défini comme la moyenne entre les deux seuils (avant et après) et est donné par :

$$\Rightarrow \frac{1}{\Delta r} = \frac{1}{\Delta r_B} = \frac{b_0}{w r_B^2} = w(r) \quad (4.5)$$

Dans cette formule b_0 est exprimée en mètres et w en radians. L'unité dans laquelle est exprimée $w(r)$ est l'inverse du mètre. $w(r)$ indique le nombre de plans que l'on peut discriminer par mètre dans le sens de la profondeur à partir d'une distance r (de l'observateur).

Exemple : En prenant l'écartement entre les yeux égal à 65 mm et le seuil physiologique égal à 0,065 milliradians, le seuil stéréoscopique est égal à :

$$w(r) = \frac{0.065}{0.000065} \cdot \frac{1}{r^2} \approx \frac{1000}{r^2} (m^{-1}) \quad (4.6)$$

Et donc $\Delta \mathbf{r} = 0.001 \mathbf{r}^2$ (en m). Cela signifie que si un objet est situé à une distance de 10m de l'observateur, ce dernier ne peut voir le relief que pour deux plans distants de 0,1m. Pour un objet à 1 m, le relief n'est perceptible que pour des plans séparés de 1 mm. En conclusion, la vision en relief n'est relativement performante que pour des distances faibles. Etant donné que cette vision procure des fatigues oculaires, ces considérations devraient inciter les concepteurs à ne pas exploiter *automatiquement* la vision stéréoscopique dans toutes les situations, mais seulement pour des cas bien justifiés. Nous en reparlerons, mais que le lecteur retienne que la vision stéréoscopique, employée d'une façon inappropriée, peut aller à l'encontre du but recherché : une meilleure perception de la profondeur pour une immersion visuelle performante.

4.2.5.4 Le champ de vision

Les deux yeux participent à l'observation de l'espace. La région de l'espace vue à chaque instant par l'un ou l'autre œil est définie comme le champ visuel. Les points de l'espace qui se situent dans la zone de recouvrement binoculaire sont acquis simultanément par l'œil droit et l'œil gauche (l'overlap, en anglais, dans les visiocasques). Les caractéristiques du champ visuel pour un œil immobile sont approximativement : horizontalement, 90° (côté tempe) et 50° (côté nez) et verticalement, 45° (haut) et 70° (bas). Mais rappelons que l'acuité visuelle n'est très précise que pour un cône central d'angle 2°. Avec les deux yeux fixes, il y a une zone de recouvrement correspondant au champ de la vision binoculaire de 100° en moyenne horizontalement.

Si on désire obtenir une immersion visuelle, il faut tenir compte du mouvement des yeux et de la tête. Un œil peut tourner dans son orbite d'environ 15° horizontalement et verticalement, avec une vitesse maximum de 600°/s environ. La tête peut tourner à la vitesse maximum de 800°/s environ. Il faut donc théoriquement au maximum un champ de vision sur les images de 170° horizontalement et de 145° verticalement. En pratique, les champs de vision dans les visiocasques sont encore bien petits (pour des visiocasques haut de gamme, on atteint 120° horizontalement). L'impression d'immersion visuelle y est donc plus ou moins forte. Il faut en conséquence bien étudier ce que les I² sensori-motrices exigent au niveau visuel. Le champ de vision total avec mouvement de la tête et des yeux est très grand (horizontalement >200 ° côté tempe et 130° côté nez, verticalement : 140° haut et 170° bas).

A noter que la Distance Inter-Oculaire (DIO, en anglais IPD, Inter-Pupillary Distance) est variable pour l'être humain. En moyenne, elle est de 65 mm pour l'homme adulte, légèrement inférieure pour la femme en moyenne, les écarts pouvant être significatifs, 50 à 70 mm environ, pour la population européenne. Cette grandeur est à prendre en compte pour l'utilisation des visiocasques, les deux optiques devant être bien centrés par rapport aux yeux et ainsi avoir un champ de vision optimum. Pour des réglages fins d'un visiocasque, nous avons mesuré la DIO pour chaque utilisateur en employant un pupillomètre plutôt que de laisser faire un réglage approximatif par l'utilisateur.

4.2.5.5 Fréquence temporelle maximum en vision

La fréquence temporelle des images pour percevoir des mouvements fluides est critique pour des valeurs inférieures à 25 images par seconde (valeur dépendant du type d'images). Les mouvements seront fluides sur des écrans monoscopiques à 50 Hz (une image est composée de deux trames). Pour les écrans stéréoscopiques à multiplexage temporel, la fréquence de balayage doit être double, 100 Hz (ou 96 Hz) pour avoir 25 (ou 24) images par seconde et par œil.

4.2.5.6 Les caractéristiques psychophysiques de la vision stéréoscopique

Nous avons explicité que le système visuel arrive à fusionner les deux images rétiniennes pour l'ensemble des points qui se situent dans l'aire de Panum. Dans le cas contraire, le cerveau est incapable de fusionner les deux images ou il le fait avec une certaine fatigue. Pour la vision en relief, l'objectif sera de créer des images stéréoscopiques créant dans les yeux de l'observateur de faibles disparités rétiniennes, tout en gardant un effet de profondeur. Mais quelles sont les limites de fusion acceptables en moyenne pour les êtres humains ?

Nous avons défini la disparité rétinienne. Pour des images en relief affichées sur un écran stéréoscopique, la disparité horizontale est égale (approximativement) à la parallaxe horizontale. Elle est définie par l'angle DPG, formé par les deux points homologues des images droite et gauche vues du point de vue de l'observateur.

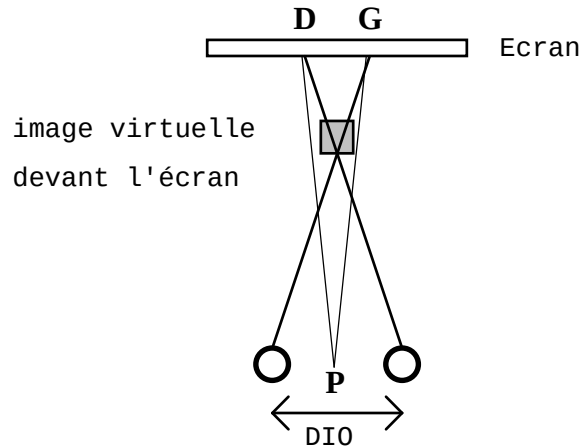


Figure 4.21 : La parallaxe horizontale est définie par l'angle DPG

En prenant que la parallaxe horizontale comme paramètre influençant la stéréopsie, des études expérimentales [Valyus, 1962] ont démontré la difficulté de fusionner deux images planes fournissant des parallaxes horizontales (et donc des disparités rétiniennes) d'angle supérieur à 1,5 degrés. Nous avons entrepris des tests en situation de travail dans le cadre de la téléopération avec des images stéréoscopiques réelles. Nos résultats donnent une limite un peu inférieure pour la limite de parallaxe horizontale : $1,2^\circ$

[Fuchs e.a., 1995]. Ces variations sont normales car elles sont variables pour chaque individu qui tolère plus ou moins les images en relief et qui force ou pas son système visuel à fusionner les images. A noter que 3 à 5% (peut être plus) de la population n'arrive pas à fusionner des images en relief sur un écran.

En faisant intervenir la durée de visualisation des images, l'étude de Yei-Yu Yeh [Yei-Yu e.a., 1990] examine la limite de la fusion et l'appréciation de la profondeur. La limite de fusion est analysée en fonction de différents paramètres des stimuli affichés sur un écran cathodique avec lunettes actives : la couleur, le temps d'affichage et la parallaxe horizontale. Les résultats impliquent que la limite de fusion (donnée en angle de parallaxe) est très faible pour un stimulus bref de 0,2 seconde : soit 27 à 24 minutes d'arc comparativement à un stimulus de 2 secondes : 1,5 degrés. Il est donc plus difficile d'observer en relief des images en mouvement rapide. La couleur rouge est plus facile à fusionner que la couleur blanche (pour des stimuli de 2 secondes : 6,2 degrés pour le rouge et 3,7 degrés pour le blanc). Les erreurs d'appréciation de la profondeur sont déterminées expérimentalement : l'erreur d'angle de parallaxe est de 2,2 minutes pour des images ayant des angles de parallaxe entre 0 et 30 minutes.

Des études ont montré qu'il est difficile de fusionner des images créant des parallaxes verticales (et donc des disparités verticales). Les yeux, généralement positionnés sur un même plan horizontal, ne sont pas aptes à percevoir des disparités verticales, sauf si elles sont très faibles. Pour les images en relief, les parallaxes verticales doivent être inférieures à un angle de 20' d'arc [Julesz, 1971]. Nous verrons, dans le paragraphe sur la création des images en relief, que cette contrainte est à prendre en compte.

Les études entreprises au Cerma (Centre d'Etudes et de Recherche de Médecine Aérospatiale) du Ministère de la Défense donnent des résultats corroborant les précédents, en étant plus précis. La limite de fusion est aussi fonction de la *fréquence spatiale des images*, ce qui n'est pas étudié dans les travaux cités précédemment. Leurs études permettent de conclure que de plus grandes disparités horizontales sont fusionnables lorsque les stimulations visuelles sont constituées de basses fréquences spatiales. Du point de vue pratique, nous pouvons en conclure que les objets représentés par leurs contours (hautes fréquences spatiales) sont plus difficilement fusionnables. Leurs études ont mis en évidence deux mécanismes impliqués dans la fusion selon le temps du stimulus : une fusion immédiate si la disparité n'est pas trop grande et une fusion non immédiate pour des disparités plus grandes, mettant en jeu des mouvements de vergence réflexes des yeux. Les résultats donnent entre autres : à la fréquence spatiale de 4,8 cycles/degré, fusion immédiate jusqu'à 20' d'arc et fusion maximale à 52' d'arc environ. A la fréquence spatiale de 0,22 cycles/degré, fusion immédiate jusqu'à 80' d'arc et fusion maximale à 176' d'arc environ, [Plantier e.a., 1990] et [Perrin, 1998].

4.2.5.7 Discrimination des couleurs

Les couleurs chaudes (rouge, jaune) font apparaître les objets plus grands que les couleurs froides (vert, bleu). Il faut rappeler que les couleurs sont perçues par des cônes spécialisés pour le rouge, le vert et le bleu. Les cellules ganglionnaires de la rétine fonctionnent par couples antagonistes (rouge-vert, bleu-jaune, intensité lumineuse). Ceci a pour conséquence qu'une altération au niveau de la vision d'une couleur par le cône spécialisé retentit sur l'ensemble de la vision des couleurs.

4.2.5.8 Dépendance et indépendance au champ

La dépendance et l'indépendance au champ est une caractéristique de la perception visuelle qui se traduit par la possibilité pour une partie des humains de reconnaître très rapidement des objets dans une scène complexe sans pratiquer d'exploration détaillée (indépendant au champ). Les autres sont obligés d'explorer l'ensemble des objets situés dans le champ visuel pour effectuer la reconnaissance des objets (dépendant au champ). Cette faculté d'indépendance s'acquière à partir de 12 ans et seulement pour un peu moins de 40 % des sujets. Environ 30 % de la population reste dépendant du champ et le reste est plus ou moins dépendant en fonction des scènes regardées. Ceci a des répercussions sur le sentiment d'immersion dans les systèmes de réalité virtuelle. Les dépendants au champ auraient besoin de champs visuels plus importants pour avoir le sentiment d'immersion visuelle.

4.3 L'OUÏE

4.3.1 INTRODUCTION

La connaissance des mécanismes sonores et de l'audition sont nécessaires pour les environnements virtuels dans trois cas :

- pour contrôler l'ambiance vibratoire des locaux où sont créés des environnements virtuels ;
- pour réaliser des systèmes de présentation de sons 3D au niveau des oreilles ;
- pour choisir entre une ambiance acoustique 3D ou une ambiance acoustique normale.

L'audition avec la vision sont les deux sens humains les plus utilisés pour s'informer. L'audition est omnidirectionnelle et par-là même capte tous les sons de l'environnement qu'ils soient utiles ou non. Sous certaines formes et à certaines intensités ils deviennent nocifs pour l'appareil auditif entraînant des lésions plus ou moins définitives qui diminuent les capacités de perception. Le son est initialement la sensation acoustique engendrée par une vibration acoustique. Le bruit est défini dans le vocabulaire de l'acoustique comme un « phénomène acoustique produisant une sensation auditive considérée comme désagréable ou gênante ». La commission électrotechnique internationale a jugé bon de la compléter par un aspect objectif « son ayant un caractère aléatoire qui n'a pas de composante bien définie ».

4.3.2 PHYSIQUE DU SON

Le son est produit chaque fois qu'un corps bouge à une vitesse suffisante pour envoyer une onde à travers le milieu dans lequel il vibre (solides, liquides ou gaz, mais pas dans le vide). Habituellement le milieu considéré est l'air. La source sonore est un objet qui vibre. Cette vibration repousse les molécules d'air qui se heurtent à leurs voisines et reviennent à leur position initiale. La perturbation se transmet ainsi de proche en proche tandis que les molécules se déplacent elles-mêmes très peu. Les endroits où les molécules se heurtent sont des zones de compression tandis qu'aux endroits quittés par les molécules se forment des zones de raréfaction. Ces zones se déplacent de

manière concentrique. La vitesse de transmission détermine la vitesse du son ou célérité. Elle dépend du milieu dans lequel le son se propage. Elle est de 340 mètres par seconde dans l'air, cependant elle varie en fonction de la température, elle augmente de 0,6m/s environ pour une élévation de température de 1°C. Dans l'eau elle est de l'ordre de 1435 mètres par seconde et de 6000 mètres par seconde dans l'acier. L'onde vibratoire ainsi obtenue se comporte suivant des lois précises de réflexion, réfraction, interférences, diffraction, effet doppler et résonance (Figure 4.22).

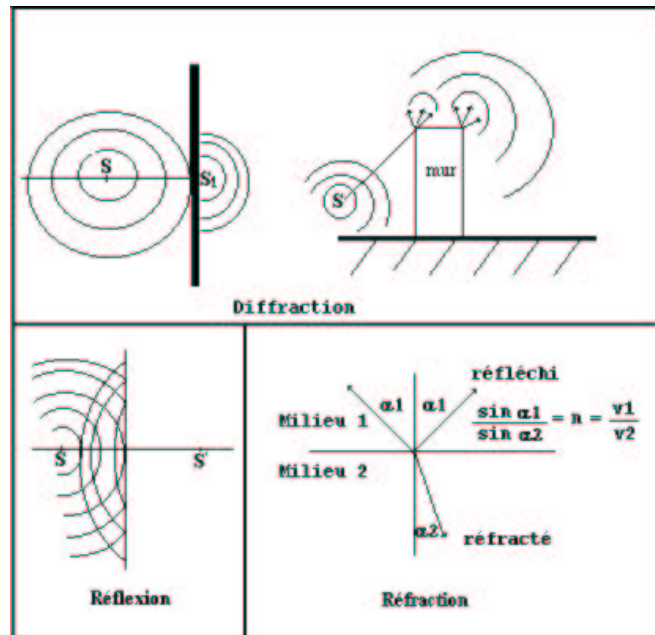


Figure 4.22 : Propagation des sons

Cette propagation répond aux lois de propagation des mouvements vibratoires dans un milieu élastique où le front de l'onde conserve une forme rigoureusement sphérique si le milieu est isotrope (dans un milieu anisotrope le front de l'onde se déforme en se propageant). La propagation se fait en ligne droite dans un milieu homogène et répond face à un obstacle par une réflexion d'une partie et par une diffraction d'une autre partie (l'obstacle se comporte alors lui-même comme une source rayonnante). Le passage d'un milieu à un autre se traduit par une réfraction, c'est-à-dire à un changement de direction en relation avec la vitesse de propagation qui est modifiée.

Un autre aspect important est la forme de cette onde de pression acoustique. Elle se caractérise par l'amplitude de la vibration (ou de la pression) qui varie au cours du temps. Cette forme peut être très simple (cas du son d'un diapason) ou très complexe. Les sons purs se caractérisent par une amplitude instantanée avec une périodicité régulière selon une fonction sinusoïdale. Cela reste vrai quand la vibration est entretenue et qu'il ne se produit donc pas d'amortissement. L'équation d'un son pur peut s'exprimer de la manière suivante :

$$D(t) = A \sin(2\pi ft + q) \quad (4.7)$$

où $D(t)$ amplitude instantanée peut être calculée à partir de A (amplitude maximale), de f (fréquence), de t (temps) et q (phase de départ). Dans la réalité, il existe une résistance à la progression de l'onde qui entraîne un amortissement de celle-ci. La psychoacoustique étudie les conséquences des variations de ces paramètres dans le domaine de la perception. Les changements d'amplitude donnent lieu à des variations de l'intensité perçue appelée sonie. Les changements de fréquence correspondent aux variations de hauteur (plus ou moins grave ou aiguë). La phase, quand elle diffère d'une oreille à l'autre, est un indice de localisation. Les sons complexes correspondent :

- soit à des sons complexes périodiques (cas des instruments de musique ou de la parole, Figure 4.23) ;
- soit à des sons complexes non-périodiques, c'est le cas des bruits.

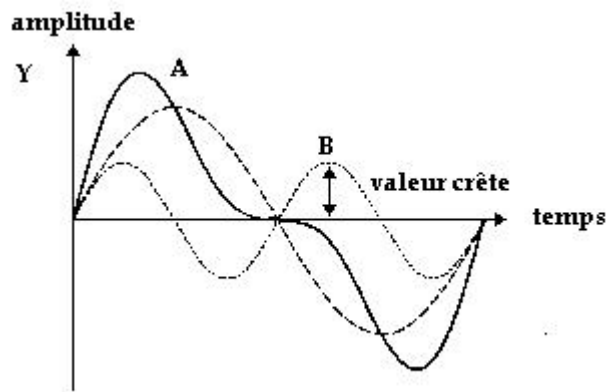


Figure 4.23 : Son périodique mais non sinusoïdal

Remarquons que le bruit se définit par rapport à l'homme. Tout phénomène périodique peut être décomposé en sinusoïdes élémentaires appelées harmoniques. Le bruit blanc comprend toutes les fréquences comprises entre 20 et 20000 Hz avec une amplitude égale, ce qui rend son enveloppe spectrale plate. Par contre l'enveloppe spectrale des voyelles comprend des pics spectraux appelés formants. La composition spectrale plus ou moins riche ainsi que la distribution de l'énergie détermine en grande partie la perception du timbre.

La pression acoustique devrait s'exprimer en Pascal. Cependant la gamme des pressions courantes des bruits va de $2 \cdot 10^{-5}$ Pa à plus de 100 Pa, ce qui n'est pas très pratique. C'est pourquoi en acoustique, on utilise le niveau de pression acoustique, qui est égal à 20 fois le logarithme décimal du rapport entre la pression acoustique considérée à une pression de référence qui a été fixée à $2 \cdot 10^{-5}$ Pa. Elle correspond au seuil de perception d'un son. Ce niveau de pression s'exprime en décibels (dB). A la pression acoustique correspond une certaine puissance acoustique exprimée en Watts et une certaine intensité acoustique exprimée en Watts/m^2 . Cependant, comme pour la pression, l'intensité est exprimée en niveau d'intensité acoustique qui correspond à 10 fois le logarithme décimal du rapport entre l'intensité considérée et une intensité de référence fixée à 10^{-12} W/m^2 . Dans le cas simple des ondes planes progressives, l'intensité acoustique est proportionnelle au carré de la pression acoustique de sorte que les niveaux de pression et d'intensité acoustiques sont égaux. D'où :

$$\text{dB} = N = 10 \lg_{10} I/I_0 = 20 \lg_{10} p/p_0$$

I et p sont respectivement l'intensité et la pression mesurées, I_0 et p_0 l'intensité et la pression de référence. C'est ainsi que 0 dB correspond environ au seuil de l'audition au tour de 1000 Hz et 140 dB au bruit d'un réacteur au banc d'essai. Il est important de retenir que la pression acoustique diminue proportionnellement à la distance, alors que l'intensité diminue proportionnellement au carré de cette distance. Ceci en champ libre, c'est-à-dire en l'absence de réverbération.

La réverbération est la persistance du son dans un espace clos après interruption de la source acoustique. En pratique, on mesure la durée de réverbération qui est le temps correspondant à une décroissance de 60 dB du niveau d'intensité acoustique. La réverbération d'un local est proportionnelle à son volume et inversement proportionnelle au coefficient d'absorption du matériau qui compose les parois. Pour augmenter l'absorption il est possible de recourir aux techniques suivantes :

- utilisation de matériaux poreux non élastiques à pores ouverts comme les panneaux à laine de verre (le polystyrène expansé n'est pas bon car à pores fermés) ;
- utilisation de résonateurs de Helmholtz qui absorbent sélectivement l'énergie acoustique dans une bande de fréquences très étroite ;
- utilisation de diaphragmes ou panneaux résonants permettant l'absorption des sons graves.

Les phénomènes décrits ci-dessus montrent que dans un espace clos les sons en provenance de deux sources différentes ou d'une même source, mais avec recombinaison après réflexion, peuvent se combiner. C'est ainsi que la combinaison de deux sons peut produire :

- une augmentation du niveau (3 dB si les deux sons sont égaux) ;
- des battements si les deux sons ont des fréquences proches ;
- des ondes stationnaires si les sons correspondent à des ondes planes, sinusoïdales, de même amplitude, même fréquence et qui progressent en sens inverse (cas d'une machine émettant des sons purs se réfléchissant sur des murs parallèles).

4.3.3 LE SYSTÈME AUDITIF

4.3.3.1 Anatomie

Le système auditif périphérique comprend :

- l'oreille externe (pavillon et conduit auditif) qui se conduit comme un résonateur et entraîne une amplification des vibrations pour celles comprises entre 1500 et 7000 Hz ;
- l'oreille moyenne (tympan et trois osselets : le marteau, l'enclume et l'étrier) transforme les vibrations aériennes en variations de pression dans les liquides de l'oreille interne ;

- l'oreille interne comprend la cochlée qui a un rôle acoustique et le système vestibulaire qui joue un rôle dans l'équilibration (ce dernier sera décrit dans une autre partie), (Figure 4.24).

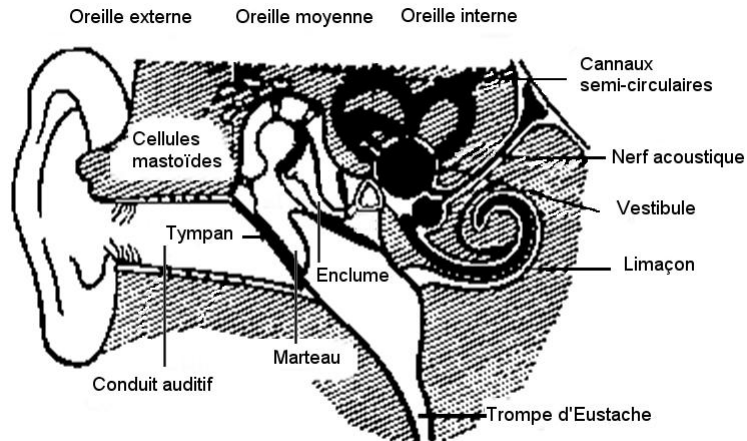


Figure 4.24 : Coupe de l'oreille

L'oreille moyenne se comporte aussi comme un amplificateur et entraîne à partir de 100 Hz un gain qui atteint 30 dB pour les fréquences comprises entre 1000 et 4000 Hz. Pour les stimulations importantes le réflexe acoustique se met en place et entraîne une contraction du muscle stapédien relié à l'étrier. Il se déclenche pour des intensités supérieures à 80 dB et peut diminuer le niveau acoustique de 20 dB. Il constitue un système de protection de l'oreille interne contre les traumatismes sonores et permet d'éviter la saturation du système en augmentant la gamme dynamique de la cochlée.

Dans l'oreille interne, la cochlée ou colimaçon est un tunnel osseux en forme de spirale, divisé sur toute sa longueur en trois compartiments par deux membranes la basilaire et celle de Reissner. On trouve le canal cochléaire contenant l'endolymphe, la rampe vestibulaire et la rampe tympanique contenant la périlymphe et communiquant au sommet par l'apex. Par l'intermédiaire de la fenêtre ovale la platine de l'étrier agit comme un piston et ses vibrations sont communiquées aux liquides cochléaires. Les cellules réceptrices sont situées sur la membrane basilaire et constituent l'organe de Corti. Il s'agit des cellules ciliées dont les cils sont excités par la vibration qui se propage avec un codage spatial (topotopie) et un codage temporel (Figure 4.25).

Le système auditif central comprend à partir du nerf auditif des voies afférentes et efférentes qui passent par le noyau cochléaire, le complexe de l'olive supérieure, le colliculus inférieur, le corps genouillé médian pour arriver au cortex auditif (Figure 4.26).

4.3.4 PHYSIOLOGIE DE L'AUDITION

Les informations sonores en provenance de l'extérieur, après transmission à travers l'oreille interne et l'oreille moyenne, arrivent à la membrane de la fenêtre ovale. Les

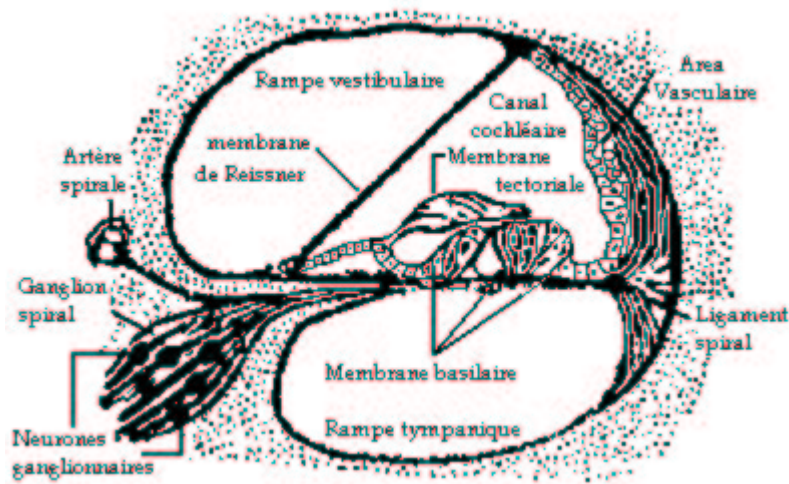


Figure 4.25 : Coupe de l'oreille interne

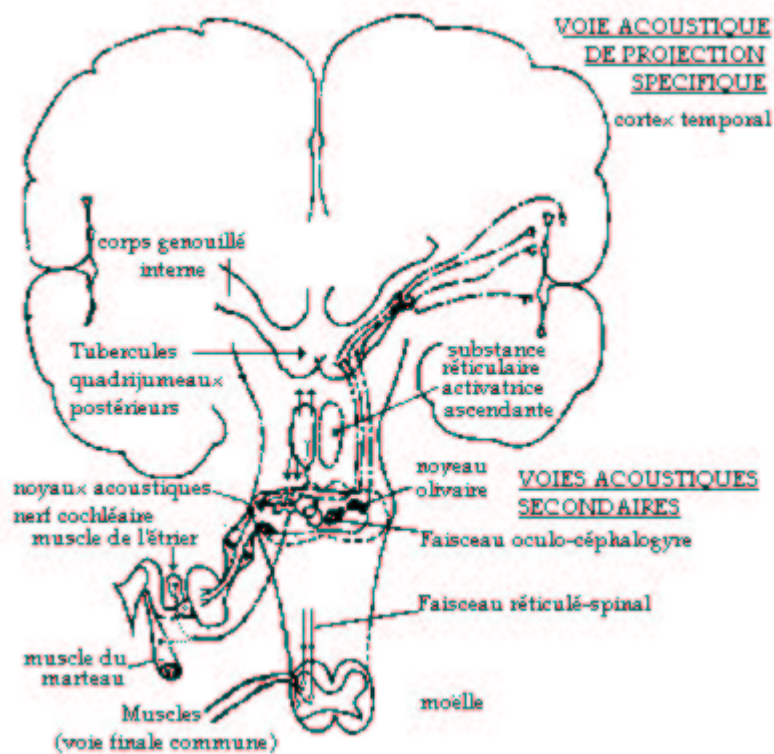


Figure 4.26 : Les voies auditives

mouvements progressent le long de la cochlée où l'énergie est transmise à la membrane basilaire. Cependant la région de la cochlée stimulée par un son dépend de sa fréquence. Les hautes fréquences induisent un déplacement plus important de la membrane basilaire à la base de la cochlée et les basses fréquences excitent surtout la partie apicale. Il en résulte une répartition des fréquences le long de la cochlée : c'est la tonotopie cochléaire (Figure 4.27).

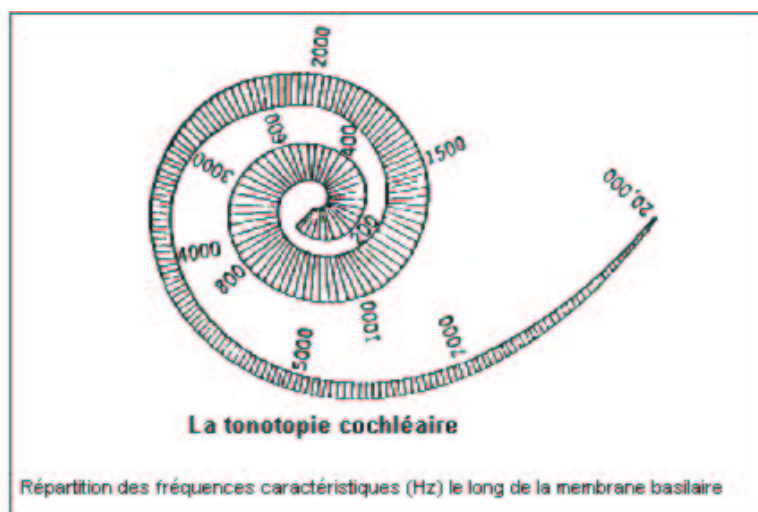


Figure 4.27 : Tonotopie cochléaire

Lorsque les cils des cellules ciliées internes subissent des déformations, les cellules sont stimulées et libèrent un neuromédiateur (glutamate) qui active les neurones du nerf auditif. Si une partie de la transmission est passive, il semble que la contraction active des cellules ciliées externes joueraient un rôle d'amplificateur pour expliquer des propriétés comme la sensibilité à 1000 Hz (capacité de détecter un mouvement de 0,1nm) ou la sélectivité en fréquence (capacité de distinguer deux sons séparés de 3 Hz). Le témoin de cette activité est l'auto émission réémise par la cochlée et enregistrable dans le conduit auditif externe (Figure 4.28).

Les informations parviennent au cortex auditif en conservant l'organisation tonotopique. Le lecteur peut consulter le site [http1], réalisé par l'Université de Montpellier 1, consacré à l'oreille et à l'audition. La localisation spatiale d'une source sonore mérite un développement particulier.

4.3.5 PRINCIPE DE LOCALISATION SPATIALE D'UNE SOURCE SONORE

Pour créer un monde virtuel visuel et sonore compatible, il est préférable de pouvoir simuler spatialement des sources sonores qui se superposent aux images des objets correspondants. Mais comment le cerveau humain localise-t-il les sources sonores virtuelles dans n'importe quel point de l'espace ? L'homme perçoit les sons différemment avec ses deux oreilles. Plusieurs différences sont inconsciemment interprétées par le cerveau. Les deux premières sont le décalage temporel et le rapport d'intensité entre les oreilles (Figure 4.29).

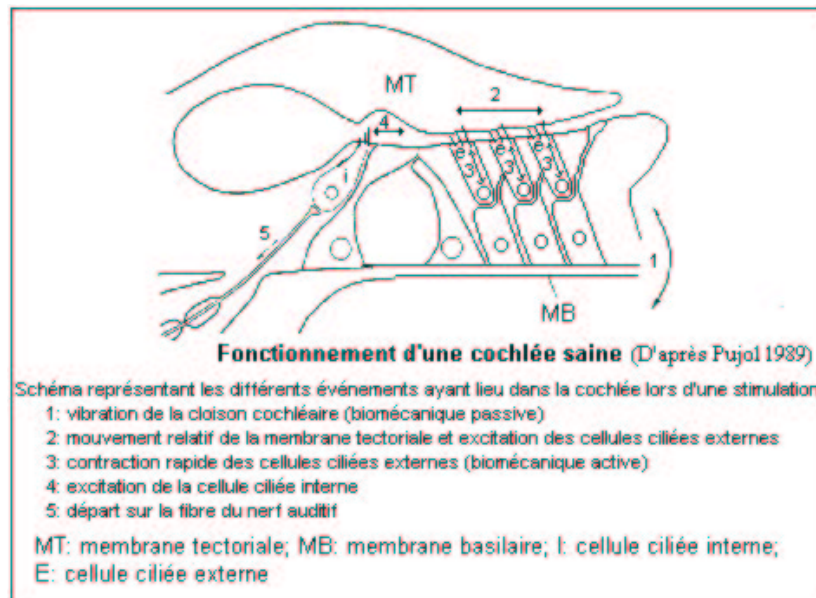


Figure 4.28 : Fonctionnement de la cochlée

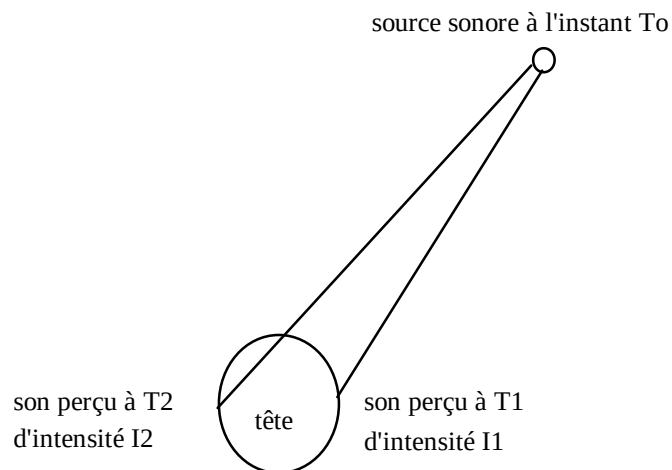


Figure 4.29 : Schéma des propagations des sons aux oreilles

Si on considère les deux oreilles comme de simples récepteurs ponctuels, le décalage de temps T_2-T_1 , de l'ordre de la milliseconde, et la différence d'intensité I_2-I_1 permettent seulement de localiser la source sonore sur un cône d'axe intra-aural, appelé cône de confusion. Les valeurs d'intensités des sons permettent d'estimer la distance de la source sonore. Mais dans ce cas de récepteurs ponctuels, il est impossible d'estimer la position avant ou arrière et en haut ou en bas du son reçu. En réalité les oreilles ne sont pas de simples récepteurs ponctuels, mais sont composées de pavillons (oreilles externes) et de conduits internes qui jouent un rôle important dans l'audition. Le son est réfléchi sur le pavillon et il est guidé par celui-ci vers l'oreille interne. Il est aussi réfléchi et en partie absorbé par la tête et les épaules. Ces phénomènes sont évidemment fonction de la position de la source par rapport aux oreilles. Tout ceci implique une troisième règle : la source sonore produit des variations d'intensité pour les deux oreilles qui sont aussi fonction de la fréquence des sons. Cette différence de perception des deux intensités dépendant de la fréquence constitue une caractéristique pour la localisation de la source.

Ces phénomènes de localisation ont été étudiés depuis quelques dizaines d'années et permettent de créer artificiellement des sources sonores virtuelles dans l'espace. Le lecteur ne doit pas confondre cette technique avec celle du son stéréophonique qui ne crée des sons que sur un plan vertical intra-aural. Les études ont montré que chaque personne a sa propre «fonction de transfert auditif de la tête» qui donne la relation entre la différence de perception des intensités I_1 et I_2 et la fréquence du son. Une source sonore peut être localisée, du moins si la personne veut bien y prêter attention et si la source n'est pas trop environnée d'autres sources sonores parasites. Des illusions intersensorielles sont toujours possibles entre la vue et l'ouïe. Par habitude, nous positionnons au niveau de la bouche la parole d'une personne. Mais le ventriloque peut donner l'illusion que sa marionnette parle à sa place en faisant bouger les lèvres de la marionnette tandis que ses lèvres restent immobiles. De même, par effet Doppler (variation de la fréquence d'un son en fonction de la vitesse de la source sonore), il est possible de suggérer que la source sonore provienne de l'objet en mouvement observé sur l'écran (voiture sur une route virtuelle, par exemple) même si la source sonore est fixe.

4.3.6 LA PERCEPTION AUDITIVE

En terme de perception, on en distingue trois types :

- la *perception tonale ou tonie* qui est la sensation auditive liée à la fréquence des sons. Le seuil différentiel de fréquence est de 3 pour mille de 500 à 4000 Hz. L'oreille humaine peut percevoir environ 1800 sons différents ;
- la *perception de la force sonore ou sonie* qui est la sensation auditive liée à la pression et à la structure du son. Pour une même pression, la sonie varie avec la fréquence. La sensibilité de l'oreille est maximale entre 1000 et 6000 Hz. Le seuil différentiel est de l'ordre de 15 à 20% de 100 à 4000 Hz ;
- la *timbre* qui est la sensation auditive permettant de distinguer divers sons complexes de même fréquence fondamentale, mais de composition spectrale différente. Il est donc lié à la richesse du son en harmoniques.

4.3.6.1 Audiométrie

Pour mesurer les sensations et les perceptions auditives plusieurs techniques existent :

- acoumétrie : test de la montre, de la voie chuchotée, du diapason (tests de Weber et de Rinné) ;
- audiométrie tonale. Elle s'attache à mesurer les seuils auditifs (le plus faible niveau de stimulation perceptible) en sons purs, c'est-à-dire pour des fréquences spécifiques ;
- audiométrie vocale. Elle s'attache à mesurer le seuil d'intelligibilité qui est le niveau de stimulation pour lequel la moitié des mots d'une liste prédéfinie sont identifiés ;
- audiométrie objective : PEA (potentiels évoqués auditifs), mesure de l'impédance de l'oreille, électrocochléogramme, réponse électrodermale.

4.4 SENSIBILITÉ CUTANÉE

Dans la plupart des tissus du corps humain (les muscles, la peau, les articulations, les tendons, etc.), il y a des récepteurs pouvant être activés mécaniquement. Leurs stimuli sensoriels constituent la somesthésie, qui est séparée en deux parties : la sensibilité cutanée et la proprioception (sensibilité interne).

La sensibilité cutanée est composée du toucher, ainsi que de la sensibilité à la pression et aux vibrations. L'absence du toucher dans un monde virtuel ôte une source d'informations, ce qui peut être gênant pour certaines simulations virtuelles ainsi qu'en téléopération. Si l'ensemble de la peau participe au sens du toucher, il est important de noter qu'il est surtout efficace au niveau des mains : grâce à celles-ci, on peut déterminer par toucher la géométrie de la surface en contact, sa rugosité et son mouvement relatif. La sensibilité superficielle englobe aussi la sensibilité thermique qui est surtout une sensibilité à la variation de température. Les sensations qui apparaissent lors d'un changement de température dépendent de l'étendue de la surface stimulée, de la température initiale et de la vitesse de changement.

Pour le sens du toucher, jusqu'à ce jour, il n'y a pas à proprement parler d'ambiance physique qui ait fait l'objet d'une réification des sensations tactiles [Piéron, 1964]. Les sensations cutanées correspondent à une stimulation mécanique de pression exercée au niveau de la surface corporelle par un corps plus ou moins solide qui rentre en contact avec le corps humain. La rencontre se fait au cours du déplacement d'un objet vers le corps, d'une partie du corps vers un objet, des deux en mouvement l'un vers l'autre. En fonction de la vitesse de rencontre et de la répétition de la rencontre, il y a des sensations, même pour des intensités faibles, dites de contact, de pression ou de vibration, cette dernière pour des variations rapides de pression. D'autres capteurs plus profonds révèlent des pressions statiques durables et correspondent au sens kinesthésique qui renseigne sur la situation des parties du corps par rapport à nous-mêmes. D'autre part les extrémités du corps, bras, mains, jambes et pieds ont une double fonction, ils sont anatomiquement faits pour sentir mais aussi pour faire. Cela est maximal au niveau des mains [Gibson, 1966]. Cependant les pieds ont pour rôle fondamental le déplacement de l'individu, c'est-à-dire la locomotion. Celle-ci pour se réaliser a besoin d'informations concernant la verticalité et les accélérations qui si elles sont en partie fournies par les capteurs kinesthésiques, le sont en plus grande part par le vestibule où se trouvent

les canaux semi-circulaires. Dans ce chapitre la locomotion ne sera pas traitée. Par contre, la manipulation manuelle sera traitée dans ce chapitre car la main est par excellence le lieu du sens du toucher, même si toute la peau y participe et même si pour manipuler il faut conserver une posture qui est contrôlée par le vestibule.

Dans le cadre de la réalité virtuelle sans retour tactile et d'effort, les objets solides qui composent les environnements virtuels sont des fantômes d'objets réels. Ils ont en effet perdu leur consistance même s'ils restent visibles. L'immersion et le déplacement virtuel dans un environnement purement visuel fait que l'on peut passer dans certains cas à travers les objets. Les interfaces avec retour tactile permettent de mieux «observer et connaître» l'objet, mais ne permettent pas de le manipuler. Par contre les interfaces avec retour d'effort permettent, dans certains cas, de mieux le manipuler. S'il existe des interfaces tactiles pour des objets solides, il n'y a pas d'interfaces qui redonnent les sensations de contact avec l'eau ou avec l'air.

C'est pourquoi ce chapitre sera consacré, après un rappel de la physique de la matière et surtout des solides, à la description anatomique de la peau et de la physiologie des capteurs sensoriels cutanés, puis à la physiologie de la main.

4.4.1 *PHYSIQUE DE LA MATIÈRE*

4.4.1.1 *Les objets solides*

En réalité virtuelle, comme dans le monde réel, on peut considérer un objet comme un solide indéformable ou comme un solide déformable. Dans le deuxième cas, les difficultés de simulation sont évidemment plus délicates à réaliser. Les éléments de catégorisation d'un objet peuvent être décrits par les propriétés de la substance (de quoi il est constitué) et de la structure (comment il est arrangé). La texture, la consistance (dureté, élasticité, friabilité), les caractéristiques thermiques (conductivité, puissance), le poids et la densité sont les propriétés de la substance. La forme, (surfactive et volumique), les dimensions et le volume sont les propriétés de la structure.

4.4.1.2 *Les échanges thermiques dans l'air*

La composition de l'air et les propriétés physiques de celui-ci sont à prendre en compte car l'homme du fait qu'il est homéotherme se comporte comme un échangeur thermique. Les échanges se font par conduction, convection et évaporation. La neutralité thermique est obtenue lorsqu'il y a équilibre thermique, c'est-à-dire lorsqu'il n'y a ni stockage ni déstockages de chaleur. Ceci se produit quand la totalité de la sueur est évaporée, ce qui correspond à une mouillure de la peau inférieure à 60 %. Les sensations de chaud ou de froid apparaissent lorsqu'on s'écarte de cette neutralité thermique. Dans ces sensations interviennent non seulement la température ambiante proprement dite mais aussi l'humidité et le vent. L'apport de chaleur se fait par radiations atmosphériques, (lumière directe et réfléchie des nuages ou de la terre ou lumière dispersée) et par les radiations thermiques du sol.

Ces phénomènes expliquent pourquoi il faut prendre en compte plus que la température donnée par le thermomètre de nos maisons pour évaluer les effets des ambiances

thermiques sur l'organisme. C'est ainsi que pour la chaleur, la température à prendre en compte est «la température effective» qui est donnée par l'index WBGT qui = $0,7 TW + 0,2 TG + 0,1 TD$ et où les températures sont exprimées en °C. TW correspond à la température humide (W pour wet). Elle est obtenue par un thermomètre entouré d'un linge humide. Elle prend en compte la vitesse du vent qui diminue l'humidité du linge. TG correspond à la température de globe, elle est obtenue par un thermomètre situé dans une sphère noire. Elle rend compte de la température radiante. TD correspond à la température sèche (D pour dry). Elle est obtenue par un thermomètre ordinaire. C'est ainsi qu'il est possible d'avoir la même sensation de chaleur par exemple avec approximativement 42 °C au thermomètre ordinaire, 20 % d'humidité et une vitesse de vent à 2 mètres par seconde (en plein désert) et 33 °C, 80 % d'humidité et un vent à 3 mètres par seconde (en bordure de mer). En température effective, dans les deux cas, elle est de l'ordre de 30 °C.

Pour la sensation de froid, du fait que ses effets sont d'autant plus marqués que la baisse de température s'accompagne de vent, on utilise pour apprécier le pouvoir de refroidissement d'une ambiance le Wind Chill Index (W.C.I.).

$$W.C.I. : Ko = (v * 100 + 10,45 - v)(33 - T)$$

où Ko est exprimé en Kcal/m².h, T en °C, v en m/s.

4.4.1.3 L'eau et les liquides

Une des propriétés intéressantes de l'eau est sa possibilité d'être pénétrée comme l'air et donc de s'y mouvoir. Cependant un corps immergé est soumis à une pression qui augmente d'un bar tous les 10 mètres. Cette pression s'exerce de manière identique sur toute la surface du corps. Cette propriété est utilisée pour donner une apesanteur relative et entraîner les cosmonautes aux vols spatiaux. Si peu ou pas de système de réalité virtuelle redonne les sensations au contact de l'eau, il existe de nombreuses réalisations pour représenter visuellement celle-ci dans les environnements virtuels. Il serait intéressant de pouvoir «toucher» ces images.

4.4.2 LA PEAU

A côté des sensations liées à l'identification des zones de contact, la peau renseigne sur les propriétés de l'objet touché (chaud, piquant, rugueux, lourd) et peut entraîner des sensations subjectives de douleur, type centalgie (piqûre), crousalgie (pincement) ou thermalgie (brûlure). En plus de recevoir des informations, la peau humaine protège les tissus profonds et régule la température. Et si sa sensibilité de contact est la plus complexe au niveau de la face palmaire de la main [Tubiana e.a., 1990], elle intéresse toute la surface du corps. Enfin la peau, organe sensoriel a un statut particulier. Anzieu [Anzieu, 1985] rappelle qu'elle est toujours en contact avec l'environnement, on ne peut pas isoler le toucher comme on peut le faire avec les autres sens : fermer les yeux pour ne pas voir, se boucher les oreilles pour ne pas entendre ou se pincer le nez pour ne pas sentir.

4.4.2.1 Anatomie

La peau se divise en derme et épiderme. L'épiderme comporte, de la profondeur vers la surface, différentes couches : la couche basale, le stratum Malpighi, la couche granuleuse et la couche cornée entre la granuleuse et la couche cornée s'interpose le stratum lucidum. Le derme est un tissu conjonctif avec une couche superficielle, un derme moyen, un derme profond et un plexus vasculaire superficiel. Il contient en plus des récepteurs sensoriels, les glandes sudoripares et les poils avec leurs glandes sébacées et leurs muscles horripilateurs.

Les mécanorécepteurs sont localisés dans toutes les régions du corps. Ils peuvent prendre la forme de terminaisons nerveuses libres, de terminaisons nerveuses encapsulées sous la forme de minuscules corpuscules, bulbes, axiales ou cylindriques. De plus il y a les cellules des poils microscopiques qui ont une sensibilité particulière au mouvement. La combinaison et la distribution de plusieurs de ces récepteurs dans une région particulière de la peau produisent une haute spécialisation des structures sensorielles. Ces structures sensorielles sont souvent réparties dans le bas de la peau, pénétrant la couche de support et le tissu conjonctif et interpénétrant tous les muscles et les tendons attachés aux os et peuvent être trouvés en écharpes autour des vaisseaux sanguins. Il y a deux types majeurs de peau (pileuse et galbe). Des récepteurs différents sont associés à chaque type. La peau galbe se trouve à la plante des pieds et à la partie palmaire des mains. La majeure partie des autres surfaces cutanées est pileuse.

Dans la peau pileuse les couches dermiques contiennent les fibres nerveuses à terminaisons libres dont les terminaisons entourent les follicules pileux. A côté des poils se trouvent les disques de Merkel et les corpuscules de Ruffini. Dans la peau galbe, en plus des fibres nerveuses à terminaisons libres, se trouvent des récepteurs encapsulés. Ainsi, la peau palmaire présente des caractères particuliers dont l'importance fonctionnelle est très grande. Elle a une structure papillaire qui fait que la ligne de jonction entre le derme et l'épiderme est marquée par des saillies dermiques et épidermiques engrenées.

Parmi ces récepteurs encapsulés, il y a les corpuscules de Meissner qui sont logés dans les papilles et permettent grâce à leur disposition perpendiculaire à la peau de participer à la précision du tact. D'autres terminaisons nerveuses spécialisées (extérocepteurs) existent dans les couches suivantes. Il est possible de distinguer : les disques de Merkel qui sont des extensions non myélinisées d'axones myélinisés et les corpuscules de Pacini qui sont des terminaisons nerveuses encapsulées correspondant à une seule fibre nerveuse qui se trouve dans la couche profonde de la peau.

En résumé on trouve :

- Au niveau de l'épiderme
 - les fibres nerveuses à terminaisons libres ;
 - les disques de Merckel.
- Au niveau du derme
 - les corpuscules de Meissner ;
 - les corpuscules de Pacini ;
 - les corpuscules de Krause ;

- les corpuscules de Ruffini.

Les fibres afférentes forment le système lemniscal (constitué par les neurones périphériques articulés aux récepteurs cutanés articulaires et musculaires) qui va se diriger via les faisceaux des colonnes de moelle épinière sur les noyaux graciles et cunéiformes. De là le faisceau formé par le lemniscus médian va rejoindre le noyau ventrobasal du thalamus et ensuite le cortex cérébral pariétal, appelé cortex somesthésique, qui est spécialisé dans la réception et le traitement des informations concernant le tact mais aussi la proprioception. Cela est schématisé sur la Figure 4.30.

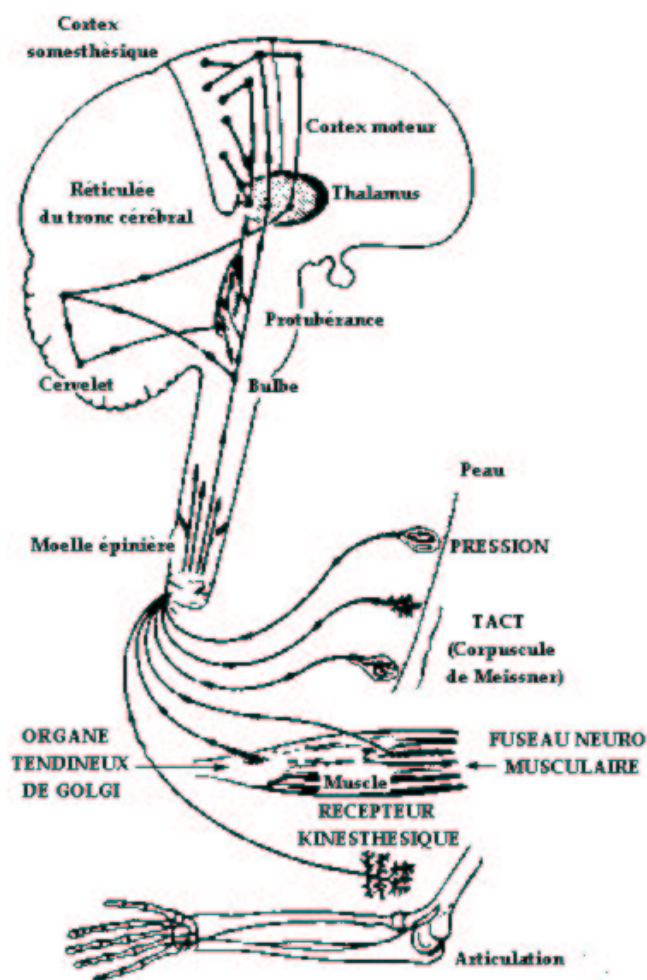


Figure 4.30 : Voies sensibles tactiles et kinesthésiques

4.4.2.2 *Physiologie*

D'une règle générale, la perception cutanée est la résultante de la déformation de la peau liée à une poussée qui excite mécaniquement les terminaisons nerveuses pour donner les sensations de pression, de tact, de température et aussi de douleur. En effet, les différentes formations nerveuses décrites sont hautement spécialisées. En fonction de leur faculté de réponse face à la vitesse, l'accélération, le déplacement ou l'ébranlement [Sherrick e.a., 1986], il est possible de distinguer :

- des mécanorécepteurs à adaptation lente ou filtre passe-bas bien adaptés à la détection de la force de déplacement :
 - disques de Merckel ;
 - corpuscules de Ruffini ;
- des mécanorécepteurs à adaptation rapide se comportant comme des filtres mécaniques passe-haut. Ils sont très sensibles à la pression :
 - corpuscules de Vater-Pacini ;
 - corpuscules de Meissner ;
 - corpuscules de Krause.
- des terminaisons nerveuses libres thermo-algésiques sensibles aux piquûres, au chaud et au froid.

4.5 LES AMBIANCES CHIMIQUES : L'ODORAT ET LE GOÛT

Les stimuli chimiques se distinguent des autres stimuli agissant sur les organes des sens car ils ne sont pas fonction de quelques paramètres, comme par exemple trois couleurs de base pour la vision. De même il n'est pas possible de les décomposer en variables numériques sur un ordinateur. Pour cela la simulation d'ambiance chimique est à classer à part dans les environnements virtuels. En effet les techniques actuelles nécessitent l'utilisation de produits naturels qui sont consommés au cours de l'immersion et qu'il faut renouveler à chaque nouvelle immersion. Cependant, et seulement pour les odeurs, certaines applications sont en train de voir le jour, c'est pourquoi, il semble intéressant de ne pas laisser ces perceptions de côté.

Les chémorécepteurs se trouvent au niveau des organes de l'olfaction et de la gustation mais aussi dans divers types de sites contrôlant la composition chimique du sang. Les organes situés dans le nez permettent de sentir les odeurs. Les organes de la bouche, mais en contact par les choanes⁵ avec le nez, permettent d'apprécier les saveurs.

4.5.1 *LES ODEURS*

Comme pour les ambiances sonores qui correspondent à ce que l'on entend, les ambiances odorantes correspondent aux produits que l'on sent. Il est nécessaire pour qu'un produit puisse être senti qu'il présente un degré moyen de volatilité. Les molécules de produits volatils véhiculées dans le courant respiratoire passent dans les fosses nasales.

⁵Orifices postérieurs des fosses nasales qui les fait communiquer avec le rhino-pharynx.

Elles sont absorbées au contact d'une région pigmentée en jaune de la muqueuse revêtue d'une couche de cellules nerveuses. La question se pose de savoir s'il existe des cellules spécialisées pour chaque odeur ou pour un groupe d'odeur. La classification d'Amoore retient 7 odeurs primaires (camphre, musc, rose, menthol, éther, vinaigre, putréfaction), qui correspondraient chacune à un récepteur spécifique. La classification de Zwaardemaker a pour catégories : étherées, aromatiques, balsamiques, ambrosiaques, alliées, empyreumatiques, capryliques, repoussantes et nauséuses. Pour Crocker et Henderson (1927), il y aurait quatre composantes fondamentales : balsamique (florale), étherée (fruité), caprylique (hircine, de bouc) empyreumatique (de brûlé). Leurs combinaisons détermineraient les quelques 3 000 odeurs reconnaissables pour certains auteurs et 400 000 pour d'autres. Les odeurs sont analysées par le rhinencéphale, qui intervient par ailleurs dans la mémoire, les émotions, les instincts, le comportement. Les études les plus récentes font penser que les odeurs sont spécifiques et non décomposables en quelques odeurs primaires.

4.5.1.1 Anatomie de l'appareil olfactif

Les fosses nasales, région la plus profonde de la cavité du nez, contiennent au niveau du plafond, de chaque côté, une petite zone d'environ 3 cm² tapissée d'une muqueuse spécialisée dans l'olfaction, la muqueuse olfactive. Cette dernière est constituée de cellules sensorielles qui sont des neurones à gros noyau. Il y aurait 20 à 40 millions de récepteurs.

La dendrite de la cellule réceptrice est munie en surface de petits cils qui baignent dans le liquide nasal humectant la muqueuse. Les axones quant à eux se regroupent en filets nerveux et, traversant de petits orifices à la base du crâne (toit des fosses nasales), se terminent immédiatement après dans les glomérules où ils rentrent en contact avec les dendrites des cellules mitrales qui constituent le bulbe olfactif (petit renflement situé à la face inférieure du lobe frontal de l'hémisphère cérébral). C'est la première partie du nerf olfactif, qui apporte les informations au cerveau au niveau du rhinencéphale (Figure 4.31).

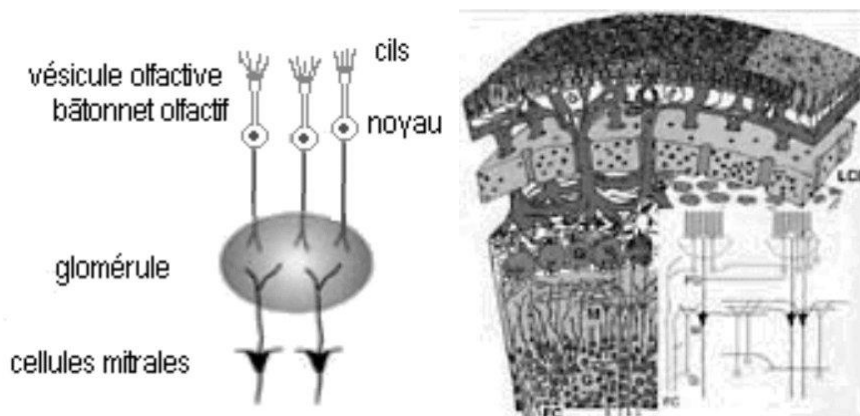


Figure 4.31 : Détails des cellules olfactives et l'organisation générale

4.5.1.2 *Physiologie de l'odorat*

La physiologie de l'odorat reposerait sur les mêmes principes que celle du goût, mais ces domaines sont encore imparfaitement connus. Plusieurs théories ont été développées. Dans la théorie physico-chimique de Bowman, les molécules volatiles doivent se dissoudre dans le mucus et, ainsi captées, pénétrer dans les cils cellulaires à constitution lipoïdique grâce à une liposolubilité supérieure à la solubilité aqueuse et qui serait caractéristique de toutes substances odorantes. Dans la théorie chimique, il y a substance odorante que si cette substance possède des valences libres. La théorie physique soutient que les cellules olfactives produisent des infrarouges qui sont plus ou moins absorbés par les substances chimiques. Ce serait la quantité des ondes infrarouges absorbées par ces substances qui donnerait la sensation d'odeur. La théorie électrostatique prétend que la sensation odorante est provoquée par une rupture de charges électriques entre les molécules du corps chimique et celles de la muqueuse olfactive.

Les théories actuelles rejettent le fait que les cellules olfactives soient des détecteurs de propriétés physico-chimiques fondamentales (poids moléculaire, densité électronique, polarisation, présence de tel ou tel groupe fonctionnel etc.). Il semblerait plutôt que les sites actifs des récepteurs sont configurés pour reconnaître des molécules et non des propriétés isolées [Holley e.a., 1994]. La fixation des molécules se ferait donc comme pour les virus sur des sites particuliers (la forme de la molécule s'emboîterait sur la surface du cil). Toutes les régions de l'épithélium olfactif ne sont pas également sensibles à tous les odorants. Il y aurait une chimiotopie qui se retrouve au niveau du bulbe olfactif.

D'une règle générale, une inspiration normale entraîne un passage de 7 % des molécules sur la muqueuse olfactive et dans l'inspiration brève et rapide (flairage) 30 %. Il n'y aurait pas de différence entre des molécules arrivant par l'air ou par l'eau. Les données quantitatives ont été réalisées pour 200 substances. La variable est la concentration [C] de la substance odorante dans l'air ou dans l'eau. La vitesse du flux porteur est importante car elle détermine la proportion de molécules arrivant sur la muqueuse. Plus la vitesse est grande, plus le pourcentage est grand et permet d'aller plus haut dans le nez. Pour les seuils, il semble que des dilutions au milliardième suffisent (seul le compteur Geiger est compétitif avec le nez). Au-dessus des seuils, des stimulations de [C] croissante donne des sensations croissantes jusqu'à un certain plafond. La différence entre le seuil et la concentration maximum est peu étendue de 1 à 50. Pour certaines substances, l'odeur perçue varie avec la concentration [C]. C'est ainsi que l'indole dilué donne une sensation de jasmin et concentré une sensation fécaloïde, l'ionone concentrée donne une sensation de terre humide et diluée une sensation de violette, la triméthylamine diluée donne une sensation de crevette cuite et concentrée une sensation de pourri ammoniacal. La latence d'apparition de la sensation serait de l'ordre de la demi-seconde et quand la stimulation persiste, la sensation disparaît. Cette adaptation existerait aussi pour les odeurs proches.

4.5.1.3 *Les méthodes de mesures de la perception olfactive*

On distingue l'olfactométrie (étude des phénomènes de la sensibilité olfactive) et l'odométrie (étude de l'activité des corps odorants) Pour cela, il existe de nombreuses techniques qui sont le plus souvent empiriques et indirectes. La technique des mouillettes.

On dispose une goutte de substance plus ou moins diluée sur du papier que l'on fait ou pas sécher. Ce papier est renflé par un groupe de personnes. Le seuil est donné pour un taux de détection de 50 %. L'olfactomètre de Zwaardemaker : la substance est déposée sur un cylindre en matière poreuse, à l'intérieur de ce cylindre il y a un tube en verre qu'on décale vers l'extérieur. On mesure en cm à partir de la perception (unité en longueurs). La technique d'insufflation d'Elsberg : on insuffle une quantité donnée d'air qui est passe dans une solution odorante. En fait comme la vitesse joue, ce n'est pas une bonne méthode. La technique de dilution gazeuse contrôlée : c'est l'olfactomètre de Woodrow qui est une adaptation de celui d'Elsberg, il y a passage à travers des flacons mais la quantité d'air est renflée par le sujet, c'est une méthode plus qualitative que quantitative.

4.5.2 LES SAVEURS

Il n'y a que 4 saveurs de base, chacune étant ressentie préférentiellement dans une zone particulière de la langue. De manière approximative, les sels, plus précisément les ions métalliques, ont un goût salé, les oses et diholosides un goût sucré, les ions H^+ un goût acide et les alcaloïdes un goût amer. Les sensations gustatives sont fournies par des substances chimiques en solution ou des substances solides qui se dissolvent dans la salive et ainsi peuvent pénétrer dans les cellules réceptrices.

4.5.2.1 Anatomie

Les cellules réceptrices sont groupées dans des les bourgeons gustatifs, organes microscopiques de forme ovale, qui sont réunis pour constituer les papilles. Celles-ci sont de petites saillies visibles sur la surface de la langue. En particulier les papilles caliciformes qui forment une douzaine de petits cercles dessinant à la base de la langue un V dont la pointe est tournée vers l'arrière. Cependant on trouve aussi de ces bourgeons dans le palais et le pharynx. Les cellules réceptrices possèdent à leur extrémité antérieure de fins bâtonnets qui se juxtaposent dans un pore du bourgeon au contact de la couche humide que revêt la muqueuse linguale. Elles sont de plus entourées par des plexus serrés formés par d'arborisations terminales de fibres nerveuses appartenant à des nerfs crâniens (lingual et glosso-pharyngien) dont les corps cellulaires sont situés dans les ganglions géniculé et pétreux. Ces fibres nerveuses se comportent comme les fibres de la sensibilité tégumentaire aux niveaux tactile, thermique et algique.

L'amer est perçu à la base de la langue, le sucré à la pointe, l'acide sur les côtés, le salé sur la plus grande partie de la surface. Ces saveurs se combinent entre elles (par exemple, le sucré masque l'amer, Figure 4.32).

4.5.2.2 Physiologie

La stimulation des cellules sensorielles du goût se ferait de la même manière que pour celles de l'odorat, c'est-à-dire par captation dans les bâtonnets de substances ayant une solubilité lipoïdique. La transmission sous forme d'influx nerveux se ferait par l'intermédiaire d'un neurone. Les neurones correspondant à la langue cheminent ensuite

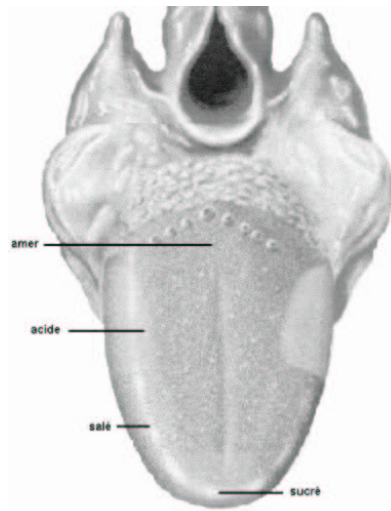


Figure 4.32 : La langue avec les différentes zones du goût

dans deux nerfs crâniens différents : le nerf facial pour la partie antérieure, le glosso-pharyngien pour la partie postérieure. Cela explique que ces deux régions de la langue pourraient présenter indépendamment des anomalies du goût. Les messages arrivent ensuite au cerveau et atteignent la conscience. Le pneumogastrique est aussi concerné. La langue est aussi tactile. Elle permet très finement de détecter, la consistance et la température des aliments. En fait la saveur d'un aliment est la résultante d'une sensation globale incluant le tactile, le thermique, le goût et l'odorat qui d'ailleurs est beaucoup plus sensible et discriminatif.

4.6 PROPRIOCEPTION

4.6.1 INTRODUCTION

Les trois domaines liés à la proprioception (sensibilité mécanique et interne du corps) sont les sensibilités à la position dans l'espace, au mouvement du corps et aux forces exercées sur les muscles. Les deux premiers domaines correspondent au sens kinesthésique.

Pour avoir l'illusion d'être immergé, il est nécessaire que le monde virtuel qui entoure le sujet soit présenté par rapport à ce dernier. Il perçoit le monde en étant au centre de ce monde. Pour cela, il peut être important de restituer dans le monde virtuel la position exacte du corps du sujet immergé. Il faudra donc disposer de capteurs permettant de transmettre à l'ordinateur générant le monde virtuel les informations concernant la position du corps et de ses segments dans le monde réel où il se trouve. D'autre part, compte tenu de la sensibilité humaine aux déplacements dans les trois axes Z, X et Y, il est aussi nécessaire de restituer au sujet immergé les sensations liées à l'équilibration (perception de la verticale) et aux mouvements simulés. Ceux-ci dépendent de la gravité, des accélérations et des déplacements relatifs. La conscience du mouvement

est donnée par la position des différentes parties du corps, leur direction, leur taux de mobilité ainsi que par les forces de contractions musculaires nécessaires à la résistance aux mouvements ou aux poids des objets [Stuart, 1996]. Les organes qui participent à ces sensations sont, en plus des organes proprioceptifs des muscles, des tendons et des articulations (voir la partie sur le toucher), ceux situés au niveau du vestibule de l'oreille interne et, pour une part non négligeable, la vision.

C'est pourquoi ce chapitre comprendra un rappel de l'anatomie et de la physiologie de l'appareil vestibulaire. L'exposé des effets des accélérations sur l'être humain sera traité au chapitre 16. Ensuite sera abordé le problème de la localisation perceptive.

4.6.2 PHYSIQUE DE LA GRAVITÉ ET DES ACCÉLÉRATIONS

La mise en mouvement d'un corps se traduit, rappelons le, par une accélération exprimée en m/s^2 ou en G ($1 \text{ G} = 9,8 \text{ m/s}^2$). L'effet sur le corps humain des accélérations dépend de plusieurs facteurs :

- l'intensité ; en général les effets sont d'autant plus marqués que l'intensité est grande, mais ceci doit être pondéré en fonction de l'interaction avec les autres facteurs,
- la vitesse d'application ; c'est-à-dire le temps mis pour atteindre le plateau d'application ; ce facteur joue pour les accélérations brèves (on en rend compte en Jolt qui s'exprime en G/s),
- la direction par rapport aux axes du corps ; la définition de ces axes est résumée sur la figure suivante. Il est possible de distinguer les accélérations liées aux trois axes X, Y, Z et par rapport à ces axes les rotations autour de X, Y, Z. Sur chacun de ces axes et sur chacune de ces rotations s'exercent des vecteurs d'accélération ou de décélération et des forces d'inertie (Figure 4.33).
- la durée d'application ; on distingue deux cas : les accélérations de longue durée (supérieures à 0,1 s) et celles de durée brève ; les secondes ont surtout des effets mécaniques alors que les premières perturbent essentiellement les fonctions hémodynamiques. Nous verrons ces effets en plus détaillé au chapitre 16.

4.6.3 L'APPAREIL VESTIBULAIRE ET LES VOIES KINESTHÉSIQUES

L'appareil vestibulaire est constitué d'une série de renflements et de trois canaux semi-circulaires orthogonaux situés dans les trois plans de l'espace. L'intérieur des renflements rempli d'un liquide gélatineux est tapissé de cellules ciliaires dont les prolongements sont courbés dans le sens inverse du mouvement. Les variations de courbure font décharger les récepteurs à la base du cil. Si la tête est au repos, le message nerveux renseigne sur l'inclinaison de la tête par rapport à la direction de la verticale. Si la tête est en mouvement, le message nerveux renseigne sur l'accélération totale subie par la tête, composée de l'accélération du mouvement et de l'attraction terrestre. Ceci permet de comprendre qu'en inclinant une personne dans une cabine de simulateur, on peut simuler indirectement une accélération de mouvement en présentant visuellement à la personne un déplacement virtuel. Dans le cas des accélérations linéaire (translation), ce sont les récepteurs maculaires (utricule et saccule) qui déchargent. Les canaux semi-circulaires sont sensibles aux accélérations angulaires (rotation).

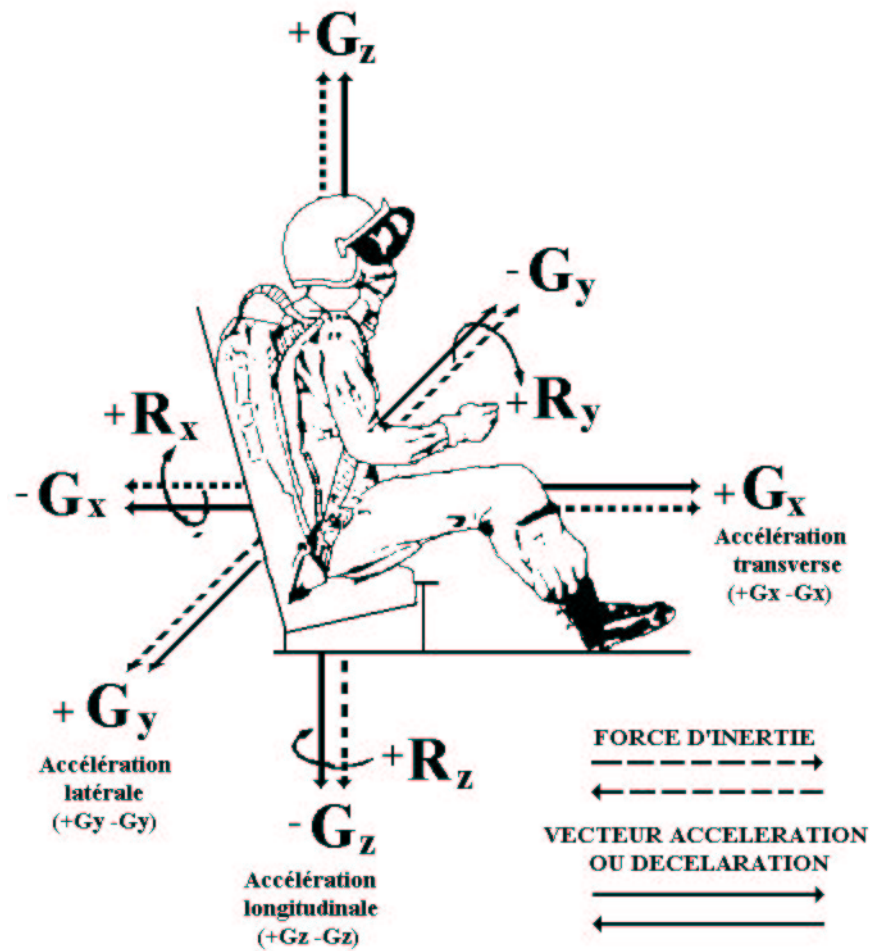


Figure 4.33 : Classification des accélérations en fonction des axes du corps

L'activité posturale anti-gravifique, de lutte contre la pesanteur permet la station érigée et demande des informations renseignant sur la surface d'appui (Figure 4.33), le poids du corps et la représentation visuelle de la verticale. D'autres récepteurs proprioceptifs, répartis dans tout notre organisme, sont sensibles à la position ou au déplacement des différentes parties du corps. Dans les muscles, les récepteurs sont les fuseaux neuromusculaires et les organes tendineux de Golgi. Les récepteurs articulaires répondent à des variations d'étirement et sont donc sensibles au mouvement relatif entre deux parties du corps. Les informations de tous les récepteurs proprioceptifs sont analysés par le cerveau avec des informations sensorielles transmissent par la peau, par les yeux et par l'oreille interne. Ainsi le cerveau en tire une connaissance de la position et du mouvement du corps.

La sensibilité aux forces exercées sur les muscles donne une information globale sur les forces de contact entre la personne et un objet. Compte tenu de l'importance de la manipulation des objets dans les environnements virtuels, ces notions sont reprises et développées au paragraphe 4.8 (Figure 4.34).

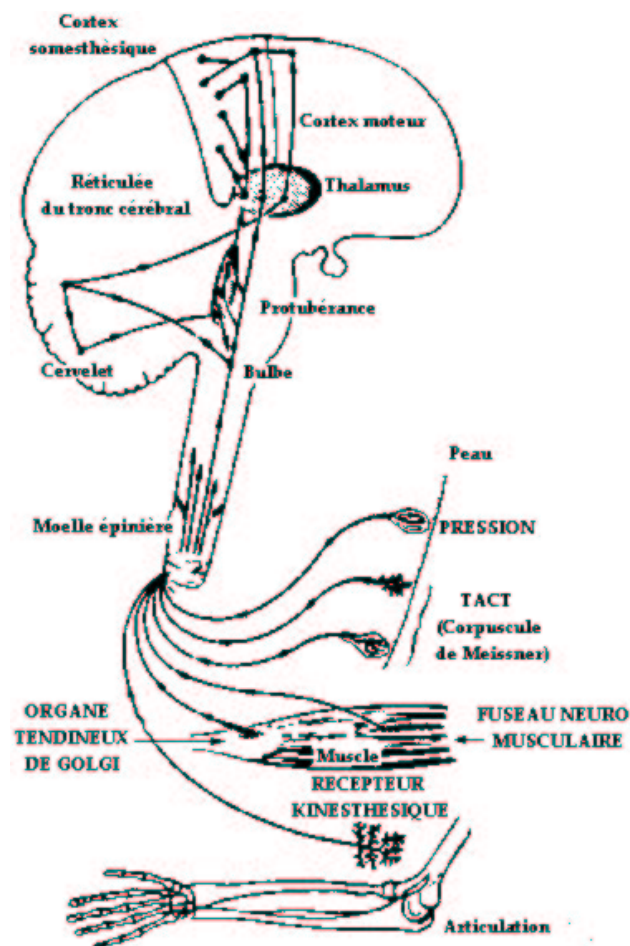


Figure 4.34 : Voies Kinesthésiques (d'après Guyton 1984)

La perception de la verticale est capitale pour se représenter son schéma corporel spatial et par-là même pouvoir évoluer dans le monde qui nous entoure. Il existe trois systèmes de perception de la verticale (visuel, vestibulaire et egocentré). Suivant les individus l'un des trois est prépondérant, mais par le phénomène de vicariance, on peut passer de l'un à l'autre. Dans certains cas, la différence d'informations fournie par ces systèmes peut conduire à des malaises du type mal des simulateurs où le mouvement des images ne s'accompagne pas d'un déplacement du corps. Parfois l'un des systèmes l'emporte sur les autres. C'est l'exemple de la sensation de déplacement ressentie au départ d'un autre train. Ne donner l'illusion du mouvement que par la vue peut entraîner des conflits entre les différents systèmes mis en jeu dans l'environnement réel et même conduire comme nous l'avons dit au mal des simulateurs.

Signalons que l'on peut créer artificiellement des messages sensoriels de mouvement d'un membre inférieur alors que le sujet est au repos et à les yeux fermés, par l'application d'une vibration de 70 Hz au niveau des muscles ischio-jambiers et quadriceps. La Figure 4.35 décrit cet état.

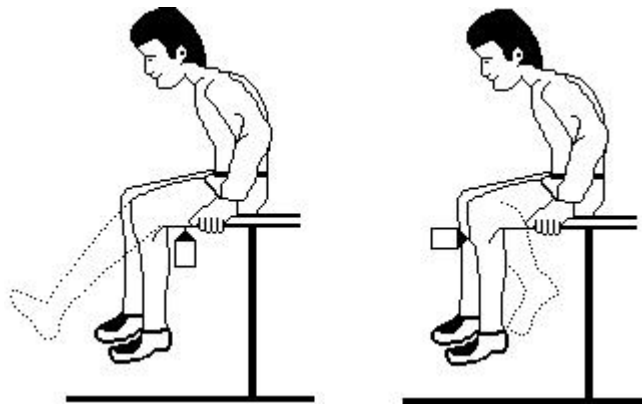


Figure 4.35 : Illusion de mouvement (d'après Roll et al. 1986)

Si les connaissances physiologiques ouvrent des voies pour leurrer les sensations kinesthésiques, les travaux effectués à ce jour, en particulier au niveau des simulateurs ne sont pas entièrement satisfaisants. En effet, si l'on peut simuler une accélération par inclinaison du véhicule, il reste difficile de la maintenir. De même il est très difficile, voire impossible, de simuler les changements de gravité. La création d'un monde virtuel de vol spatial n'est pas encore pour demain.

4.7 RÉPONSES MOTRICES

Les muscles constituent une grande partie du corps (environ 50% en poids). A part des muscles aux fonctions spécifiques (muscles des cordes vocales, muscles du cœur, des yeux, etc.), les autres sont des muscles squelettiques. Ils peuvent fournir un travail mécanique et des mouvements d'une partie du corps par rapport à une autre, permettant à l'homme d'agir physiquement. Ils peuvent aussi fournir des forces importantes sans qu'il y ait mouvement, comme dans la position immobile debout. Les muscles, composés de fibres musculaires, peuvent se contracter mais ils sont aussi élastiques et peuvent s'étirer.

Une réaction motrice involontaire à un stimulus sensoriel (réflexe myotatique) est contrôlée par la moelle épinière. Les fuseaux neuromusculaires dans un muscle transmettent l'information d'étirement du muscle à une synapse dans la moelle épinière. En retour, la synapse active des fibres musculaires et ainsi la contraction du muscle avec un temps de réponse très court. Ces réflexes myotatiques ont un rôle important pour garder une posture immobile pour une personne. Le maintien d'une posture se traduit toujours par une contraction d'un groupe musculaire et une relaxation du groupe musculaire antagoniste. Le fait qu'il y ait constamment des muscles au travail explique que le maintien d'une posture entraîne toujours une fatigue. La meilleure posture est celle qui permet d'en changer ! Par contre, un mouvement volontaire est contrôlé par le cerveau. Ce mouvement nécessite un ensemble de processus qui se suivent. Dans le cas classique de prise volontaire d'un objet, plusieurs étapes interviennent : en premier l'objet est localisé, ensuite le mouvement à réaliser est programmé et enfin le mouvement est exécuté. Cette exécution est réalisée par un mouvement d'approche rapide sans bouclage de correction (sans «feed back»), ensuite par un mouvement plus lent et ajusté grâce aux informations sensorielles (feed back), enfin l'objet est pris avec des ajustements fins des doigts autour de l'objet. A noter que suivant la perception du poids apparent de l'objet, celui-ci est pris avec les doigts de différentes façons. D'une façon générale, les perceptions et l'action sont liées. Une étude détaillée sur le couplage entre perception et action peut être consultée dans [Berthoz, 1998].

4.8 LA MAIN, ORGANE D'INFORMATION ET D'ACTION

4.8.1 ANATOMIE DE LA MAIN

Le système moteur de la main est constitué par une structure comprenant 18 os reliés entre eux par des articulations sans friction et recouverts par des tissus mous et de la peau. Des tendons rattachent ce squelette à des muscles qui permettent de donner à la main plus de 20 degrés de liberté [Srinivasan, 1993].

Le système sensitif possède un grand nombre de récepteurs et de terminaisons nerveuses dans la peau, les articulations, les tendons et les muscles. Il y aurait 17000 mécanorécepteurs dans la paume de la main et dans les bouts des doigts. Quand ces récepteurs sont activés, ils transmettent des impulsions électriques via les neurones afférents au système nerveux central duquel les instructions sont transmises via les neurones efférents aux muscles moteurs. Ceci permet la palpation, c'est-à-dire la recherche d'informations tactiles concernant un objet.

Sur le plan des récepteurs sensoriels : les fibres à terminaisons libres de la peau de la main comprennent quatre groupes de fibres mécanoréceptrices à bas seuil qui contribuent aux sensations tactiles non douloureuses [Barfield e.a., 1995]. Elles diffèrent en fonction de leur rapidité d'adaptation pour répondre à une gamme de fréquences vibratoires et de leur capacité à répondre à une bande passante spécifique. De plus, elles se caractérisent par la largeur du champ tactile qu'elles peuvent appréhender.

Les quatre groupes sont :

- fibres à adaptation rapide avec un champ de réception localisé {fast adapting (FAI)} ;
- fibres à adaptation rapide avec un champ de réception diffus {fast adapting (FAII)} ;

- fibres à adaptation lente avec un champ de réception localisé {slow adapting (SAI)} ;
- fibres à adaptation lente avec un champ de réception diffus {slow adapting (SAII)}.

En combinant avec la perception somatosensorielle en provenance des fibres afférentes, la perception vibrotactile provoque la perception tactile en passant par quatre voies correspondant à quatre canaux. [Bolanowski, 1988] :

Canal pacinien (P)

- actives dans la bande de fréquences vibratoires 40-800 Hz ;
- à un seuil de sensibilité aux fréquences en forme de U avec un maximum de sensibilité vers 300 Hz ;
- grande sensibilité aux changements de température de la peau ;
- grande sensibilité aux changements de forme et de durée ;
- capable de sommation spatiale et temporelle ;
- cette sensibilité concerne les vibrations.

Premier canal non-pacinien (NP I)

- à une réponse relativement plate vis-à-vis des fréquences vibratoires ;
- fonctionne nominalement pour une bande de fréquences vibratoires entre 10 et 100 Hz ;
- n'est pas affecté par les variations de température ;
- ne présente pas de sommation spatiale et temporelle ;
- sa qualité sensorielle est d'être sensible au frémissement ;
- en terme de physiologie, il s'agit de fibres à adaptation rapide ;

Second canal non-pacinien (NP II)

- actives dans la bande de fréquences vibratoires 15-400 Hz ;
- sensible à la température ;
- présente une sommation temporelle ;
- sa capacité à sommer spatialement n'est pas connue ;
- sa spécificité sensitive n'est pas connue ;
- en terme de physiologie, il s'agit de fibres à adaptation lente ;

Troisième canal non-pacinien (NP III)

- actives dans la bande de fréquences vibratoires 0,4-100 Hz ;
- peut être excité par des stimuli vibratoires sur une gamme considérable d'intensités et de fréquences ;
- sensible à la température ;

- ne possède pas de sommation spatiale.

L'information tactile donnée par un contact initial renseigne sur les caractéristiques de surface de l'objet comme la géométrie de surface, sa texture, sa température et son lissage plus ou moins important et par la même son glissement. Ensuite cette sensation est complétée par l'interaction avec les données concernant le retour d'effort (force feedback) qui renseignent sur toute la force d'appui, l'élasticité et le poids. Ces caractéristiques peuvent être interprétées comme forces de résistance. Finalement, c'est le retour kinesthésique ou proprioceptif qui indique à l'individu la position de ses membres et de leur prolongement. Ainsi le toucher et la manipulation implique le corps tout entier.

4.8.2 FONCTION DE LA MAIN

Dans une tâche manuelle la main sert à prendre de manière active de l'information et à agir. La notion de retour tactile et de retour kinesthésique devient donc fondamentale. Par exemple, l'évaluation de la longueur et de la consistance d'un objet entre le pouce et l'index demande des informations d'origine kinesthésique mais aussi tactile [Srinivasan, 1993].

La main permet la prise et la préhension. *«Prendre n'est pas simplement saisir, la préhension est intentionnelle. La préhension est l'ensemble des fonctions mises en jeu dans le but de saisir à l'aide des mains. Elle comporte une intention, un mécanisme de prise et une information sensorielle permanente. La prise est la partie mécanique de la préhension»*. [Tubinar e.a., 1990]. Manier, soulever et soupeser des objets sont des actions courantes que la plupart des individus font avec leurs mains pour avoir une idée précise de leur taille, forme et orientation, sans avoir recours à la vision. Cette interaction active avec le monde physique est désignée par le terme de toucher dynamique par Gibson [Gibson, 1966].

La préhension manuelle comporte plusieurs phases :

- l'approche qui peut se faire selon 3 modalités : à vue, la plus précise, par tâtonnement et enfin sur ordre pour des objets mémorisés ;
- la prise comporte 3 stades : la présentation et l'ouverture des doigts, la fermeture des doigts et enfin la régulation de la force de prise. On distingue les prises de force et les prises de précision qui dépendent non de l'objet mais de ce que l'on se propose de faire. La prise de force peut se faire avec une prise digito-palmaire (exemple : la tenue d'un marteau). La prise de précision se trouve dans la prise pollici-digitale ;
- la régulation de la force de prise est essentielle. La force de serrage doit s'adapter au poids de l'objet, à sa fragilité, à sa surface plus ou moins glissante et à son utilisation. Les prises de force mesurées en Kg force à l'aide d'un dynamomètre sont indiquées dans le Tableau 4.1.

Le lâchage de prise est assuré par l'ouverture des doigts. Il est à noter que cette ouverture est sous la dépendance du nerf radial alors que la fermeture est sous la dépendance du nerf cubital. Il apparaît ainsi bien difficile de parler de la manipulation d'un objet sans parler de l'objet lui-même. C'est ainsi que s'appuyant sur Gibson [Gibson, 1966], Turvey et ses collègues ont analysé les invariants mécaniques qui existent lorsqu'un objet est manipulé pour en connaître la taille, la forme et l'orientation [Turvey e.a.,

	chez l'homme	chez la femme
digito-palmaire	45kg	25kg
pollici-digitale pulpaire	8kg	6kg
prise latérale	7,5kg	5 kg

Tableau 4.1: Prises de force moyenne (mesurées en kg)

1992 ; Turvey e.a., 1989 ; Solomon e.a., 1988 ; Barac-Cikoja e.a., 1991 ; Carello e.a., 1992 ; Kinsella-Shaw e.a., 1992 ; Chan e.a., 1991]. Il apparaît que la force mécanique exercée sur les tissus et sur les muscles soit suffisante pour renseigner sur la forme, la taille et l'orientation. Par contre, pour estimer une longueur d'une barre, ce n'est pas la longueur qui est détectée mais les forces existantes autour du poignet et plus spécifiquement le moment d'inertie autour du poignet. Ce moment d'inertie correspond à la résistance de l'objet qui est tourné dans les trois dimensions. Mathématiquement, le moment d'inertie rend compte de la masse d'un objet et de la distribution de cette masse autour de l'axe de rotation. Ce genre de considération est à prendre dans le développement des interfaces haptiques dans les environnements virtuels, en particulier au niveau des logiciels.

Lederman et Klatzky (1987) ont classé les mouvements de la main par rapport au mode d'exploration et au type d'informations recueillies par ces mouvements [Lederman e.a., 1988 ; Klatzky e.a., 1987 ; Klatzky e.a., 1992 ; Reed e.a., 1990]. Les mouvements de déplacements latéraux où les doigts glissent sur la surface d'un objet induisent une sensation de texture par frottement de la peau sur l'objet. Une pression appuyée sur un objet grâce au retour de force et au couple de celle-ci permet de déterminer la compliance (faculté de se déformer plus ou moins à la pression) de l'objet. Cela renseigne sur la dureté ou la mollesse de l'objet. Le contact statique de large surface de peau renseigne sur les contours comme si la peau formait un moule autour de la forme de l'objet. Il fournit aussi l'information de température de l'objet. La tenue non appuyée et le levage d'un objet sont ressentis comme un poids. Le serrage d'un objet, et par conséquent le modelage de cet objet ou d'une partie de celui-ci avec la main, permet d'apprécier la forme de ce dernier. Et si le serrage de l'objet est total, il devient aussi possible de déterminer les dimensions de cet objet. Alternativement il est possible de déterminer la forme de l'objet par le traçage de ces bords (poursuite de contour).

Deux de ces procédures peuvent être mises en jeux simultanément. C'est le cas de l'exploration « en tandem » décrite par Reed, Lederman et Klatzky. Par exemple les mouvements latéraux de glissade peuvent être combinés avec des pressions locales et ainsi donner simultanément des informations de texture et de dureté. Une procédure peut ainsi donner deux informations primitives. De même l'enfermement peut donner une grossière information de forme mais aussi d'autres renseignements comme par exemple la température. Pourtant, il doit être signalé que quelquefois la procédure en tandem devient un travail contre-productif. Par exemple l'utilisation de la pression pour détecter la dureté distord la forme de l'objet [Reed e.a., 1990].

Ces procédures peuvent être considérées comme suffisantes et nécessaires pour apprécier les propriétés d'un objet [Klatzky e.a., 1992]. Par exemple dans l'identification de la texture, on doit frotter la surface. Ainsi le mouvement latéral est nécessaire. La poursuite de contours est une méthode très suffisante pour déterminer la forme d'un objet. De plus le suivi de contour est plus minutieux que le serrage pour identifier une forme. Ainsi la poursuite de contours devient nécessaire pour apprécier la forme exacte. Pour-

tant, ce mouvement prend du temps pour être accompli et beaucoup de gens ne l'essayent pas en premier lieu. Ensermer et maintenir un objet sont généralement suffisant pour faire une discrimination grossière des propriétés de cet objet. De plus un maintien non appuyé peut donner des informations sur les dimensions des surfaces et leur forme sans perdre du temps en employant une procédure de suivi des contours. Enfin, l'identification des objets est facilitée par l'utilisation du maximum de doigts. L'utilisation de plusieurs doigts permettent une plus grande exploration spatiale.

4.9 CLASSIFICATION DES INTERFACES

Les interfaces comportementales de la réalité virtuelle peuvent être classées en interfaces sensorielles ou interfaces motrices, sauf le cas des interfaces sensori-motrices. Nous proposons la classification suivante, qui découle évidemment des différents sens et réponses motrices de l'homme. Et nous avons tenu compte aussi que plusieurs types d'interfaces sont exploitées par la main, d'où le regroupement des interfaces **manuelles**. C'est suivant cette classification que sont présentées les interfaces comportementales dans la suite des chapitres :

LES INTERFACES MOTRICES :

LES CAPTEURS DE LOCALISATION - TRAQUEURS
LES INTERFACES SPECIFIQUES DE LOCALISATION CORPORELLE
LES INTERFACES **MANUELLES** MOTRICES

LES INTERFACES SENSORI-MOTRICES :

LES INTERFACES **MANUELLES** SENSORI-MOTRICES
(les interfaces à retour d'effort)

LES INTERFACES SENSORIELLES :

LES INTERFACES **MANUELLES** SENSORIELLES
(les interfaces à sensibilité cutanée)
LES INTERFACES VISUELLES
LES INTERFACES A SIMULATION DE MOUVEMENT ET LES INTERFACES A
SIMULATION DE CLIMAT
LES INTERFACES OLFACTIVES

4.10 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Anzieu, 1985] D. Anzieu, «*Le Moi-peau*». Dunond, Paris(1985).
- [Barac-Cikoja e.a., 1991] D. Barac-Cikoja, & M.T. Turvey, «*Perceiving aperture size by striking*». Journal of Experimental Psychology : Human Perception and Performance, 17(2),pp. 330-346. (1991).
- [Barfield e.a., 1995] W. Barfield, C. Hendrix, O. Bjorneseth, K.A Kaczmarek, and W. Lottens, «*Comparison of human sensory capabilities with technical specifications of virtual environment equipment*». Presence, 4(4), 329-356. (1995).

- [Bolanowski e.a., 88] S.J.Bolanowski, G.A. Gescheider, R.T. Verillo, and C.M. Checkosky, «Four channels mediate the mechanical aspects of touch». *Journal of the Acoustical Society of America*, 84, 1680-1694. (1988).
- [Bonnard e.a., 1989] M. Bonnard, C. Bonnet, M-C. Botte, J. Pailhous, J.B. Segui, «Perception, action, langage : Traité de psychologie cognitive 1». Dunod Bordas, Paris, (1989).
- [Bullier, 1998] J. Bullier, « *Architecture fonctionnelle du système visuel, Vision : aspects perceptifs et cognitifs* », pages 11–41. Solal, Marseille (France), Boucart, M., Hénaff, M. A., et Belin, C., éditeurs, (1998).
- [Burdea, 1996] G. C. Burdea, «Force and touch rétroaction for virtual reality». New York : John Wiley & Sons. (1996).
- [Burniston, 1996] A. Burniston, «The human side of life : The analysis of training requirements to aid training equipment design». *Military Simulation & Training*, March, 42-48. (1996).
- [Buser e.a., 1987] Buser et Imbert, « neurophysiologie fonctionnelle IV », éditeur Hermann, Paris, 1987.
- [Campbell e.a., 1968] F. W. Campbell et J. Robson, « Application of Fourier analysis to the visibility of gratings ». *Journal of Physiology (London)*, 197 :551—566, (1968).
- [Carello e.a., 1992] C Carello, P. Fitzpatrick & M.T Turvey, «Haptic probing : Perceiving the length of a probe and the distance of a surface probed». *Perception and Psychophysics*, 51(6), 580-598. (1992).
- [Chan e.a., 1991] T.C. Chan & M.T. Turvey, «Perceiving the vertical distances of surfaces by means of a hand-held probe». *Journal of Experimental Psychology : Human Perception and Performance*, 17(2), 347-358. (1991).
- [DeValois e.a., 1988] R. L. DeValois, et K. K. DeValois, « Spatial vision ». Oxford University Press, New York, (1988).
- [Enroth-Cugel e.a., 1966] C. Enroth-Cugel et J.D. Robson, « The contrast sensitivity of retinal ganglion cells of the cat ». *Journal of Physiology*, 187 :517—522, (1966).
- [Ehrlich e.a., 1997] J. Ehrlich, B. Knerr, D.Lampton, & D. McDonald, «Team situational awareness training in virtual environments». U. S. Army Research Institute Technical Report (1997).
- [Ehrlich e.a., 1999] J. Ehrlich, E.M. Rodriguez, S. L. Goldberg ,«The Haptic Interface in Virtual Reality» U. S. Army Research Institute Simulator Systems Research Unit Orlando, Florida For NATO Research Study Group 28 Human Factors Issues in the Military Use of Virtual Reality Technology. (1999).
- [Fuchs e.a., 1995] P. Fuchs, D. Ernadotte, D. Maman, C.Laurgeau, J.P. Bordas, « Téléprésence visuelle stéréoscopique », congrès «Interface des mondes réels et virtuels», Montpellier 95, p77-p91.
- [Gibson, 1996] J.J. Gibson, «The Senses Considered as Perceptual Systems». New York : Houghton Mifflin. (1996).
- [Hollands, 1996] R. Hollands, «The virtual reality homebrewer's handbook». Chichester : John Wiley & Sons. (1966).
- [Holley e.a., 1994] A.Holley et G. Sicard, «Les récepteurs olfactifs et le codage neuronal de l'odeur». *Médecine/sciences* ; 10 1091-8 (1994).
- [Iggo, 1973] A. (Ed.) Iggo, «Handbook of sensory physiology : Somatosensory system», vol.2 New York : SpringerVerlag. (1973).
- [Julesz, 1971] B. Julesz, « Foundation of Cyclopean Perception », university of Chicago press, Chicago, (1971).

- [Kinsella-Shaw e.a., 1992] J.M. Kinsella-Shaw & M.T. Turvey, «Haptic perception of object distance in a single-strand vibratory web». *Perception and Psychophysics*, 52(6), 625-638. (1992)
- [Klatzky e.a., 1992] R.L. Klatzky & S.J. Lederman, «Stages of manual exploration in haptic object identification». *Perception and Psychophysics*. 52(6), 661-670. (1992).
- [Klatzky e.a., 1987] R.L. Klatzky, S.J. Lederman & C. Reed, «There's more to touch than meets the eye» *Journal of Experimental Psychology : General*, 116(4), 356-369. (1987).
- [Latham, 1996] R. Latham, «Now that VR is officially dead, perhaps we can get on with doing it». *Real Time Graphics*, 4(9), 4. (1996).
- [Lederman e.a., 1988] S.J. Lederman, R.A. Browse, & R.L. Klatzky, «Haptic processing of spatially distributed information». *Perception and Psychophysics*, 44(3), 222-232. (1988).
- [Lederman e.a., 1987] S.J. Lederman & R.L. Klatzky, «Hand movements : A window into haptic object recognition». *Cognitive Psychology*, 19, 342-368. (1987)
- [Livingstone e.a., 1988] M. Livingstone et D. Hubel, « Segregation of form, color, movement and depth : anatomy, physiology and perception ». *Science*, 240 :740—749, (1988).
- [Marbach, 1972] Marbach, «Cours d'ergonomie des environnements», Strasbourg. (1972)
- [Marr e.a., 1980] D. Marr, T. Poggio, et E. Hildreth, « The smallest channel in early human vision ». *Journal of the Optical Society of America*, 70 :868—870, (1980).
- [Marr, 1982] D. Marr, « Vision : a computational investigation into the human representation and processing of visual information ». Freeman, New York, (1982).
- [Mery, 1971] J. Mery, «Le bruit, aspects physiques et physiologiques». *Cours d'ergonomie*, Strasbourg, (1971).
- [Papin, 1992] J.P. Papin, « cours d'ergonomie », Angers, ETAS, 1992.
- [Perrin, 1998] J. Perrin, « Profondeur et binocularité : algorithmie, étude psychophysique et intérêt pour l'ergonomie des interfaces stéréoscopiques ». Thèse de doctorat de l'Ecole des Mines de Paris, 1998.
- [Piéron, 1964] H. Piéron, «La sensation». QSJ n° 555 PUF, Paris 4ème édition, (1964).
- [Plantier e.a., 1990] J. Plantier, J.P. Menu, C. Roumes, « caractérisation optique d'un système de génération d'images stéréoscopiques », rapport Cerma 90-48, 1990.
- [Reed e.a., 1990] C.L Reed, S.J. Lederman, & R.L. Klatzky, «Haptic integration of planar size with hardness, texture, and planar contour». *Canadian Journal of Psychology*, 44(4), 522-545. (1990)
- [Rodieck e.a., 1965] R.W. Rodieck, et J. Stone, « Analysis of receptive fields of cat retinal ganglion cells ». *Journal of Neurophysiology*, 28 :833—849, (1965).
- [Richard, 2000] P. Richard., «Le Net et la télé ont aussi une odeur!». Ouest France, multimédia du 8 octobre (2000).
- [Sanders e.a., 1993] M. S. Sanders & E.J. McCormick, «Human factors in engineering and design». New York : McGraw-Hill. (1993)
- [Schneider, 1971] O. Schneider, «Physique du bruit» *Cours d'ergonomie*, Strasbourg 1971.
- [Sherrick e.a., 1986] C.E. Sherrick, and R.G. Cholewiak, «Sensory processes and perception», in *Handbook of perception and human performance*, K.R. Boff, L. Kaufman, and J.P. Thomas (Eds.) Vol. 1, Ch. 12. New York : Wiley & Sons. (1986)

- [Solomon e.a., 1988] H.Y. Solomon & M.T. Turvey, «Haptically perceiving the distances reachable with hand held objects». *Journal of Experimental Psychology : Human Perception and Performance*, 14(3), 404-427. (1988).
- [Srinivasan, 1993] M.A.Srinivasan, «The human haptic system». MIT : Department of Mechanical Engineering. (1993).
- [Stuart, 1996] R.Stuart, «The design of virtual environments». New York : McGraw Hill. (1996).
- [Trotter, 1995] Y. Trotter, « Bases neuronales de la perception visuelle chez le primate ». *Journal Français d'Orthoptique*, 27, 1995.
- [Tubinar e.a., 1990] R. Tubinar R. et J-M. Thomine, «La main anatomie fonctionnelle et l'examen clinique». Masson, Paris. (1990).
- [Turvey e.a., 1992] Turvey, M.T., G. Burton, , C.C. Pagano, H.Y. Solomon, & S. Runeson, «Role of inertia tensor in perceiving object orientation by dynamic touch». *Journal of Experimental Psychology : Human Perception and Performance*, 18(3), 714-727. (1992).
- [Turvey e.a., 1989] M.T.Turvey, H.Y. Solomon, & G. Burton, «An ecological analysis of knowing by wielding». *Journal of the Experimental Analysis of Behavior* 52(3), 387-407. (1989).
- [Tyler, 1991] C. W. Tyler, « The horopter and binocular fusion, Binocular vision, volume 9 de Vision and visual dysfunction », pages 19–37. Macmillan Press, London, Regan, D., éditeur, (1991).
- [Valyus, 1962] N.A. Valyus, « Stereoscapy », (London, New York), Focal Press (1962).
- [Wable e.a., 1996] J. Wable et S. Chery-Croze, «Presbyacousie et appareillage auditif». Lyon, GIRA éditions (1996).
- [Watson e.a., 1982] A. Watson, H. Barlow, et J. Robson, « What do the eye see the best ? » *Nature*, 302(5907) :419—422, (1983).
- [Wilson e.a., 1979] H. R. Wilson, et J. R. Bergen « A four mechanism model for spatial vision. *Vision Research* », 19 :19—32, (1979).
- [Wilson e.a., 1983] H. R.Wilson, D. McFarlane et G. Phillips, « Spatial frequency tuning of orientation selective units estimated by oblique masking ». *Vision Research*, 23(9) :873—882, (1983).
- [Wong-Riley e.a., 1979] M. Wong–Riley, « Changes in the visual system of monocularly sutured or enucleated cats demonstrated with cytochrome oxidase histochemistry ». *Brain Research*, 171 :11—18, (1979).
- [Yei-Yu e.a., 1990] Y. Yei-Yu, L. Silverstein, « Limits of fusion and depth judgment in stereoscopic color displays », *the human factors society*, 21(1), 45-60, (1990).
- [Zeki e.a., 1978] S. M. Zeki, « Fonctionnal specialisation in the visual cortex of the rhesus monkey ». *Nature*, 274(5670) :423—428, (1978).

Adresse Web de centres de recherche :

[http1] <http://www.iurc.montp.inserm.fr/cric/audition>

Troisième partie

Les interfaces motrices

5 LES CAPTEURS DE LOCALISATION

Philippe Fuchs

5.1 GÉNÉRALITÉS

5.1.1 LA LOCALISATION SPATIALE

L'objectif de ce chapitre est de présenter les capteurs permettant une **localisation en temps réel** de tout objet pour le positionnement et(ou) le suivi de sa trajectoire. Ce sont des interfaces motrices indispensables pour les techniques de réalité virtuelle. Dans ce domaine, l'objet à localiser est souvent une partie du corps, tout le corps ou un objet manipulé par l'utilisateur. Ces dispositifs, interfaces motrices pour la réalité virtuelle, peuvent être aussi exploités dans d'autres secteurs d'activité (métrologie, robotique, etc.). Quand ils sont employés pour la réalité virtuelle, on les nomme par le néologisme «traqueur». Ces capteurs sont conçus pour mesurer seulement la position spatiale d'un repère mobile, lié à l'objet observé, par rapport à un repère fixe. Dans le chapitre suivant, nous présentons les interfaces spécifiques de localisation du corps ou d'une partie de celui-ci, telles que les combinaisons de données et les gants de données. Ces interfaces spécifiques de localisation peuvent exploiter dans leur fonctionnement des traqueurs.

Le positionnement d'un repère lié à un objet est parfaitement déterminé si le traqueur permet de mesurer les 6 paramètres (3 angles et 3 distances) liés aux 6 degrés de liberté (ou ddl) du repère. Ces 6 paramètres sont définis par la matrice de transformation géométrique entre le repère lié à l'environnement et le repère lié à l'objet. Cette transformation est composée d'une rotation et d'une translation. Les paramètres peuvent être définis de différentes façons, et en général les systèmes de mesure laissent à l'utilisateur le choix entre les différentes représentations. Le capteur de positionnement est soit lié à l'objet ou soit immobile par rapport à l'environnement. Dans ce dernier cas, c'est par l'observation de l'objet que le capteur détermine les paramètres de localisation. Le traqueur parfait (ou «presque parfait» en ne tenant pas compte du prix) n'existe pas. Tous les traqueurs présentés dans ce livre ont au moins une contrainte d'exploitation ou des performances restrictives. Les techniques de réalité virtuelle sont pourtant très demandeuses de ce type d'interface. Mais concevoir et réaliser un capteur de localisation bon marché qui fournisse en temps réel (à une fréquence supérieure à 100 Hz) six mesures très précises, sans contraintes d'exploitation, est un challenge technique très difficile. Il peut être étonnant qu'à notre époque aux techniques très évoluées, donner six mesures en temps réel semble très complexe. Depuis la première version de mon livre [Fuchs, 1996], il n'y a pas eu d'avancée importante pour les traqueurs, malgré ce que nous avions espéré. Et même, en ce qui concerne les principaux traqueurs employés en réalité virtuelle, les traqueurs électromagnétiques, il n'y a eu ni amélioration importante de leurs caractéristiques métrologiques, ni baisse de leur prix.

C'est sous l'essor de la réalité virtuelle que sont apparus les capteurs de localisation les plus évolués. Ces traqueurs sont souvent employés dans les systèmes de réalité virtuelle pour :

- la localisation de la tête de l'utilisateur portant un visiocasque ;
- la localisation (6 degrés de liberté) de la main avec les gants de données ;
- la mesure du déplacement spatial de l'extrémité de la main permettant de manipuler un objet virtuel ou de diriger *naturellement* un curseur dans un environnement virtuel ;
- la localisation d'une partie du corps, utilisée entre autres, pour la «motion capture» (localisation de tout le corps), que nous présenterons au chapitre suivant : les interfaces spécifiques de localisation corporelle.

Les véritables capteurs de localisation spatiale déterminent normalement les 6 paramètres liés aux 6 degrés de liberté. Par contre certains capteurs ne mesurent que 3 paramètres :

- les 3 coordonnées liées à la translation de l'objet par rapport au référentiel ;
- les 3 coordonnées liées à la rotation de l'objet ;
- les 3 coordonnées liées à la localisation dans un plan.

Ce livre présente les capteurs de localisation *spatiale*. Nous ne présentons pas les simples capteurs de mesure de distance ou d'orientation angulaire bien qu'à partir d'un ensemble de ce type de matériel, il soit toujours possible de déterminer la position spatiale d'un objet en mesurant plusieurs de ses points par rapport à un référentiel (méthode de triangulation). De même nous n'étudions pas les capteurs de «navigation» permettant la localisation terrestre, exploités, entre autres, en robotique mobile et en aéronautique. Par exemple, le système GPS de navigation utilise les mesures de distances entre une balise et des satellites géostationnaires. Nous ne présentons donc que les capteurs localisant un objet dans une zone restreinte, telle que l'espace environnant d'un ordinateur ou d'une pièce (quelques centimètres à quelques décimètres). Nous ne présentons pas non plus dans ce chapitre les interfaces manuelles sensori-motrices à retour d'effort, qui toutes obligatoirement mesurent la position de la main en même temps qu'elles appliquent des efforts en retour sur la main (voir le chapitre sur ces interfaces). Pour ce type d'interfaces à retour d'effort, les mesures de position sont en général basées sur le même principe que celui des traqueurs mécaniques mesurant des distances, présentés en 5.3.1.

5.1.2 CAPTEUR DE LOCALISATION ET INTERFACE DE COMMANDE

Nous venons d'indiquer qu'un traqueur peut servir à manipuler un objet virtuel ou à diriger un curseur dans un environnement virtuel. Dans le premier cas, l'utilisateur agit «*naturellement*» sur le monde virtuel, le traqueur se contentant de suivre le mouvement de sa main. Dans le deuxième cas, il s'agit en général de commander une action spécifique : par exemple, se déplacer dans le monde virtuel en indiquant une direction, pointer sur un menu déroulant, commander symboliquement l'ouverture d'une porte, etc. Dans ce deuxième cas, les traqueurs peuvent être remplacés par *des interfaces de commande* classiques (joysticks et souris 3D). Celles-ci, posées en général sur une table, permettent à l'utilisateur de commander manuellement le déplacement d'un curseur grâce aux mouvements relatifs de l'interface par rapport à une position neutre. Ces mouvements ont de faibles courses (quelques millimètres à quelques centimètres). Il ne s'agit donc pas de localisation dans l'espace de la main pour ces interfaces de

commande et elles sont basées évidemment sur d'autres principes que les traqueurs. A noter qu'il s'agit dans le cas d'un traqueur de mouvoir *naturellement* un curseur, donc sans grand effort intellectuel. A la différence des souris 3D et joysticks avec lesquels le curseur est mû grâce à une action manuelle, réfléchie et peu naturelle de l'utilisateur, sauf pour les habitués de ce type d'interface. Si un traqueur peut donc être utilisé en curseur 3D, les joysticks et les souris 3D ne peuvent être employés en capteur de localisation de la main de l'utilisateur. Ces derniers seront présentés dans le chapitre des interfaces spécifiquement manuelles : interfaces (motrices) de commande.

5.2 PRINCIPES DE POSITIONNEMENT

Avant de présenter les phénomènes physiques qui régissent les différents types de traqueurs, nous explicitons les principes géométriques de positionnement qu'ils exploitent.

5.2.1 PRINCIPE PAR MESURES DE DISTANCE ENTRE POINTS

Comme nous l'avons signalé au paragraphe précédent, la mesure de localisation d'un solide indéformable (6 ddl) peut être déterminée au minimum à partir des positions de trois points S_1 , S_2 et S_3 non alignés du solide. La position du point S_1 est calculable par triangulation grâce à trois distances mesurées entre ce point S_1 et trois points fixes F_1 , F_2 et F_3 liés au repère de référence. La position du point S_2 peut être déterminée par deux distances au minimum avec deux points F_i , car la distance entre S_1 et S_2 est connue, et de même pour S_3 une seule distance reste à déterminer avec un point F_i . Dans ce cas minimum de six mesures, on peut calculer les 6 ddl du solide. Mais concrètement, si nous pouvons mesurer trois distances pour le point S_1 , nous pouvons le faire aussi facilement matériellement avec les deux autres. Nous obtenons alors un système de neuf équations à six inconnues, dont nous établissons la solution par minimisation des erreurs. De plus la mesure concrète des distances peut introduire parfois une inconnue complémentaire, (la vitesse du son dans l'air peut être considérée comme une variable complémentaire dans les méthodes à ultrasons). Nous avons dans ce cas un système de neuf équations à sept inconnues (Figure 5.1).

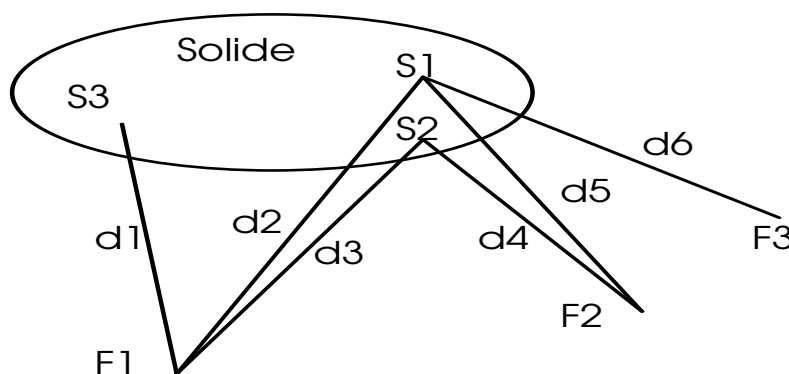


Figure 5.1 : Principe basé sur les mesures de distances entre points

En conclusion pour ce principe, il s'agit de pouvoir mesurer physiquement des distances entre points. Comme nous l'avons écrit précédemment, toute personne peut mettre en œuvre elle-même ce dispositif avec de simples capteurs de mesure de distance. Mais des constructeurs proposent de tels capteurs de localisation bien optimisés, il est donc vain d'en concevoir d'autres, sauf pour des applications bien spécifiques. Ce principe est employé dans les traqueurs à ultrasons, les traqueurs optiques et les traqueurs mécaniques à câbles.

5.2.2 PRINCIPE DE MESURE AVEC SOURCE ÉMETTRICE

Une source émettrice fixe crée un champ variable d'une grandeur physique dans son espace environnant. Un récepteur lié à l'objet à localiser mesure ce champ, dont les valeurs sont fonction de la position spatiale de l'objet (Figure 5.2). Ce principe est utilisé dans les traqueurs électromagnétiques.

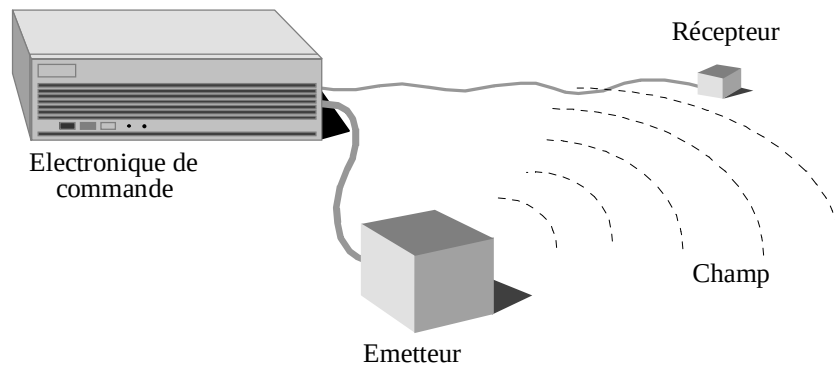


Figure 5.2 : Principe basé sur les mesures de champ

Remarque : le récepteur peut théoriquement être fixe et la source liée à l'objet observé. Mais pour des raisons de poids et d'encombrement, la source est généralement fixe.

5.2.3 PRINCIPE DE MESURE SANS SOURCE ÉMETTRICE ARTIFICIELLE

Il n'y a pas de source émettrice artificielle créant un champ, ce qui facilite la technique de mesure. On peut exploiter les champs gravitationnel ou magnétique du globe terrestre, qui est ici la source émettrice naturelle. Dans ce cas, on détermine les modifications d'orientation par rapport aux lignes du champ, permettant de retrouver l'orientation du capteur. On peut aussi exploiter le principe du gyroscope : l'immobilité dans l'espace de l'axe de rotation d'une masse tournant à grande vitesse.

5.2.4 PRINCIPE DE MESURE DE CARACTERISTIQUES DU MOUVEMENT

Si on connaît en temps réel les vitesses de rotation d'un objet, on peut déterminer son orientation à chaque instant à partir de son orientation initiale (gyromètre). Si on mesure en temps réel les accélérations d'un objet en translation, il suffit de connaître

sa position et sa vitesse initiales pour connaître sa position tridimensionnelle (3ddl) à chaque instant (accéléromètre). Par intégration simple (gyromètre) ou double (accéléromètre), on obtient la position mais les erreurs sont amplifiées par ces intégrations.

5.2.5 PRINCIPE PAR MESURES D'ANGLES ENTRE SOLIDES

Cette méthode est évidemment exploitée pour les traqueurs mécaniques avec lesquels il existe en permanence une liaison entre l'objet détecté et son environnement. Pour laisser 6 ddl à l'objet (ou 3 ddl à un point), la liaison est réalisée par un ensemble de barres ayant des liaisons pivot entre elles. Il suffit de mesurer six angles (ou trois) pour déterminer la position de l'objet. Certains capteurs optiques sont aussi basés sur des mesures d'angles.

Les principes physiques exploités actuellement sont variés, chacun ayant au moins un avantage par rapport aux autres. Les phénomènes physiques exploités sont :

- la mécanique ;
- l'électromagnétisme ;
- l'acoustique ;
- l'optique.

5.3 TRAQUEURS MÉCANIQUES

5.3.1 TRAQUEURS MÉCANIQUES MESURANT DES DISTANCES

Avec ce type de traqueur, il existe en permanence une liaison mécanique entre l'objet détecté et son environnement tout en laissant 6 ddl à l'objet ou 3 ddl à un point. En général, la liaison mécanique est composée d'une chaîne de trois barres articulées à l'extrémité de laquelle se trouve une rotule permettant une orientation angulaire à 3 ddl (Figure 5.3). L'architecture mécanique est souvent identique à celle de robots manufacturiers, mais en plus légère pour supprimer le plus possible les contraintes de mouvement et d'inertie de l'objet observé.

Les mouvements relatifs entre les pièces mécaniques sont des rotations. Celles-ci sont mesurées grâce à des potentiomètres ou des capteurs optiques incrémentaux. Ces mesures aisées procurent les principaux avantages de ce type d'appareil. Les mesures sont précises, très rapides (jusqu'à 300 mesures par seconde) et le temps de réponse est très court (2 à 3 ms). Le prix relativement faible est un autre avantage. Par contre, le seul inconvénient majeur, inhérent à ce principe, est la contrainte de liberté de mouvement : l'étendue de mesure est limitée et le capteur crée des obstacles aux déplacements de l'objet observé. Ce type de traqueur est principalement exploité pour mesurer le mouvement de la tête, mais il peut enregistrer des mouvements d'objet quelconque. Il y a peu de développement sur ce type de traqueur, qui sont peu exploités, malgré leur qualité métrologique intrinsèque. Mais ils font si peu «high tech» qu'on hésite à les utiliser (ou à en concevoir pour des applications spécifiques). Si, après analyse des I² mentales et sensori-motrices, le concepteur déduit que la tête de l'utilisateur se déplace peu, ces traqueurs mécaniques sont à considérer attentivement comme solution, car ils ont les

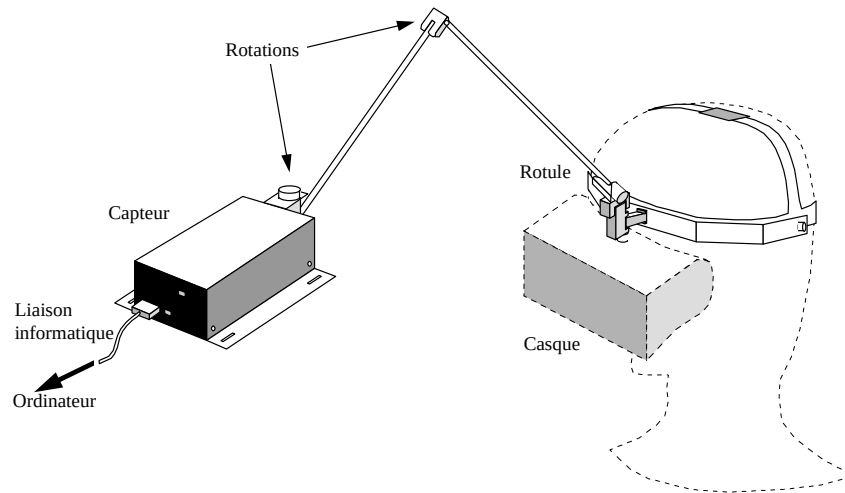


Figure 5.3 : Traqueur de localisation mécanique à leviers articulés

meilleures caractéristiques métrologiques (et tant pis, si le côté mécanique «démodé» apparaît).

Toujours basé sur des mesures mécaniques de distances, différents traqueurs mécaniques ont été conçus dans le passé. La présentation de solutions variées qui va suivre est plutôt écrite pour montrer au lecteur que des solutions originales peuvent être réalisées, que pour en faire un inventaire exhaustif.

Un principe mécanique original a été envisagé pour détecter le déplacement de la tête d'une personne portant un visiocasque. Le dessus du visiocasque est relié par quatre câbles aux quatre coins d'une salle. Les câbles sont toujours légèrement tendus par des moteurs électriques. En mesurant leurs rotations, on peut connaître les distances entre le sommet du visiocasque et les quatre coins de la salle. La rotation autour de l'axe vertical du visiocasque est toujours possible car les quatre câbles sont fixés sur une petite plaque pouvant tourner relativement au visiocasque. Cette rotation est aussi mesurée. Avec ces cinq mesures, on peut calculer les déplacements dans la salle et la rotation verticale. Le principal inconvénient du dispositif est l'obligation d'équiper la salle et d'étalonner le dispositif. Celui-ci n'est pas commercialisé [PCRVR, 1994].

Un prototype similaire (Spidar II) à plusieurs fils est employé par Hirata [Sato, 2001], pour mesurer les mouvements de translation de l'extrémité d'un doigt (3 ddl). Celle-ci est attachée à quatre fils reliés à quatre moteurs-encodeurs qui maintiennent une tension sur les fils. Le doigt peut se mouvoir dans un volume d'un mètre cube environ (Figure 5.4). Grâce aux encodeurs, on relève à chaque instant la longueur des parties déroulées des fils et on détermine géométriquement la position du doigt. Les moteurs peuvent appliquer une tension variable aux fils, ce qui donne aussi en complément une interface originale à retour d'effort (Figure 5.5). Le dispositif peut être utilisé avec deux doigts, en doublant les fils et les moteurs-encodeurs. Mais dans ce cas, il y a évidemment des incompatibilités de mouvement entre les deux doigts.

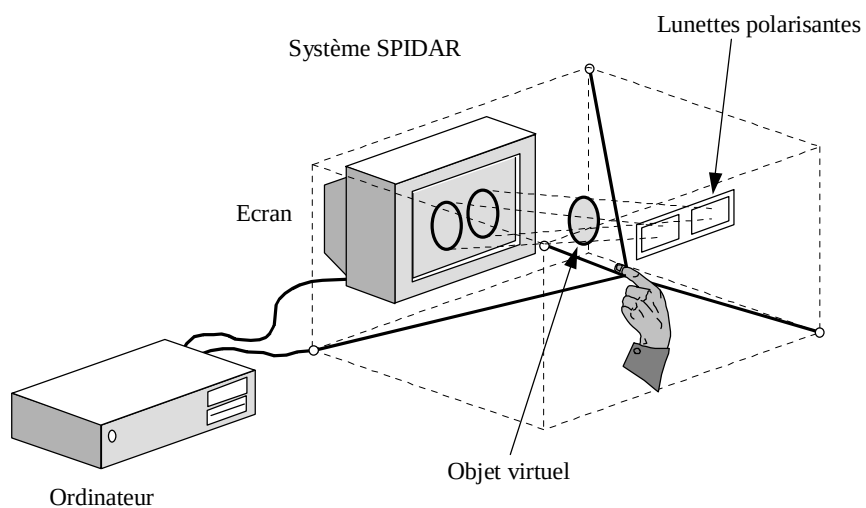


Figure 5.4 : Mesure de position d'un doigt et retour d'effort sur celui-ci

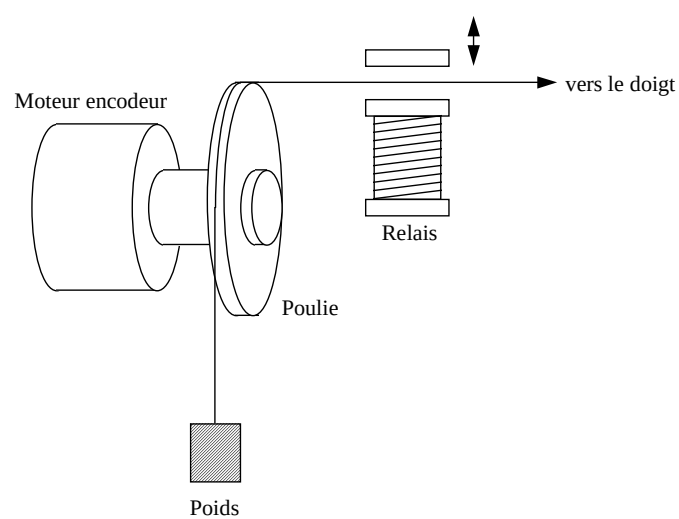


Figure 5.5 : Mécanisme de mesure et de tension d'un fil par moteur-encodeur

Avec cette interface, on peut faire de la «poterie virtuelle» en déplaçant le doigt pour façonner le pot virtuel en rotation. Il est possible aussi de saisir un objet en sentant les efforts sur deux doigts.

5.3.2 CAPTEURS MÉCANIQUES DÉTERMINANT UNE ORIENTATION, UNE VITESSE OU UNE ACCÉLÉRATION

Plusieurs types de traqueurs mécaniques sont regroupés sous cette même partie car ils sont basés sur des principes mécaniques ne fournissant, directement ou indirectement, la mesure *que d'un degré de liberté*. Ils font donc les associer entre eux ou avec d'autres traqueurs «non-mécaniques» pour réaliser des capteurs de localisation à 3 ou 6 ddl. Certains de ces traqueurs ont été développés pour les systèmes de guidage des avions et missiles. Sur des principes similaires, on a conçu des traqueurs pour le marché de la réalité virtuelle en les miniaturisant avec un prix de revient abordable. Certains commencent à apparaître sur des visiocasques à faible prix. Ils sont rarement vendus séparément, mais ils sont associés pour donner des traqueurs à 3 ou 6 ddl. Ils ont l'avantage d'avoir moins de limitation dans l'étendue de mesure, à l'inverse de la plupart des autres traqueurs. Les différents types de capteurs mécaniques exploités pour la réalité virtuelle sont les inclinomètres, les gyroscopes et gyromètres et les accéléromètres.

5.3.2.1 Inclinomètres

Pour ce type de traqueur, c'est la terre qui est la source émettrice du champ gravitationnel. Sans mouvement de l'inclinomètre, il indique sa position angulaire avec la verticale du lieu. Mais s'il y a mouvement, il indique la direction de la force extérieure totale sur lui, due à l'accélération du mouvement et à l'attraction terrestre. Un constructeur propose un inclinomètre basé sur le principe d'un contenant parallélépipédique rempli de liquide visqueux. Sur une face, il y a émission de lumière qui est recueillie sur la face opposée, en quantité fonction de l'angle d'inclinaison (Figure 5.6).

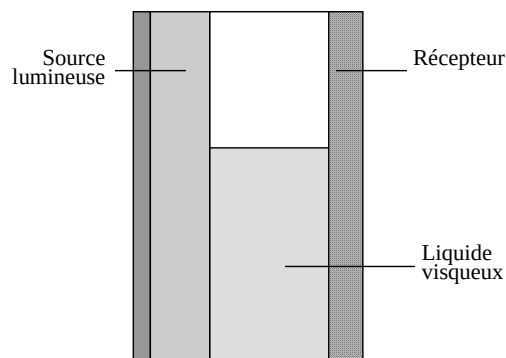


Figure 5.6 : Principe d'un inclinomètre

5.3.2.2 Gyroscopes et gyromètres

Le plus commun des gyroscopes est celui basé sur le principe mécanique d'un rotor tournant à grande vitesse, dont l'axe garde une direction constante. Mais d'autres phénomènes physiques sont exploités pour mesurer des vitesses angulaires de rotation. Nous ne présentons à la suite que les appareils susceptibles d'être miniaturisés et d'un coût acceptable pour être utilisés en réalité virtuelle.

Le gyroscope à suspension de Cardan est un dispositif indiquant une direction constante grâce à une masse (toupie) en rotation autour de son axe de révolution. La toupie du gyroscope est reliée au niveau de son centre d'inertie à une suspension de cardan qui permet à la toupie de s'orienter dans toutes les directions. Les frottements au niveau des liaisons sont réduits au minimum. On prévoit souvent un moteur entraînant la toupie pour corriger la dérive. Le principe de fonctionnement repose sur la propriété suivante (l'approximation gyroscopique) : étant donnée une masse de révolution d'axe Δ ayant une vitesse angulaire relativement grande autour de Δ , on admet que le moment cinétique de cette masse reste parallèle à Δ . Par application du théorème du moment cinétique, on trouve que la direction de Δ est fixe. En théorie, un gyroscope pointe donc toujours dans la même direction. En pratique, cela est difficile à réaliser à cause des frottements au niveau de la suspension. Beaucoup d'appareils ne mesurent pas l'angle de rotation du support par rapport à l'axe Δ , mais sa vitesse angulaire de rotation : le gyromètre. La mesure de localisation de l'objet en mouvement est obtenue à partir de la vitesse par intégration, d'où une augmentation des erreurs de position dans le temps. Des constructeurs ont réussi à miniaturiser ces dispositifs, qui peuvent être exploités dans différents domaines (modélisme, réalité virtuelle, etc.). Ces capteurs ont le défaut d'avoir une dérive dans le temps, de l'ordre de plusieurs degrés par minute. Mais dans les applications en réalité virtuelle, telles que la mesure de rotation de la tête, il est plus important de connaître la rotation relative que la rotation absolue. Le poids peut être inférieur à 50 grammes tout en gardant une bonne résolution pour le traqueur.

Le gyroscope piézoélectrique de la société Murata est basé sur le principe de l'effet Coriolis. Il consiste en un petit prisme triangulaire piézoélectrique, avec une excitation sur une face et réception sur les deux autres. L'excitateur fournit des impulsions périodiques qui sont recueillies en réception sur les deux autres faces de façon identique quand l'appareil est au repos. Un quelconque mouvement de rotation du prisme produit des forces de Coriolis qui causent une augmentation d'amplitude du signal de réception sur un détecteur par rapport à l'autre. La différence entre les deux détecteurs est proportionnelle à la vitesse de rotation. Celle-ci doit donc être intégrée pour obtenir la rotation angulaire, avec les problèmes inhérents à cette méthode : la dérive des mesures implique de faire un recalage, si on a besoin d'une mesure absolue de position. Le traqueur gyroscopique de la société VR Systems exploite deux de ces appareils. Il existe d'autres principes de gyroscopes comme les gyroscopes à lames vibrantes (à diapason), mais ceux-ci ne sont pas employés actuellement dans les dispositifs de réalité virtuelle.

Une exploitation intéressante d'un gyroscope peut servir de traqueur à 3 ddl qui se tient dans la main. L'utilisateur peut ainsi manipuler aisément en rotation un objet virtuel. C'est par exemple le cas du traqueur «Gyropoint II Pro» de IXMicro Gyropoint (Figure 5.7). Ce capteur est «sans fil» et est basé sur des gyroscopes à «état solide».



*Figure 5.7: Le traqueur Gyropoint II Pro sans fil de IXMicro Gyropoint, copyright
photo IXMicro Gyropoint*

5.3.2.3 Accéléromètres

Les accéléromètres sont bien sûr basés sur la mesure d'une force provenant de l'accélération d'une masse. Les accéléromètres peuvent être en boucle ouverte ou en boucle fermée : pour la mesure en boucle ouverte, la force est mesurée par le faible allongement d'un ressort supportant la masse. Cette mesure est détectée par principe piézo-électrique, piézorésistif, à jauges de contrainte ou à variation de capacité. Pour ce type d'accéléromètre en boucle ouverte, il y a le problème classique de compromis à faire entre précision et rapidité de réponse. Pour les accéléromètres en boucle fermée, on crée une force opposée à la force inertielle, annulant le déplacement de la masse. Cette force d'opposition est créée par un système moteur dont la commande est fonction de cette force, donc de l'accélération. Ce principe permet d'obtenir des accéléromètres de meilleure précision (Figure 5.8).

La mesure de localisation de l'objet en mouvement est obtenue par double intégration, d'où une augmentation quadratique des erreurs de position. Les dérives dans le temps sont donc le principal inconvénient de ce type de traqueurs. L'accélération terrestre est à prendre en compte dans les mesures, par exemple avec un inclinomètre. Des accéléromètres miniatures existent, mais sont encore peu exploités en réalité virtuelle actuellement.

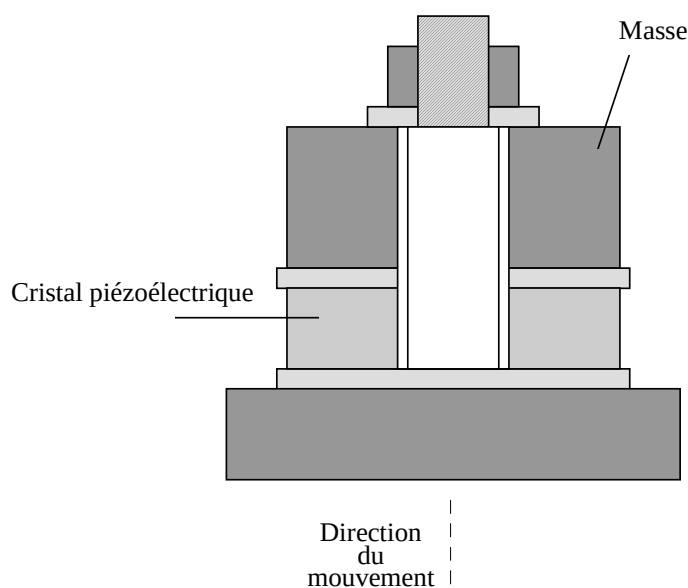


Figure 5.8 : Principe d'un accéléromètre

5.4 TRAQUEURS ÉLECTROMAGNÉTIQUES

5.4.1 TRAQUEURS ÉLECTROMAGNÉTIQUES À CHAMP ALTERNATIF

Ce type de traqueur est composé d'un émetteur, d'un récepteur et d'une unité électronique. L'émetteur est un ensemble de trois bobines, qui propagent des champs électromagnétiques dans une sphère environnante de quelques mètres. Le récepteur est lui aussi composé de trois petites bobines qui recueillent des flux magnétiques fonction de la position de celles-ci par rapport à l'émetteur [Raab, 1979].

Pour un point éloigné d'une bobine, où passe un courant alternatif

$$I = I_0 \exp(-i\omega t) \quad (5.1)$$

le champ magnétique B créé a pour valeur en coordonnées polaires r et θ dans un plan contenant l'axe de la bobine (Figure 5.9).

$$B_r = \mu_0 N I R^2 \cos(\theta) / 2r^3 \quad (5.2)$$

et

$$B_\theta = \mu_0 N I R^2 \sin(\theta) / 4r^3 \quad (5.3)$$

avec les conditions géométriques : $r \gg R$ et $r \gg L$
bobine de N spires, de rayon R et de longueur L .

μ_0 perméabilité du vide, λ longueur d'onde

Avec une bobine de l'émetteur, on crée un champ magnétique alternatif, dont l'intensité est fonction de r et de θ . La direction et l'intensité de ce champ sont déterminées avec les trois bobines du récepteur, orientées orthogonalement entre elles, car celles-ci recueillent des courants électriques mesurés, fonction de r et de θ . De même on crée deux champs magnétiques avec les deux autres bobines de l'émetteur, orientées aussi orthogonalement entre elles (Figure 5.10). On peut penser que pour déterminer les 6ddl du récepteur, on crée isolément avec chaque bobine un champ magnétique, chacun à un instant différent. Dans ces conditions, on aurait un système d'équations non linéaires, difficile à linéariser. En pratique, pour contourner cette difficulté, l'émetteur crée un champ dans la direction supposée du récepteur et dans les deux directions orthogonales correspondantes, à deux autres instants. Dans ces conditions, le système d'équations est facile à linéariser. La méthode de calcul permettant de calculer les 6 degrés de liberté du récepteur est donnée en annexe, à la fin du chapitre.

L'unité électronique, qui gère l'ensemble, a les fonctions suivantes :

- alimentation des trois bobines de l'émetteur par des courants alternatifs de fréquence porteuse voisine de 10 KHz ;

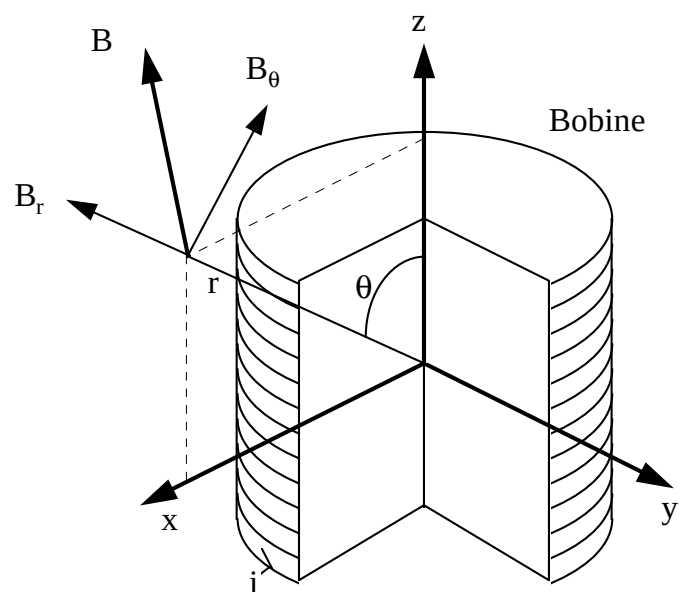


Figure 5.9: Champ électromagnétique autour d'une bobine

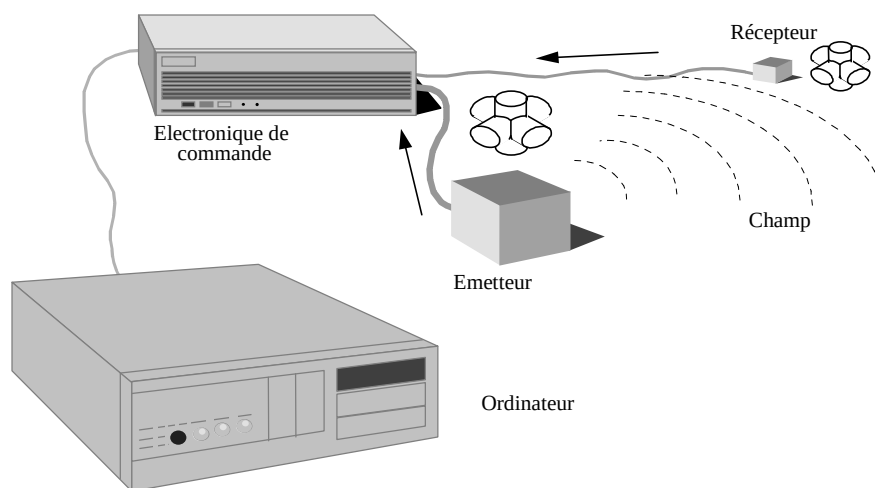


Figure 5.10: Dispositifs d'un capteur de localisation électromagnétique

- mesure des courants circulant dans les bobines du récepteur ;
- calcul des paramètres de localisation en fonction des mesures effectuées ;
- filtrage éventuel des mesures pour supprimer les bruits parasites, mais celui-ci augmente le temps de réponse ;
- transmission des valeurs calculées à l'ordinateur en communication avec l'appareil.

L'unité électronique peut gérer plusieurs récepteurs (4 par exemple) et le constructeur a prévu des fréquences porteuses différentes pour associer dans un même espace jusqu'à 4 émetteurs commandés par 4 unités différentes (16 récepteurs).

Les mesures et les calculs doivent être effectués le plus rapidement possible pour obtenir une localisation en temps réel. Le traqueur est d'autant plus efficace qu'il peut réaliser un grand nombre de mesures précises par seconde (120) et qu'il transmet les valeurs obtenues le plus rapidement possible. Les ordres de grandeur du temps de latence : en pratique 9 ms, avec un seul récepteur et sans filtrage, à 45 ms avec filtrage (voir au paragraphe suivant pour la comparaison avec l'autre type de traqueur électromagnétique). Un logiciel de communication et d'exploitation est fourni avec le dispositif pour faciliter l'utilisation du matériel.

La contrainte principale de ce type de traqueur est qu'il ne peut pas fonctionner dans un espace comportant des pièces métalliques. En présence de pièces conductrices, le champ magnétique variant dans le temps crée des courants induits dans ces pièces (les courants de Foucault). Les intensités de ces derniers sont proportionnelles à la fréquence du champ magnétique et créent des champs magnétiques complémentaires qui perturbent les mesures.

Le traqueur Ultratrak est la version permettant une utilisation sur une grande étendue de mesure, jusqu'à cinq mètres, et avec au maximum seize récepteurs en fonctionnement. L'émetteur est une sphère d'un diamètre de 18 pouces (Figure 5.11).

5.4.2 TRAQUEURS ÉLECTROMAGNÉTIQUES À CHAMP IMPULSIONNEL

Pour supprimer ou du moins atténuer les courants de Foucault, on peut réaliser les mesures avec des champs magnétiques constants non générateurs de tels courants. Les formules précédentes du champ magnétique restent valables en prenant $I = I_0$ constant.

$$B_r = \mu_0 N I_0 R^2 \cos(\theta) / 2r^3 \text{ et } B_\theta = \mu_0 N I_0 R^2 \sin(\theta) / 4r^3 \quad [4]$$

avec les conditions géométriques : $r \gg R$ et $r \gg L$
bobine de N spires, de rayon R et de longueur L .

μ_0 perméabilité du vide

Le champ magnétique d'une bobine de l'émetteur est impulsionnel pendant 2,5 ms environ. Au début de cet intervalle de temps, il y a création de courants de Foucault qui diminuent ensuite quand le champ magnétique devient constant. A ce moment les mesures sont effectuées sur les trois bobines du récepteur. Ce cycle séquentiel est répété deux fois pour les deux autres bobines de l'émetteur. Un quatrième intervalle de 2,5 ms est réservé à la mesure du champ électromagnétique du lieu qui dans ce type de

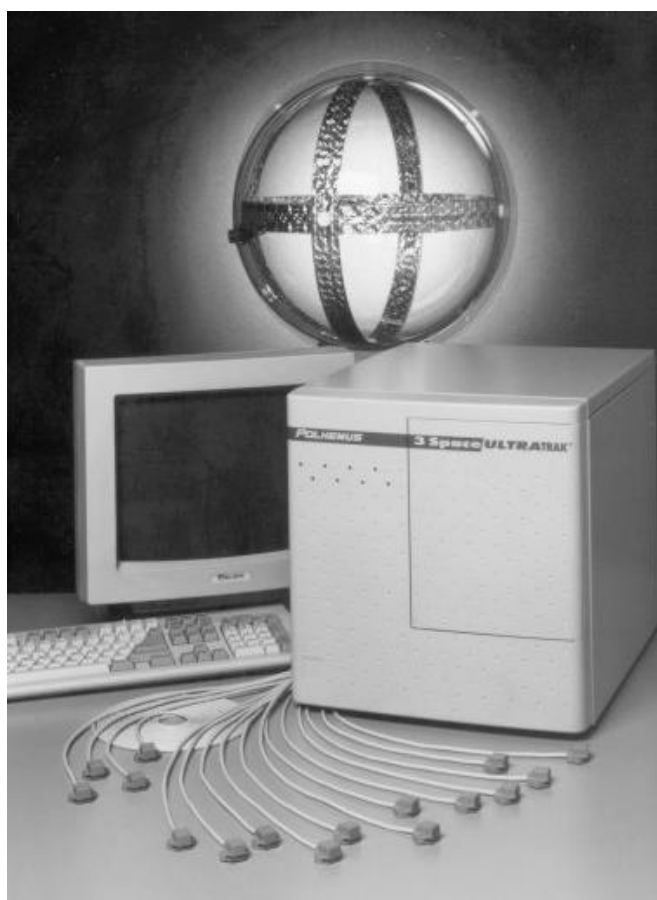


Figure 5.11 : Traqueur Ultratrak de Polhemus, copyright photo SimTeam

traqueur est déterminé (il peut donc être variable). Le cycle total des mesures d'une position du récepteur dure moins de 10 ms. La structure matérielle est similaire à celle du traqueur précédent, sauf dans la forme du bobinage. L'émetteur comme le récepteur sont composés de trois bobines enroulées sur un cube. Le récepteur a en plus au centre du cube un bobinage cylindrique complémentaire pour augmenter l'énergie magnétique réceptionnée. Le constructeur de ce type de traqueur propose l'emploi d'un émetteur avec plusieurs récepteurs (jusqu'à 30) sans diminution du nombre de mesures par seconde. Les ordres de grandeur du temps de latence : en pratique 8 ms, avec un seul récepteur et sans filtrage, à 47 ms avec filtrage. On peut aussi exploiter plusieurs émetteurs pour augmenter l'étendue de mesure (Figure 5.12).



Figure 5.12 : Traqueur Bird d'Ascension, copyright photo Immersion S.A.

Ce même constructeur propose, à tout concepteur de système de réalité virtuelle, des boucles magnétiques servant d'antennes à positionner soi-même (SpacePad). Les antennes peuvent être disposées horizontalement et verticalement dans une pièce ou sur tout dispositif (de jeux, par exemple), permettant une localisation à la fréquence de 120 Hz avec un récepteur (Figure 5.13). Ascension commercialise aussi un traqueur de petites dimensions (9 x 6 x 6 mm pour le récepteur) : le miniBird (Figure 5.14).

La société Northern Digital Inc (NDI) propose un nouveau capteur électromagnétique Aurora, ayant des récepteurs de 0,9 mm de diamètre, 10 pouvant être suivis simultanément. Les précisions annoncées sont de 1 mm en translation et de 0,5° en rotation. La fréquence de mesure est comprise entre 20 et 60 Hz, indépendant du nombre de récepteurs. Un logiciel détecte et informe des perturbations électromagnétiques.

5.4.3 CARACTÉRISTIQUES DES TRAQUEURS ÉLECTROMAGNÉTIQUES

Les traqueurs électromagnétiques présentent de bonnes performances quant à la précision des mesures. Ils sont les plus employés pour les applications liées à la réalité virtuelle, car ils sont transparents pour l'utilisateur. Ils ont pour principaux inconvénients :

- leur coût élevé. Une diminution du coût n'a pas eu lieu malgré le marché supposé



Figure 5.13 : Les antennes SpacePad de Ascension, copyright photo Immersion S.A.

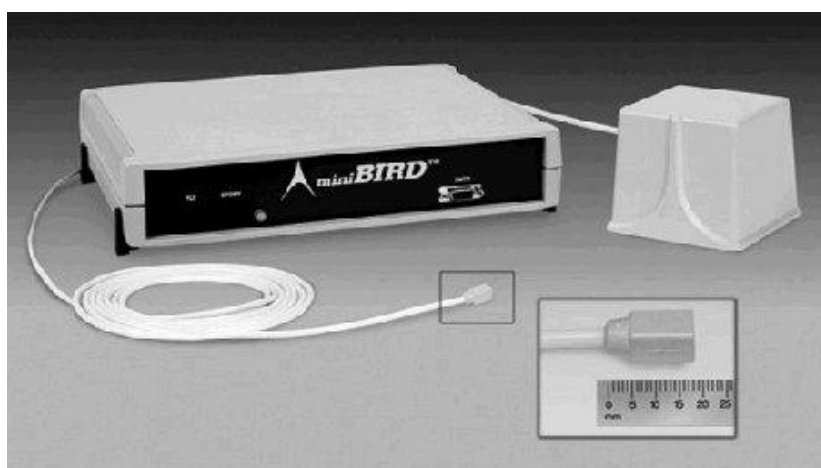


Figure 5.14 : Très petit récepteur du traqueur miniBird d'Ascension, copyright photo Immersion S.A.

porteur de traqueurs à 6 ddl peu onéreux et malgré la composition du matériel principalement électronique. Les prix de ces traqueurs sont actuellement trop élevés pour des applications «grand public» ;

- la liaison électrique par câble avec le récepteur peut être une entrave à la liberté de mouvement du solide observé. Des systèmes sans liaison par câble entre le récepteur et le boîtier électronique existent maintenant depuis 3 à 4 ans sur le marché. Ils sont présentés dans le chapitre des interfaces spécifiques de localisation corporelle car ils servent principalement pour la capture de mouvements du corps (motion capture) ;
- et surtout, l'impossibilité d'employer les appareils à champ magnétique alternatif dans un environnement métallique. Sauf si celui-ci est fixe : dans ce cas, il est possible de faire un étalonnage préalable pour compenser les perturbations. L'étalonnage est délicat à mettre en œuvre et n'est réalisé que pour des cas très particuliers, comme l'intérieur d'un cockpit d'avion.

Les temps de latence sont relativement courts et comparables entre les deux types de traqueurs. Ces temps augmentent s'il y a plusieurs récepteurs en fonctionnement et surtout si un filtrage des mesures est effectué. N'oublions pas que le temps de réponse global dépend aussi du temps de transmission des données (par liaison série ou parallèle) entre le contrôleur et l'ordinateur. Pour le lecteur désirant plus d'informations sur les temps de latence des différents modèles de traqueurs, ceux-ci ont été étudiés et mesurés par Adelstein et al [Adelstein, 1994]. Un gain de temps dans la transmission est réalisé si on incorpore l'électronique de mesure sur une carte directement connectable dans l'ordinateur. L'entreprise Polhemus le propose avec le traqueur InsideTrak et l'entreprise Ascension avec le traqueur pcBird, mais leurs caractéristiques sont de moins bonne qualité (Figure 5.15). Pour conclure sur cette catégorie de traqueur : ils sont les plus employés jusqu'à ce jour en réalité virtuelle, même si depuis dix ans leurs performances ont peu progressées¹ et leur prix peu diminué, à notre grande déception.

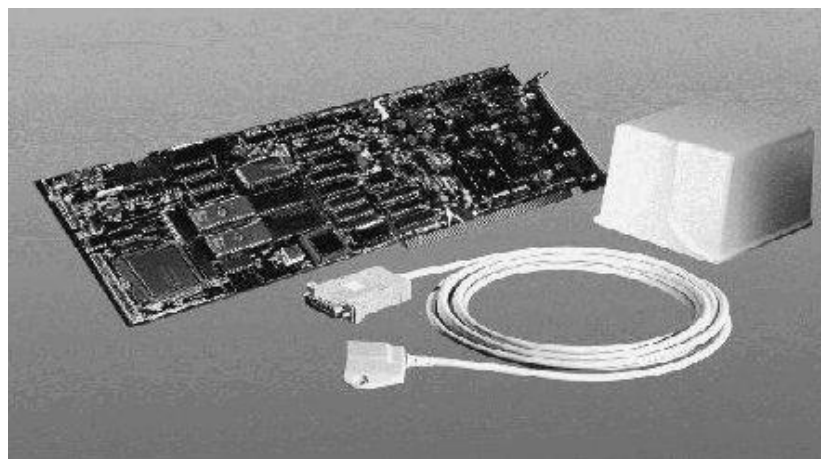


Figure 5.15 : Le traqueur sur carte PC : pcBird d'Ascension

¹Mais, pas d'inquiétude, elles sont relativement satisfaisantes pour la plupart des applications RV.

5.4.4 COMPAS

La terre crée un grand champ magnétique qui peut être utilisé pour des mesures de localisation. Les compas électroniques miniatures exploitent principalement l'effet Hall. Ce phénomène consiste en une déviation de courant dans une plaque métallique très mince, quand on place celle-ci dans un champ magnétique normal à son plan. Avec deux capteurs à effet Hall perpendiculaires entre eux, on peut déterminer le nord magnétique terrestre. Combinés avec un inclinomètre, on obtient un traqueur d'orientation à trois degrés de liberté. Par exemple, comme nous l'avons déjà indiqué, l'entreprise Intersense a plusieurs traqueurs d'orientation à technologie inertielle (Figure 5.16). Le dispositif inertiel mesure les vitesses angulaires de rotation avec un gyroscope, l'attraction terrestre avec un inclinomètre et le champ magnétique terrestre avec *des compas*.



Figure 5.16: L'InterTrak2 à 3 ddl d'orientation à technologie inertielle d'InterSense, copyright photo SimTeam

Mais ce type de traqueur a pour principal inconvénient d'être perturbé par les pièces métalliques environnantes et par les variations locales du nord magnétique. Dans le cas de l'emploi de mesures seulement relatives, ce problème est moins crucial.

Association de capteurs

Par exemple, l'entreprise Intersense a plusieurs traqueurs d'orientation à technique inertielle. Le capteur inertiel InterTrax2 de Intersense mesure 3 degrés de liberté (les 3 rotations sur 360°). C'est un traqueur bon marché, qui est fourni en standard avec des visiocasques bas de gamme, Glasstron de Sony et Scuba de Philips. Avec un dispositif inertiel miniature, le IS 300 Pro mesure les vitesses angulaires de rotation avec un gyroscope, l'attraction terrestre avec un inclinomètre et le champ magnétique terrestre avec des compas (voir paragraphe suivant). Les vitesses de rotation dans les trois directions sont intégrées pour obtenir les trois rotations. L'inclinomètre et le compas sont utilisés pour compenser les dérives du gyroscope. Un algorithme de prédiction du mouvement est fourni avec ce capteur, ce qui peut être intéressant pour diminuer le temps de retard dans la boucle «action - perception» de tout système de réalité virtuelle. La portée des appareils est illimitée par principe (limitée en pratique par la longueur des câbles) et il n'y a pas de problème d'occultation.

L'IS-300 et l'IS-300 Pro sont des capteurs d'orientation plus rapides et plus précis que

l'Intertrax. La fréquence d'échantillonnage de l'IS-300 Pro monte jusqu'à 500Hz. Sa vitesse maximale de rotation permise est de 1200°/sec. Sa résolution est de 0,02° RMS et sa précision est de 3° RMS. Intersense commercialise aussi un traqueur à six degrés de liberté à technologie hybride inertielle et ultrasons, le 600 Mark 2 (Figure 5.17).

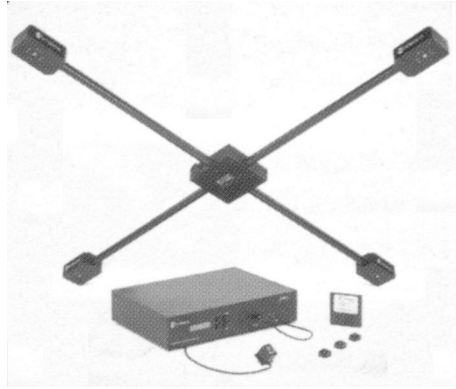


Figure 5.17 : Le traqueur 600 Mark2 de Intersense, copyright photo SimTeam

5.5 TRAQUEURS ACOUSTIQUES

Ce type de traqueur a pour principe la mesure de distances déterminées par des temps de propagation d'ultrasons dans l'air. Les lois de l'acoustique indiquent que la vitesse du son C dans l'air sec, à la pression normale, est fonction de la température :

$$C = 331.(T/273)^{1/2} \text{ en m/s (et } T, \text{ température ambiante, en degrés Kelvin)}$$

La vitesse est aussi fonction de l'humidité de l'air.

La technique consiste à exciter une céramique piézoélectrique de manière impulsionnelle, à une fréquence voisine de 40 KHz. Cette fréquence résulte d'un compromis entre des fréquences plus élevées où l'atténuation de l'onde dans l'air limite la portée (2 mètres environ à 80 KHz) et des fréquences plus basses pour lesquelles la longueur d'onde λ ($\lambda = C/f$) est trop grande, d'où aussi la résolution qui est égale à λ . Dans l'air à 40 KHz, la longueur d'onde est de $331/40000 = 8,3\text{mm}$. (à 200 KHz, $\lambda = 1,66 \text{ mm}$ - à 20 KHz, $\lambda = 16,6 \text{ mm}$).

Une autre contrainte concerne l'angle solide du faisceau sonore qui diminue quand la fréquence augmente. Une pastille de 20 mm de diamètre excitée à 40 KHz émet dans un cône de 30 ° environ. Pour une distance de 3,3 mètres, le temps de vol est de 10 ms, ce qui n'est pas négligeable si on veut un temps de réponse très court du capteur. Il faut ajouter au temps de vol le temps de l'émission impulsionnelle (environ 1 ms). Comme il faut au moins trois émetteurs, la mesure totale est trois fois plus grande si les mesures de distances sont faites cycliquement. On peut coder différemment les impulsions, soit en fréquence, soit en durée pour diminuer le temps de réponse du capteur, comme le proposent des chercheurs [Villarroya, 1994].

Le traqueur type est composé d'un ensemble triangulaire fixe de trois émetteurs d'ul-

trasons (céramiques piézoélectriques) et d'un ensemble triangulaire de trois récepteurs qui est lié à l'objet (céramiques identiques, la piézoélectricité étant un phénomène réversible). Les mesures des différents temps de vol des ultrasons entre les émetteurs et les récepteurs permettent de déterminer la position de l'objet (Figure 5.18). Il faut noter que cela n'est possible que si les émetteurs soient orientés vers les récepteurs. La zone d'étendue de mesure pour les orientations est donc plus limitée que pour les traqueurs électromagnétiques. En distance, la zone d'étendue est par exemple pour le S20 Sonic Digitizers de science Accessories Corporation de 16x8x8 pieds pour une précision 1,3 mm.

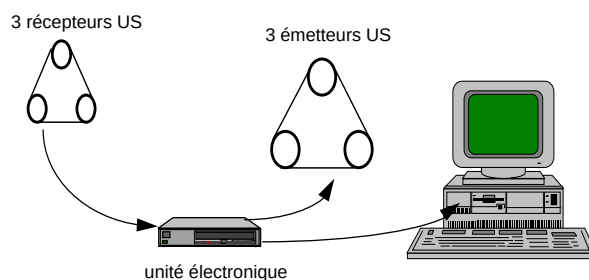


Figure 5.18 : Dispositifs d'un traqueur de localisation à ultrasons

Un système électronique gère les émissions d'ultrasons, les mesures de temps de propagation, les calculs des paramètres et la communication avec l'ordinateur exploitant les mesures. Le traqueur type peut réaliser 50 mesures par secondes. Son faible coût est un avantage par rapport au traqueur électromagnétique (Figure 5.19).

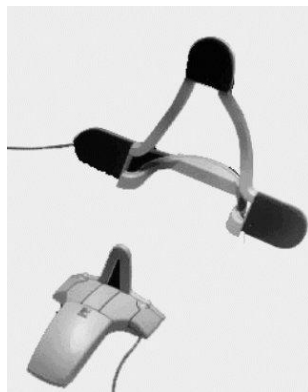


Figure 5.19 : Traqueur acoustique de Logitech qui n'est plus commercialisé, copyright photo Immersion S.A.

La liaison électrique par câble avec le récepteur peut être une entrave à la liberté de mouvement du solide observé, comme pour les traqueurs électromagnétiques, sauf dans le cas original suivant :

Normalement, la sonde émettrice et la sonde réceptrice doivent être reliées toutes les deux au boîtier électronique de commande pour la synchronisation des impulsions et

des mesures des temps. En associant les ultrasons et les infrarouges, ces derniers servant à synchroniser les émissions ultrasonores, on peut s'affranchir du câble entre le boîtier de commande et la sonde portée par la personne observée. Il fonctionne de la façon suivante :

- on dispose d'une part d'un petit boîtier portable, de la taille d'une boîte d'allumette, qui comprend un émetteur d'ultrasons et un récepteur d'infrarouges. Dans la pièce on place aux moins trois boîtiers fixes comprenant chacun récepteur d'ultrasons et un émetteur d'infrarouges ;
- le micro-ordinateur PC demande une mesure ;
- chaque récepteur d'ultrasons des boîtiers fixes envoie une impulsion infrarouge, se place en attente de réception d'une impulsion ultrasonore en chronométrant le temps d'attente ;
- si le boîtier portable reçoit au moins une des impulsions infrarouges, il émet une impulsion ultrasonore ;
- chaque récepteur d'ultrasons enregistre le temps d'arrivée de l'impulsion ultrasonore ;

l'ordinateur calcule par triangulation la position du boîtier portable, sachant que les temps de propagations des ondes infrarouges sont négligeables devant ceux des ondes ultrasonores. Ce système présente plusieurs avantages :

- il n'y a pas de liaison matérielle entre le boîtier portable et l'environnement ;
- l'étendue de mesure peut être grande en augmentant le nombre de récepteurs d'ultrasons ;
- le prix est faible, comparativement aux autres capteurs ;
- l'encombrement et le poids sont faibles ;
- la précision est de l'ordre du millimètre dans une aire de plusieurs mètres.

Ce système présente plusieurs inconvénients :

- le récepteur IR (le boîtier portable) doit voir au moins un émetteur IR ;
- la température influence les mesures car la vitesse du son est fonction de la température. Le dispositif doit être étalonné.

Mais ce traqueur est resté à l'état de prototype non commercialisé par Contrôle US. Un autre constructeur propose un traqueur, reposant sur le même principe, que l'on fixe à un doigt. Cette interface, a priori employée en pointeur 3D, exploite les ultrasons et les infrarouges. Ces derniers servent à indiquer au boîtier électronique l'instant du début de l'émission ultrasonore. Il n'y a donc pas de fil entre la main et le dispositif de mesure. Le capteur est composé d'un émetteur et de trois récepteurs d'ultrasons, qui sont fixés sur le bord de l'écran. La fréquence de mesure est de 50 Hz, la résolution est de 0,3 mm et l'étendue de mesure est limitée à 90 cm.

Il existe aussi des traqueurs acoustiques dont le principe est basé sur la mesure de déphasage entre l'émission sonore et la réception sonore. Dans ce cas les mesures sont

continues, ce qui permet un meilleur temps de latence et aussi de pouvoir filtrer les données pour diminuer les bruits.

Rappelons aussi que l'entreprise Intersense commercialise un traqueur à six degrés de liberté à technologie hybride inertielle et ultrasons, le 600 Mark 2 (voir le paragraphe correspondant). Les 4 récepteurs à ultrasons sont aux extrémités d'une croix et complètent les mesures données par la centrale inertielle (trois rotations).

Les traqueurs à ultrasons ont pour principal défaut d'être influençables par la température ambiante, la vitesse du son dépendant de la température. Il pourrait être prévu un étalonnage automatique par la mesure de la vitesse du son entre deux points fixes, comme cela existe sur certains traqueurs à ultrasons (mesure de niveau). Ces capteurs sont facilement bruités par tous les appareils qui émettent des ultrasons et par les échos sur les parois. Une entreprise a conçu des algorithmes de filtrage pour diminuer l'influence des bruits et des échos. Comme indiqué précédemment, l'émission d'ultrasons est très directive, ce qui limite l'espace de mesure. En contrepartie, les traqueurs à ultrasons ont pour avantage leur faible coût comparativement aux traqueurs électromagnétiques et, évidemment, de fonctionner dans un environnement métallique. Le couplage des techniques ultrasons et infrarouges peut être une solution d'avenir qui demande à être encore développée techniquement et commercialement.

5.6 TRAQUEURS OPTIQUES

Il n'y a pas dans cette catégorie de traqueur standard. La lumière est en général émise par l'intermédiaire de diodes électroluminescentes (LED) dans les infrarouges pour que la lumière ambiante ne perturbe pas les mesures. On peut aussi observer des cibles ponctuelles dans le spectre visible, la lumière ambiante étant mise à profit. Les traqueurs sont classés en deux groupes :

- les capteurs qui mesurent des angles à partir de rayons lumineux reçus sur des senseurs ponctuels infrarouges (phototransistors) ;
- les capteurs qui utilisent des récepteurs plans (caméras).

5.6.1 TRAQUEURS AVEC RÉCEPTEURS PONCTUELS

Des traqueurs ont été conçus pour détecter le mouvement de la tête d'un pilote d'avion, avec le principe suivant : sur le côté du cockpit, deux fins faisceaux lumineux plans tournent autour de deux axes de rotation parallèles. Leurs passages sur le casque sont détectés par des senseurs infrarouges, ce qui permet de connaître la rotation angulaire des faisceaux quand ceux-ci pointent sur les senseurs. Ces deux mesures d'angles permettent de déterminer chaque droite, parallèle aux axes de rotation, sur laquelle se trouve chaque senseur. Avec deux senseurs, on peut connaître deux angles d'orientation du casque. Pour augmenter la qualité de détection des mouvements, on peut augmenter le nombre de senseurs ou mettre un système identique de l'autre côté du casque [Ferrin, 1990] (Figure 5.20).

Un autre type de traqueur basé sur un ensemble de LED et de récepteurs ponctuels a pour principe la variation d'intensités reçues sur deux LED en fonction de leur position relative angulaire. Ce traqueur est encore sous la forme d'un prototype et pourrait

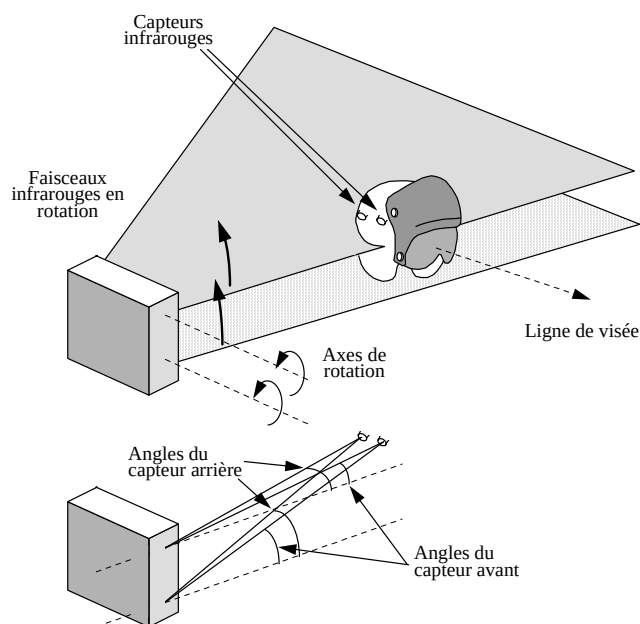


Figure 5.20 : Capteur de localisation optique avec récepteurs ponctuels

être employé dans le domaine des jeux virtuels. Il a été conçu par David Fenner, son inventeur. Le principe de base est le suivant [Fenner, 1993] :

Un transmetteur est composé de deux LED infrarouges et le récepteur a deux photodétecteurs adjacents (figure 1). Si la LED 1 envoie une impulsion, le rapport des intensités reçues sur les photodétecteurs est fonction de l'angle entre la LED 1 et les récepteurs 3 et 4 (figure 2). En plaçant un autre photodétecteur 5, la distance entre le transmetteur et le récepteur peut être calculée (figure 3). Par addition d'un autre récepteur, on peut déterminer la position spatiale de la LED 1 (figure 4). Les LED envoient des impulsions qui peuvent transmettre un codage correspondant à l'identification de la LED. Dans ce cas, les photodétecteurs peuvent être utilisés pour plusieurs LED. Cinq LED suffisent pour la détermination des 6 degrés de liberté du transmetteur. Le traqueur devrait être de faible coût. Le transmetteur a un faible poids et est alimenté par une petite batterie (Figure 5.21).

5.6.2 TRAQUEURS AVEC RÉCEPTEURS PLANS

A partir du traitement d'images acquises par des caméras, il est toujours possible de déterminer la position d'un objet à l'aide de deux ou trois caméras. Les solutions obtenues ne sont exploitables en général qu'au cas par cas. Ce vaste domaine scientifique de la reconstruction tridimensionnelle d'un environnement par stéréovision ou par plus de deux caméras n'est pas abordé dans son ensemble dans cet ouvrage. Nous en parlerons en partie au chapitre suivant pour la reconstruction du corps et pour la détermination des mouvements corporels. Il est aussi exploité pour le suivi du mouvement de la tête de l'observateur (face à un écran souvent). Mais dans ce cas, cela concerne plus la télécon-

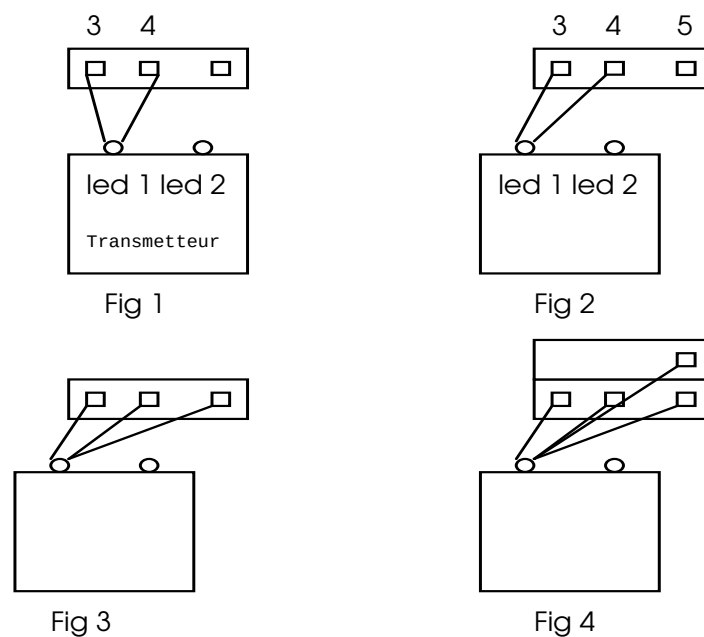


Figure 5.21 : Traqueur de localisation optique avec récepteurs ponctuels

férence, Internet et aussi la télévirtualité, en faisant vivre à distance un avatar (clone) de l'utilisateur. Nous ne présentons que des dispositifs susceptibles d'être exploités en interface comportementale. Certaines ne sont pas commercialisées ou anciennement, mais leur description peut aider l'ingénieur à en concevoir pour son application spécifique. On peut considérer deux principes de mesure suivant le type d'acquisition des images [Wang, 1990] :

- soit une ou plusieurs caméras sont fixes et observent des mouvements de repères ponctuels (LED ou marquage). C'est le principe de l'extérieur vers l'intérieur (outside-in) ;
soit une ou plusieurs caméras sont en mouvement sur l'objet
- à localiser et elles observent des repères ponctuels fixes. C'est le principe de l'intérieur vers l'extérieur (inside-out).

5.6.2.1 Principe de l'extérieur vers l'intérieur (outside-in) :

Il existe des capteurs spécifiques de localisation angulaire exploitant les mêmes dispositifs techniques que les caméras de traitement d'images. À l'aide d'un détecteur et d'un système optique associé, on détermine la position angulaire d'une balise ponctuelle émettrice en infrarouge par rapport à l'axe optique du dispositif. Plusieurs balises peuvent être repérées en ayant des codages lumineux différents. Les balises sont indépendantes des capteurs. Ce capteur est surtout exploité pour la détermination de trajectoires.

Ce principe a été utilisé pour suivre le mouvement d'une paire de lunettes (et donc de la tête), quand l'observateur regarde des images stéréoscopiques sur un écran. Pour augmenter la perception de la profondeur, comme nous l'avons indiqué au chapitre sur les sens humains (vision), nous pouvons changer en temps réel le point de vue des images stéréoscopiques en fonction du mouvement de la tête devant l'écran. C'est l'objectif du traqueur optique DynaSight de Origin Instruments Corporation. Le traqueur est un capteur compact qui mesure à 65 Hz la position tridimensionnelle (3ddl) d'une cible passive très légère (10 mg) qui se met sur la paire de lunettes stéréoscopiques (Figure 5.22).



Figure 5.22 : Le traqueur optique DynaSight pour le suivi de mouvement de paires de lunettes, copyright photo Origin Instruments Corporation

Il fonctionne près des infrarouges sur une étendue de mesure de 75° en élévation par 75° en azimut et en profondeur de 1,5 m (cible de 7 mm) à 5 m (cible de 50 mm). Le temps de latence est satisfaisant (9 à 28 ms). La précision est de 2 mm latéralement et de 8 mm en profondeur. Un dispositif adaptatif avec 4 cibles permet d'obtenir un traqueur à 6 ddl (Figure 5.23).

Avec un dispositif similaire, pour détecter le mouvement d'un visiocasque avec le principe outside-in, on place quatre LED sur le casque (théoriquement trois suffisent, mais la redondance de mesure augmente la précision, Figure 5.24). Connaissant les caractéristiques intrinsèques et extrinsèques de la caméra, on peut déterminer la position du casque à partir des huit coordonnées sur le plan image [Ferrin, 1990].

Au lieu de LED, on peut placer sur le visiocasque quatre marques. On n'est plus obligé d'alimenter les LED, mais le traitement des images est plus délicat. Plusieurs caméras sont nécessaires si on veut augmenter l'étendue de mesure.

A noter que certains dispositifs, conçus pour le «motion capture», peuvent être utilisés pour détecter seulement une partie du corps, comme deux doigts avec le système Coda de Charnwood Dynamics. Le dispositif, composé de trois caméras, suit en temps réel

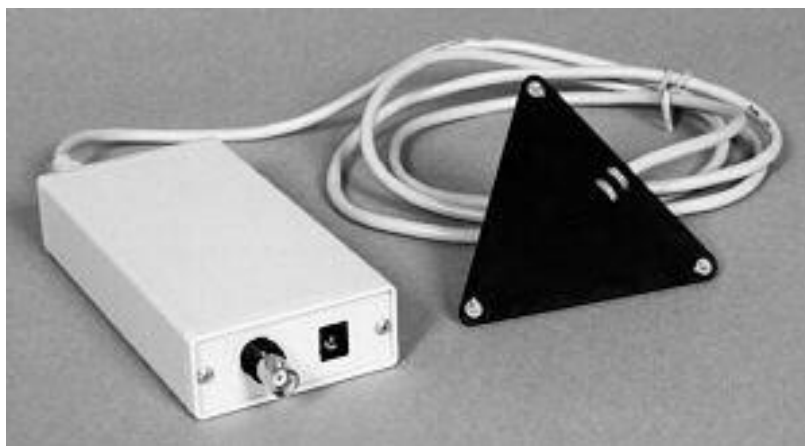


Figure 5.23 : Le traqueur optique à 6 ddl Active Target Adapter pour le suivi de mouvement de visiocasque, copyright photo Origin Instruments Corporation

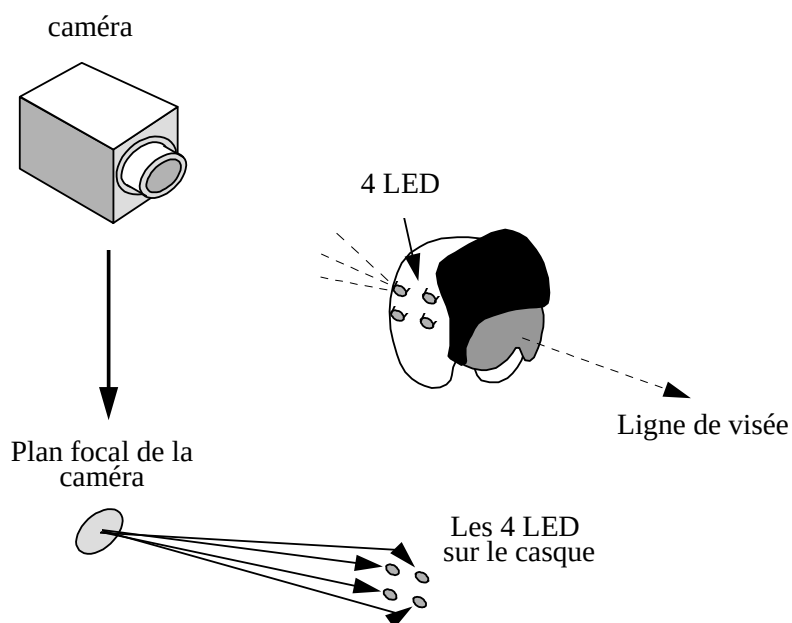


Figure 5.24 : Principe de l'extérieur vers l'intérieur pour un capteur de localisation optique

la position tridimensionnelle de cibles : des petites LED infrarouges, chacune étant identifiable, qui sont alimentées par des petites batteries portables (autonomie d'une journée de travail). Le temps de latence est de 5ms et la fréquence de mesure monte jusqu'à 800 Hz. 56 LED sont identifiables et la résolution est de 0,1 mm (précision inconnue, Figure 5.25).



Figure 5.25 : Le dispositif de 3 caméras et deux cibles sur deux doigts du système Coda motion capture pour le suivi de doigts, copyright photos Charnwood Dynamics

La difficulté du principe outside-in réside dans le fait que les repères ponctuels se déplacent peu relativement à la rotation de la cible, d'où une précision relativement faible en rotation. Dans le cas de mouvements de translation, la caméra doit avoir un grand champ de vision relativement à ces mouvements. Des capteurs de mesures tridimensionnelles exploitent le même principe dans le domaine de la métrologie. Avec trois caméras linéaires (2048 éléments) montées sur une poutre d'un mètre environ, le système Optotrak de Northern Digital Inc (NDI) peut donner la position de diodes émettrices infrarouges (jusqu'à 250). Les caméras linéaires, équipées de lentilles cylindriques, ont des résolutions «subpixel» meilleures que les caméras à CCD plan (1/200000 au lieu de 1/4000). Des pointeurs «crayons» sont recouverts de 24 diodes, permettant de supprimer les problèmes d'occultation, quand on manipule les pointeurs (Figure 5.26). Les caractéristiques sont très performantes : précision de 0,1 mm pour une zone à 2,25 m et d'étendue de 1,3 m sur 1,3 m, fréquence de transfert des données jusqu'à 400 Hz. Ces systèmes de localisation précise sont relativement chers et servent au départ dans le domaine de la métrologie. Mais NDI a développé son système aussi pour la capture de mouvements du corps pour la conception des jeux vidéo, de la réalité virtuelle, etc. (voir le chapitre des interfaces spécifiques de localisation corporelle).

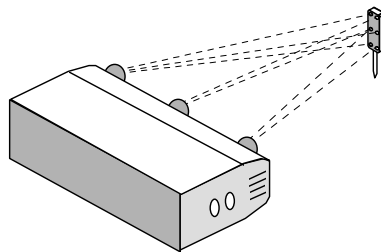


Figure 5.26 : Capteur de localisation optique : Système Optotrak

5.6.2.2 Principe de l'intérieur vers l'extérieur (inside-out) :

Avec le principe inside-out, en plaçant suffisamment de diodes fixes dans l'environnement, les mouvements relatifs des diodes par rapport aux caméras sont grands et les champs de vision de celles-ci peuvent être limités. Par contre, ce principe a pour inconvénient d'obliger la personne, dont on veut localiser la tête, à supporter le poids des caméras. Avec la miniaturisation des caméras «crayons» (moins de 100g), cet inconvénient peut disparaître. Il reste que le traitement des images exige une grande puissance de calculs et que la salle où évolue l'utilisateur doit être spécialement aménagée.

Avec trois caméras sur un visiocasque et une série de diodes fixées sur le plafond d'une pièce, des chercheurs [Wang, 1990] ont obtenu il y a plusieurs années un traqueur ayant de bonnes performances, fonction de la station de travail employée :

- précision : au mieux 1 cm et $0,3^\circ$;
- étendue de mesure : pièce de $4 \times 4 \text{ m}^2$.

Mais ce type de traqueur est bien complexe à mettre en œuvre, ce qui est son principal défaut (Figure 5.27).

Les traqueurs optiques ont le même inconvénient que les traqueurs à ultrasons concernant l'espace limité de mesure, tout récepteur optique devant «voir» un émetteur. On peut toutefois augmenter l'étendue de mesure en rajoutant des récepteurs ou des émetteurs, mais cela au détriment du coût. Ils sont rarement commercialisés en tant que traqueurs standards de localisation à six degrés de liberté, sauf pour le «motion capture», présentés au chapitre suivant.

5.7 COMPARAISON ENTRE LES DIFFÉRENTS TYPES DE TRAQUEURS

Les traqueurs mécaniques, mesurant des distances, sont intrinsèquement les meilleurs quant aux caractéristiques métrologiques, mais ils sont peu employés par manque de qualité ergonomique. Les traqueurs mécaniques, sans source émettrice artificielle, commencent à apparaître pour des applications en réalité virtuelle. Ils sont souvent vendus avec des visiocasques bas de gamme.

Les traqueurs électromagnétiques sont performants. Ils sont souvent employés dans les

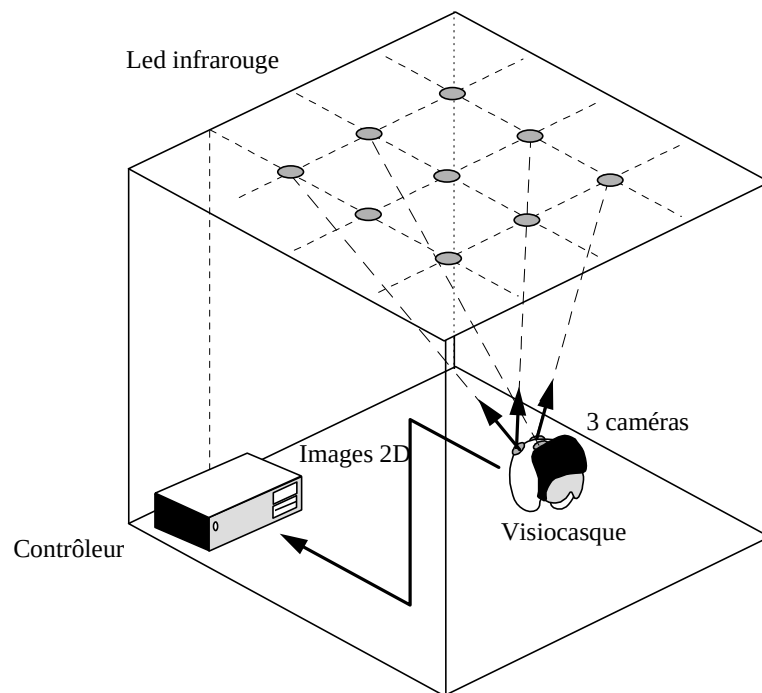


Figure 5.27 : Principe de l'intérieur vers l'extérieur pour un capteur de localisation optique

applications liées à la réalité virtuelle, sauf pour les produits «grand public» à cause de leur prix élevé. L'impossibilité d'exploiter ces appareils dans un environnement métallique reste leur principal défaut.

Les traqueurs à ultrasons ont pour principal défaut d'avoir une précision faible, d'être influençables par la température et l'humidité ambiantes et d'être perturbés par les bruits ultrasonores parasites. Ils ont pour avantages leur faible coût comparativement aux traqueurs électromagnétiques, et évidemment leur possibilité de fonctionner dans un environnement métallique. Le couplage ultrasons et infrarouges peut être une solution d'avenir bon marché qui demande à être encore développée techniquement et commercialement.

Les traqueurs optiques ont le même inconvénient que les traqueurs à ultrasons concernant l'espace limité de mesure. Il existe peu de traqueurs optiques standards pour la réalité virtuelle, ils sont réservés en général à des applications spécifiques.

Pour tous les traqueurs, il faut principalement comparer les caractéristiques métrologiques suivantes :

- le temps de réponse global (jusqu'au lieu d'exploitation des données);
- les précisions obtenues;
- la fréquence de transmission des mesures à l'ordinateur (à ne pas confondre avec la fréquence d'échantillonnage, qui peut être plus élevée);
- l'étendue de mesure.

Ces quatre paramètres sont les principaux critères quantitatifs de choix d'un traqueur. Après, on doit faire intervenir des considérations ergonomiques, en pensant aux contraintes d'utilisation du traqueur qui doivent être compatibles avec les I² sensori-motrices et mentales.

Comme nous l'avons déjà signalé, le traqueur universel parfait n'existe pas, mais le choix entre les différents types devient maintenant satisfaisant. Les problèmes techniques à résoudre ne sont pas aussi complexes que pour les autres interfaces, comme les visiocasques et les interfaces à retour d'effort. Le besoin d'interfaces à bon marché pour des produits «grand public» aurait dû permettre l'essor et la commercialisation de nouveaux traqueurs ou l'amélioration de ceux déjà existants, mais l'évolution est relativement lente.

Pour terminer, un petit schéma (Figure 5.28) résume symboliquement les différences entre les différents traqueurs.

5.8 ANNEXE : CALCUL DES PARAMÈTRES DU CAPTEUR ÉLECTROMAGNÉTIQUE FASTRAK

5.8.1 PRINCIPE

A l'aide des lois en électromagnétisme, on peut déterminer les courants électriques recueillis dans les bobines du récepteur à partir des courants électriques des bobines de

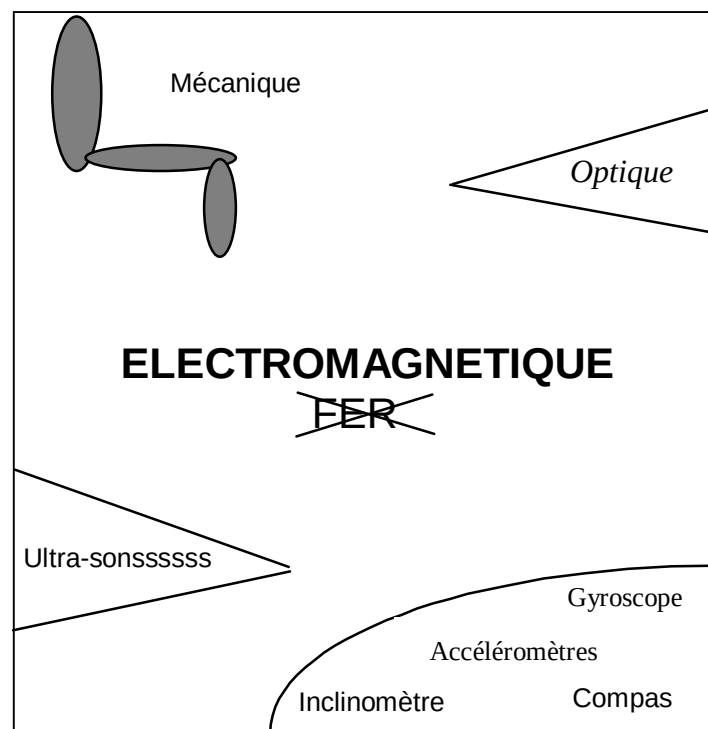


Figure 5.28 : Schéma résumant symboliquement les différences entre les différents traqueurs

l'émetteur et des 6 coordonnées de la position du récepteur par rapport à l'émetteur. Les relations n'étant pas linéaires, la détermination de la position (6 ddl) en fonction des courants dans les bobines est établie par approximation linéaire à partir de la position précédente.

5.8.2 DÉTERMINATION DES COURANTS DANS LE RÉCEPTEUR

Le repère de référence est celui lié à l'émetteur (X_1, Y_1, Z_1).

Le repère 5 est lié au récepteur.

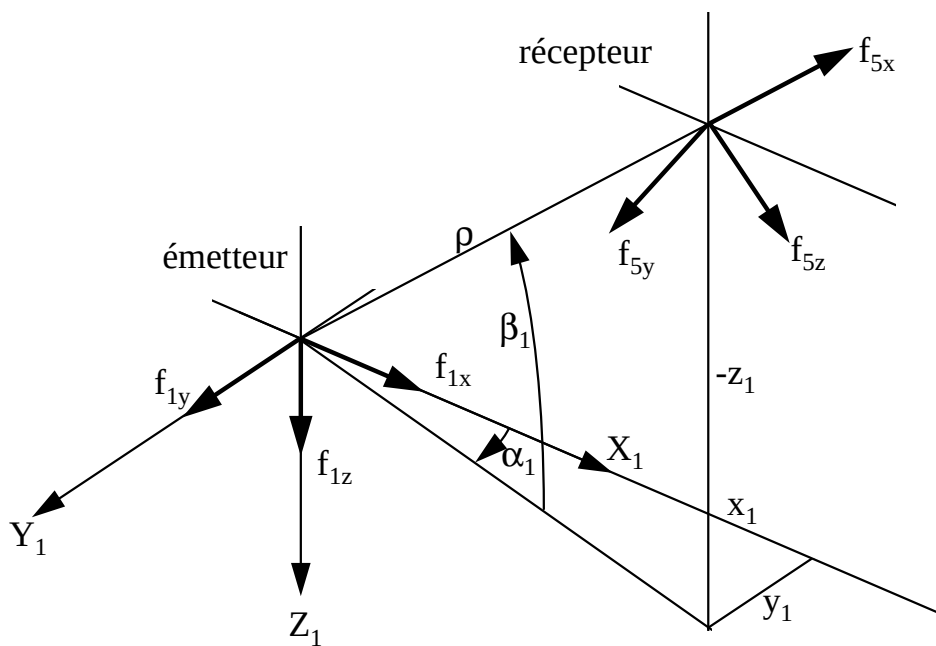


Figure 5.29 : Repères de référence

Les valeurs des dipôles magnétiques 3-axes de l'émetteur et du récepteur sont décrites par des vecteurs f_1 et f_5 : $f_1 = [f_{1x}, f_{1y}, f_{1z}]^T$. Les 3 bobines étant identiques, les coordonnées du vecteur f_1 correspondent aux amplitudes des courants.

Si on considère le couplage entre un émetteur 2 et un récepteur 3 ayant même alignement, on a la relation entre f_3 et f_2 :

$$f_3 = \frac{C}{\rho^3} S f_2 = \frac{C}{\rho^3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} f_2 \quad (5.4)$$

Le couplage entre l'émetteur et le récepteur dans une position quelconque fait intervenir les matrices de rotations :

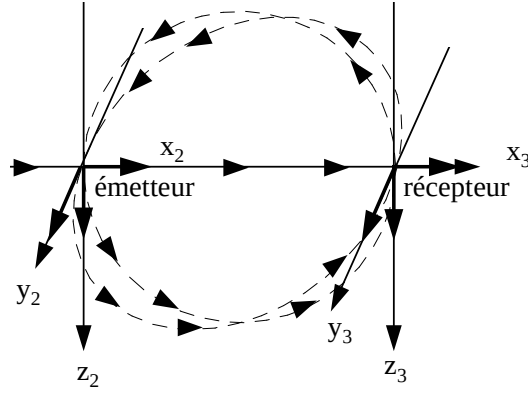


Figure 5.30 : Couplage entre l'émetteur et le récepteur

On définit les matrices suivantes pour les positions :

$$T_{\alpha} = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, T_{\beta} = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \quad (5.5)$$

Pour les rotations, on définit,

$$T_{\psi} = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, T_{\theta} = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \quad (5.6)$$

$$T_{\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (5.7)$$

On obtient donc f_5 en fonction de f_1 :

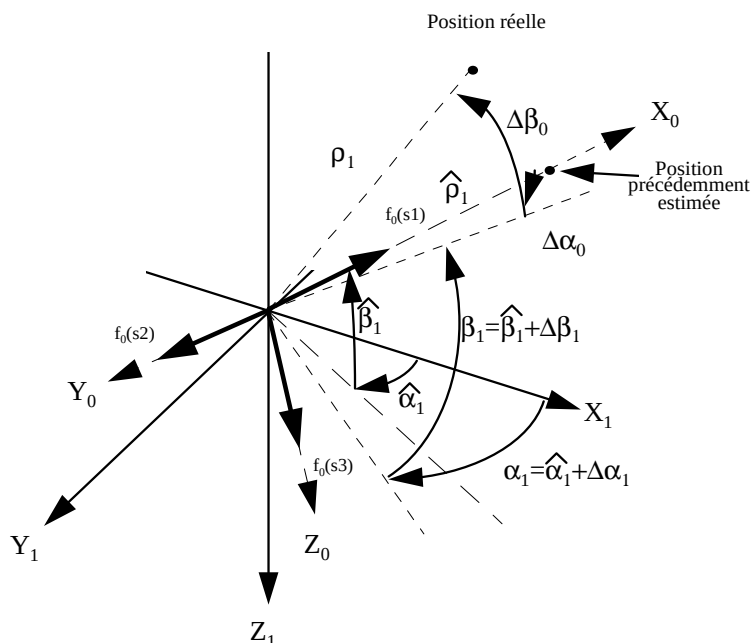
$$f_4 = \frac{C}{\rho^3} T_{-\alpha_1} T_{-\beta_1} S T_{\beta_1} T_{\alpha_1} f_1 = \frac{C}{\rho^3} Q f_1 \quad (5.8)$$

et

$$f_5 = T_{\phi_1} T_{\theta_1} T_{\psi_1} f_4 = \frac{C}{\rho^3} A f_4 = \frac{C}{\rho^3} A Q f_1 \quad (5.9)$$

5.8.3 DÉTERMINATION DES CHANGEMENTS EN POSITION ET ORIENTATION

La relation précédente n'est pas linéaire. Elle peut être linéarisée en faisant intervenir des petits changements en position et en orientation par rapport à la position et à l'orientation précédemment calculées. Les relations entre les positions actuelle et précédente, ainsi que pour les orientations, sont représentées sur la figure 5.8.3.



Le vecteur émetteur f_1 peut être défini dans le repère de «poursuite» estimé précédemment : vecteur f_0 .

$$f_1 = T_{-\hat{\alpha}_1} T_{-\hat{\beta}_1} f_0 \quad (5.10)$$

$$f_5 = \frac{C}{\rho^3} A T_{-\alpha_1} T_{-\beta_1} S \left(T_{\beta_1} T_{\alpha_1} T_{-\hat{\alpha}_1} T_{-\hat{\beta}_1} \right) f_0 = \frac{C}{\rho^3} A T_{-\alpha_1} T_{-\beta_1} S \Delta P f_0 \quad (5.11)$$

Le vecteur f_5 a été défini dans le repère récepteur. En le définissant dans le repère de «poursuite» estimé du récepteur : vecteur f_7 , on obtient une relation entre f_0 et f_7 qui ne fait intervenir que des petites variations angulaires :

$$f_7 = T_{\hat{\beta}_1} T_{\hat{\alpha}_1} A^{-1} f_5 = T_{\hat{\beta}_1} T_{\hat{\alpha}_1} T_{-\hat{\psi}_1} T_{-\hat{\theta}_1} T_{-\hat{\phi}_1} f_5 \quad (5.12)$$

$$f_7 = \frac{C}{\rho^3} \Delta A (\Delta P)^{-1} S \Delta P f_0 = \frac{C}{\rho^3} R f_0 \quad (5.13)$$

$$R = \begin{bmatrix} 1 & \frac{3}{2} \Delta \alpha_0 - \frac{1}{2} \Delta \psi_0 & -\frac{3}{2} \Delta \beta_0 + \frac{1}{2} \Delta \theta_0 \\ \frac{3}{2} \Delta \alpha_0 - \Delta \psi_0 & -\frac{1}{2} & -\frac{1}{2} \Delta \phi_0 \\ -\frac{3}{2} \Delta \beta_0 + \Delta \theta_0 & \frac{1}{2} \Delta \phi_0 & -\frac{1}{2} \end{bmatrix} \quad (5.14)$$

A partir d'un jeu de trois vecteurs excitations $f_0(S1)$, $f_0(S2)$ et $f_0(S3)$, on mesure trois vecteurs $f_7(S1)$, $f_7(S2)$ et $f_7(S3)$ qui permettent de calculer les variations de la distance et des angles. Celles-ci sont additionnées aux précédentes valeurs pour obtenir les nouvelles valeurs des 6 ddl :

$$f_0(S1) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, f_0(S2) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, f_0(S3) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (5.15)$$

$$\hat{\rho} = \sqrt[3]{C/f_{7x}(S1)} = \sqrt[3]{-C/2f_{7y}(S2)} = \sqrt[3]{-C/2f_{7z}(S3)} \quad (5.16)$$

$$\Delta \alpha_0 = \frac{2}{3} \cdot \frac{\hat{\rho}^3}{C} [2f_{7x}(S2) - f_{7y}(S1)] \quad (5.17)$$

$$\Delta \beta_0 = \frac{2}{3} \cdot \frac{\hat{\rho}^3}{C} [f_{7z}(S1) - 2f_{7x}(S3)] \quad (5.18)$$

$$\Delta \psi_0 = 2 \frac{\hat{\rho}^3}{C} [f_{7x}(S2) - f_{7y}(S1)] \quad (5.19)$$

$$\Delta \theta_0 = 2 \frac{\hat{\rho}^3}{C} [f_{7z}(S1) - f_{7x}(S3)] \quad (5.20)$$

$$\Delta \phi_0 = 2 \frac{\hat{\rho}^3}{C} f_{7x}(S2) = -2 \frac{\hat{\rho}^3}{C} f_{7y}(S3) \quad (5.21)$$

Remarques :

On a neuf équations pour six inconnues, ce qui permet de minimiser le bruit sur les mesures de deux variables. A l'initialisation, quelques itérations suffisent pour obtenir des mesures correctes.

5.9 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Adelstein, 1994] Adelstein et al., Presence, Vol. 5, N°3, MIT, 1994.
- [Ferrin, 1990] F. Ferrin, «*Survey of helmet tracking technologies*» SPIE proceedings, Large screen projection, avionic, and HMD, vol. 1546, 1990.
- [Fuchs, 1996] P. Fuchs, Les interfaces de la réalité virtuelle, AJIIMD, 1996, ISBN 2-9509954-0-3, les Presses de l'Ecole des Mines de Paris.
- [Fenner, 1993] D. Fenner, «*Pladar, capteur de localisation*», revue News VR, the Virtual Reality Newsletter, Cydata Limited, P.O. Box 2515, London N4 4JW, England, vol. juin 1993.
- [PCRV, 1994] Revue PCVR, N° 13, janvier 1994. PO. Box 475, Stoughton, WI 53589, USA.
- [Raab, 1979] F. Raab et E. Blood, «*Magnetic position and orientation tracking system*», IEEE Transaction on Aerospace and Electronic systems, N° 15, vol. 5, pp 709-18, septembre 1979.
- [Sato, 2001] M. Sato, «*Evolution of Spidar*», VRIC, Virtual Reality International Conference, proceedings, p 81-84, Laval Virtual, 2001
- [Villarroya, 1994] A. Villarroya Arranz et J. W. Flanigan, «*New tracking system for Virtual Reality based in ultrasonics and using transputers*» Actes des Journées «interface des mondes réels et virtuels», Montpellier, pp 359-368, février 1994.
- [Wang, 1990] J. F. Wang, V. Chi et H. Fuchs, «*A real-time optical 3D tracker for HMD systems*», Computer graphics, N° 2, p 24, 1990.

Adresses Web des entreprises :

Capteurs mécaniques

Fakespace	http://www.fakespace.com
Gyration	http://www.gyration.com
Systron Donner Inertial Div.	http://www.systron.com
VR Systems	http://www.ndirect.co.uk/vr-systems

Capteurs électromagnétiques

Ascension Technolgy Corp.	http://www.ascension-tech.com
General Reality Company	http://www.genreality.com
InterSense	http://www.isense.com
Polhemus	http://www.polhemus.com
Precision Navigation Inc.	http://www.precisionnav.com

Capteurs acoustiques

Alps Electric LTD	http://www.alps.ie
InterSense	http://www.isense.com
Logitech	http://www.vrdepot.com/vrteclg.htm
Pegasus Technologies	http://www.pegatech.com
Science Accessories Corp.	http://bei.moscow.com/sac.html
Visualization and Media Systems Design	http://viswiz.gmd.de

Capteurs optiques

Charnwood Dynamics	http://www.charndyn.com
Northern Digital Inc.	http://www.ndigital.com
Origin Instruments Corp.	http://www.orin.com/3dtrack/dyst.htm

6 LES INTERFACES SPÉCIFIQUES DE LA LOCALISATION CORPORELLE

Philippe Fuchs

6.1 INTRODUCTION

L'objectif de ce chapitre est de présenter les interfaces spécifiques de la localisation et de la détermination du mouvement du corps ou d'une partie de celui-ci, telles que les combinaisons de données ou les oculomètres. L'objectif de localisation est en général quelconque, sauf pour les interfaces spécifiques de locomotion. Dans ce cas précis, l'objectif principal de l'interface n'est pas de déterminer précisément les mouvements des pieds, mais de permettre grâce à ces derniers une *locomotion* de l'utilisateur dans l'environnement virtuel.

Toutes les interfaces spécifiques de la localisation sont exposées dans ce chapitre, sauf les gants de données (ou gants numériques). Ceux-ci sont présentés dans un chapitre inclus dans la partie des interfaces manuelles. Nous avons souhaité regrouper toutes les interfaces concernant la main, aussi bien les interfaces sensorielles manuelles, les interfaces sensori-motrices manuelles que les interfaces motrices manuelles. En effet, vu les fonctionnalités importantes de la main comme «outil» performant et privilégié de l'homme, beaucoup d'interactions et de sensations passent par la main. De plus, certains types d'interfaces manuelles, tels que les gants de données, peuvent évoluer d'interfaces motrices vers d'interfaces sensori-motrices. Pour analyser toutes ces considérations, les interfaces manuelles ont donc été regroupées, mais les gants de données auraient pu être inclus dans ce chapitre des interfaces spécifiques de localisation corporelle.

Ces interfaces motrices mesurent en temps réel la position de la partie concernée du corps permettant le suivi de sa trajectoire. Ces dispositifs sont nécessaires dans les applications de réalité virtuelle d'immersion corporelle. Ils ont été spécialement conçus pour la réalité virtuelle, sauf les oculomètres (capteurs détectant les mouvements orbitaires des yeux) et les interfaces de «motion capture» (mesure du mouvement de l'ensemble du corps). Ces dernières ont été développées pour le marché important des jeux vidéo et pour les études biomécaniques en médecine et en sport.

Ces interfaces spécifiques améliorent l'interaction du sujet dans le monde virtuel, car ce dernier peut ainsi agir de façon assez naturelle : par des mouvements gestuels, identiques si possible à ceux qu'il exécute dans le monde réel. Il augmente ainsi sa qualité d'immersion. Les mouvements corporels, que l'on désire détecter, sont principalement les suivants :

- de la main, organe performant pour la manipulation ;
- de l'ensemble du corps¹, pour une immersion totale du corps ;
- de la locomotion pour permettre à l'utilisateur des déplacements naturels dans le

¹Mais seulement pour ces grands constituants : la tête, le tronc et les différents segments des membres ; donc par exemple pour les mouvements des doigts, du visage, etc.

monde virtuel ;

- des yeux, pour améliorer la restitution visuelle ;
- du visage et des lèvres, pour animer le visage et(ou) la bouche, de la représentation virtuelle de la personne.

L'interface universelle permettant une détection *complète et fine* du corps entier (jusqu'aux extrémités des doigts !), regroupant tous les cas précédents, est utopique actuellement et serait fort coûteuse.

Dans sa globalité, il s'agit de la problématique de détermination du positionnement et du mouvement pour des solides déformables. Ce problème est complexe à résoudre, surtout si nous désirons que les dispositifs gênent peu l'utilisateur, en particulier dans ses mouvements. C'est la notion de transparence des interfaces, importante en réalité virtuelle, qui implique que l'ergonomie des dispositifs soit satisfaisante. Pour atteindre ce but, il est donc souhaitable que les réponses motrices soient transmises, si possible, sans support matériel entre l'homme et la machine. Nous pensons évidemment à l'emploi de caméras permettant, par traitement d'images, de suivre le mouvement du corps ou d'une de ses parties. Cette technique est la plus prometteuse sur le plan théorique, mais elle demande encore à être développée. Nous ne développons pas dans ce livre les problèmes complexes que ces traitements d'images induisent ; nous nous contentons d'indiquer au lecteur que les recherches sont toujours en cours dans ce domaine pour suivre les mouvements du corps, de la main ou des yeux, *sans marquage* sur ceux-ci. Si, en fonction de l'application projetée, il est possible de mettre des marques ponctuelles sur les parties en mouvement, le problème de traitement d'images s'en trouve facilité, mais reste encore délicat pour des détections en temps réel. Dans ce cas, des dispositifs commercialisés existent.

Notre présentation suit une approche anatomique descendante allant du général au particulier : localisation du corps, interfaces de locomotion, localisation des membres supérieurs, expression du visage ou des lèvres et localisation des yeux. Un paragraphe à part concerne le cas particulier des biocapteurs, d'emploi réservé au corps médical.

6.2 INTERFACES DE LOCALISATION DU CORPS «MOTION CAPTURE»

6.2.1 DEUX PRINCIPES DE «MOTION CAPTURE»

Les interfaces de localisation du corps sont utilisées pour détecter les mouvements de tout le corps et aussi souvent pour déterminer la configuration du squelette à tout instant. Dans le langage habituel, on emploie en général pour désigner ce type d'interface le terme anglais «motion capture», plutôt que capture du mouvement. Il existe deux familles de solution, totalement différentes : les interfaces portables de localisation du corps et les interfaces de localisation optique.

La première solution consiste à employer des interfaces portables, soit :

- la combinaison de données ayant un grand nombre de capteurs de localisation fixés sur la combinaison ;
- l'exosquelette, système mécanique composé de barres articulées portées par les différentes parties du corps.

Mais les exosquelettes sont souvent destinés à fournir en premier lieu des retours d'effort sur le corps, tout en mesurant les mouvements relatifs entre les membres. Dans ce chapitre, nous ne parlerons que des exosquelettes qui sont seulement des interfaces de localisation du corps : exosquelettes de localisation. Les problèmes logiciels de modélisation de personnages virtuels sont exposés au paragraphe 15.10.

6.2.2 INTERFACES PORTABLES DE LOCALISATION DU CORPS

La plupart du temps, on utilise les combinaisons de données dans cette catégorie d'interfaces, les exosquelettes posant des contraintes ergonomiques importantes. Toutefois, comme pour les traqueurs mécaniques, les qualités métrologiques des exosquelettes sont un avantage : précision élevée et temps de latence très court. Les combinaisons de données utilisent évidemment les classiques traqueurs électromagnétiques.

6.2.2.1 Interfaces de localisation électromagnétique

On emploie les capteurs électromagnétiques car ceux-ci n'ont pas de problème d'occlusion entre l'émetteur et les récepteurs. Ceci est le principal avantage par rapport aux interfaces basées sur une localisation optique par caméras. Grâce aux traqueurs électromagnétiques, la combinaison de données donne des résultats satisfaisants pour le suivi du mouvement du corps, si celle-ci est équipée d'un assez grand nombre de traqueurs judicieusement placés. Les qualités métrologiques des mesures dépendent directement de celles des traqueurs, sous réserve que ces derniers soient bien fixés sur le corps. Mais la limite d'étendue de mesure peut poser problème dans ce cas de motion capture, car les espaces de travail sont souvent assez grand par rapport aux capacités de ces traqueurs. Les sociétés Ascension et Polhemus, qui fabriquent les capteurs de position à 6 degrés de liberté, proposent des combinaisons de données comprenant jusqu'à une centaine de traqueurs. Le nombre de câbles liés à la personne est une gêne non négligeable. Compte tenu du nombre important de capteurs mis en œuvre, Ascension et Polhemus commercialisent des versions «wireless»² pour les combinaisons. Les capteurs ont toujours pour principal défaut d'être perturbés par les objets métalliques et les champs magnétiques.

Le «MotionStar Wireless» sans fil d'Ascension utilise les champs magnétiques générés par son émetteur longue portée pour mesurer la position et l'orientation de ses capteurs. Les récepteurs sont positionnés sur le corps aux points où l'on souhaite mesurer les mouvements. Les données issues des capteurs sont transmises par des câbles vers un émetteur radio fonctionnant sur batteries et porté à la ceinture (Figure 6.1). Les données des récepteurs et d'autres périphériques (tel un gant de données) sont ensuite transmises à la station de base.

Les caractéristiques du «MotionStar Wireless» sont les suivantes : 14 traqueurs à 6 ddl par acteur ; étendue de mesure de 3 m dans chaque direction ; précision en position de 0,75 cm RMS à une distance de 1,5 m et de 1,5 cm RMS à une distance de 3 m ; précision en orientation 0,5° à 1,5 m et de 1° à 3 m ; la fréquence d'échantillonnage monte Jusqu'à 120 mesures par seconde.

²Cela signifie en pratique qu'il n'y a pas de liaison électrique entre l'utilisateur et le boîtier électronique relié à l'ordinateur.



Figure 6.1 : La combinaison de données MotionStar sans fil d'Ascension, copyright photo Immersion S.A.

6.2.2.2 Localisation par exosquelette

Le produit de la société allemande X-IST Realtime Technologies change des combinaisons traditionnelles : celle-ci propose un exosquelette souple fixé à l'ensemble du corps. Le dispositif n'est pas perturbé par un environnement métallique et même les gants de données associés peuvent s'utiliser sous l'eau !

Le «Body tracker» de X-IST Realtime Technologies est une interface de capture de mouvements par un système mécanique de tiges et de codeurs que porte la personne. Cet exosquelette est performant et adaptable à la morphologie de la personne. Son étendue de mesure est de 20 mètres. Le poids total est de 6,8 kg et il n'y pas beaucoup de friction au niveau des articulations. Ce système est proposé par l'entreprise avec les suivis de mouvements de la tête («Face tracker», voir plus loin) et des mains, pour donner un système complet : le «Full Body tracker». Le dispositif est relié à l'ordinateur PC par une simple liaison RS232.

La capture de mouvement par exosquelette Gypsy de Analogus Corporation est basé sur le même principe mécanique de tiges et de codeurs que porte la personne. Ce système n'a pas de limite en étendue pour l'enregistrement de mouvement, avec une autonomie de 4 heures, fournie par une batterie portable. Pour l'animation en temps réel, le système est relié à un ordinateur PC par un simple câble Ethernet. Son prix est de 10 000 \$ US environ (Figure 6.2) :

6.2.3 INTERFACES DE LOCALISATION OPTIQUE

La deuxième solution, par traitement d'images, est plus ergonomique puisqu'elle laisse toutes libertés de mouvement pour la personne. Si nous ne pouvons pas poser de marques sur le corps observé, les difficultés de traitement d'images sont importantes, comme nous l'avons déjà signalé. Par contre, si on peut coller des repères sur le corps (pastilles réfléchissantes, cibles électroluminescentes ou LED infrarouges), les traitements sont plus aisés et plus rapides mais la mise en œuvre peut être délicate, surtout pour la gestion de l'éclairage, des occultations et de la zone d'observation des caméras. Les produits sont de prix trop élevés pour une utilisation non professionnelle, mais nous pouvons en espérer progressivement la baisse des prix. Un exemple de caractéristiques avec un système de caméras et d'éclairage par infrarouges de petites marques passives : de 2 à 8 caméras, ayant des champs de vision de 0,2 à 10 mètres en profondeur, la précision est de 1/2800 du champ de vision en 3D dynamique, ce qui donne une précision de 1 mm environ pour un champ de 2,8 mètres, avec un nombre illimité de marqueurs.

Cette méthode peut souffrir de l'occultation des parties du corps qui se retrouvent masquées dans certaines positions. Une autre difficulté technique est le nombre imposant de calculs, engendrés par l'analyse des images, mais la puissance actuelle des PC commencent à estomper ce nœud technologique. L'éclairage est aussi délicat à contrôler car il peut perturber les mesures. Enfin, des caméras très rapides doivent être employées en général, car les 25 acquisitions par seconde d'une caméra classique ne donnent pas souvent de résultats satisfaisants, surtout pour des mouvements rapides de la personne.

La capture de mouvements par le «Vicon Motion Capture systems» de Vicon est basée sur la détection par caméra de marqueurs passifs réfléchissants. Il n'y a donc pas de



Figure 6.2 : La capture de mouvement par exosquelette Gypsy de Analogus Corporation, copyright photo Analogus Corporation

batteries à porter et de câbles de liaison. Les caméras précises fonctionnent jusqu'à 120 images par secondes. Elles peuvent suivre 50 marqueurs pour déterminer les mouvements de 20 à 30 segments du corps. Les positions 2D des marqueurs permettent de déterminer les positions 3D des segments du corps, en tenant compte de considérations anatomiques. Les personnages peuvent être animés ensuite jusqu'à la fréquence de 30 images par secondes sans temps de latence perceptible.

Un nouveau produit a été développé par Ascension sur un principe optique : ReActor. Ce système est basé sur un ensemble de 12 barres qui forment les arêtes d'un parallélogramme de 3m x 3m x 4m. Des caméras numériques infrarouges détectent 30 marqueurs actifs et infrarouges. La vitesse de capture est de 900 par seconde et la transmission des données peut monter jusqu'à 60 Hz. Des batteries doivent être portées dans un sac de 900g par l'utilisateur. Le système mesure les mouvements des marqueurs et détermine en temps réel les mouvements du corps.

La capture de mouvements par le «Eagle Digital System» de Motion Analysis est aussi basée sur la détection de marqueurs passifs par caméra spécifique, à résolution de 1280 x 1024 à la fréquence de 600 millions de pixels par seconde. L'affichage et l'animation du squelette se font en temps réel, à la fréquence des caméras vidéo. La précision est bonne, inférieure à 1%, et le temps de latence est imperceptible.

La société Northern Digital Inc. (NDI) propose un système complet de capture de mouvement temps réel pour le visage, les mains et le corps : le système Optotrak, composé d'un banc de trois caméras infrarouges. Nous avons présenté ce dispositif au chapitre sur les capteurs de localisation. Le système comprend jusqu'à 256 marqueurs et il est capable d'acquérir jusqu'à 150 images par seconde, permettant des captures sans câble de plusieurs acteurs. Le système Optotrak suit les déplacements des marqueurs infrarouges, qui sont tous identifiables (utile pour résoudre les problèmes d'identification après occultation). Le capteur est calibré en usine. Les caractéristiques sont très performantes : précision de 0,1 mm en X et Y et 0,15 mm en Z pour à une distance de 2,5 m avec une étendue de mesure de 1,3 m sur 1,3 m. Les données 2D sont traitées à 450 Hz et fournissent des données 3D à 60 Hz (Figure 6.3). Plusieurs systèmes de capture peuvent être couplés pour augmenter la flexibilité et l'étendue de mesure. La capture de mouvement peut fonctionner avec la lumière, puisqu'il travaille en infrarouge (mais les reflets de ces derniers doivent être évités). NDI commercialise aussi un autre système de motion capture optique à plus bas prix : le Polaris.

La société Hypervision propose un système de capture de mouvement à partir d'une combinaison possédant des marqueurs actifs prédisposés, ce qui permet un calibrage rapide de l'interface. Les marqueurs sont alimentés électriquement par une batterie portée. Il n'y a pas de câble qui empêchent les mouvements de la personne. Les positions XYZ des marqueurs sont calculées en quelques millisecondes à une précision de quelques millimètres.

Le système Coda de Charnwood Dynamics est composé de trois caméras mises dans un seul boîtier (Figure 6.4). Il suit en temps réel la position tridimensionnelle de cibles : des petites LED infrarouges, chacune étant identifiable, qui sont alimentées par des petites batteries portables (autonomie d'une journée de travail). Le temps de latence est de 5 ms et la fréquence de mesure monte jusqu'à 800 Hz. 56 LED sont identifiables et la résolution est de 0,1 mm (précision non communiquée).

Différentes techniques sont utilisées en motion capture. Tous les concepteurs de nou-

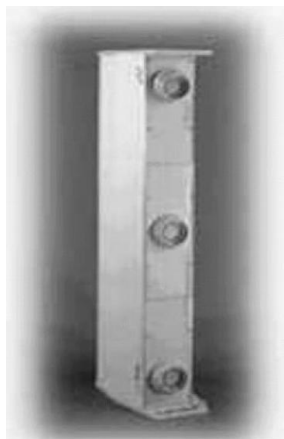


Figure 6.3 : Système Optotrak de Northern Digital Inc., copyright photo Northern Digital Inc.



Figure 6.4 : Le dispositif de 3 caméras du système de Coda motion capture, copyright photo Charnwood Dynamics

veaux systèmes recherchent à supprimer les insuffisances des autres capteurs. Nous pouvons citer, par exemple, les ambitions de l'entreprise XD Productions qui conçoit un système efficace de motion capture «le Cyberdôme», tout en recherchant à minimiser les coûts, grâce entre autres par l'exploitation de caméras classiques peu onéreuses. Le Cyberdôme est un outil de recherche et de production installé dans les locaux CNBI à Angoulême, constitué par une structure en bois de forme hémisphérique de 10 mètres de diamètre sur laquelle sont fixées, à espace régulier, 14 caméras vidéo (Figure 6.5). Ce dôme est monté sur un plateau de 400 m², doté d'une hauteur de 5,5 mètres. Connectées à des ordinateurs PC, les caméras ont ainsi la capacité, grâce aux calculs de paires stéréoscopiques de calculer les coordonnées de plusieurs points lumineux positionnés sur un acteur en mouvement. Ces coordonnées sont ensuite appliquées sur le squelette virtuel d'un personnage 3D synthétique, dont l'animation suit automatiquement les mouvements de l'acteur.

Les points du corps de référence sont facilement repérés en temps réel par des cibles électroluminescentes de couleurs suffisamment différenciables par les caméras, ce qui est une des originalités du Cyberdôme. Ces cibles sont alimentées à base tension par une batterie que porte la personne. Le problème de reconnaissance des marqueurs ne pose donc pas de problème insurmontable dans ce type d'approche avec des caméras vidéo à prix modéré. La structure du dôme et le nombre de caméras, qui peut facilement être augmenté sans contrainte technique, est bien approprié pour la capture de mouvement dans un espace suffisamment grand.

6.3 INTERFACES DE LOCALISATION DES MEMBRES SUPÉRIEURS OU INFÉRIEURS

Il n'existe pas beaucoup d'interfaces spécifiques de détection du bras ou de l'avant bras. La société Exos se distingue dans la capture spécifique du mouvement du bras par exosquelette. Il permet de mesurer 5 degrés de liberté pour le bras et l'avant-bras, grâce à la détection des mouvements relatifs de l'épaule (3 degrés de liberté), du coude (1 degré de liberté) et de l'avant-bras (1 degré de liberté). Les mécanismes sont modulaires et peuvent donc être utilisés séparément. Les capteurs sont des potentiomètres de précision. Le poids des deux modules sur le bras est relativement faible, environ 500 g, permettant d'améliorer la transparence de l'interface. Les précisions sont de 0,1 ° au niveau des articulations, ce qui implique des précisions de l'ordre de 1,5° au niveau des articulations (Figure 6.6). Il existe une variante de cet exosquelette avec retour d'effort (voir le chapitre correspondant).

La capture par exosquelette Upper Body Waldo de Character Shop est basée sur un principe mécanique de tiges en plastique, pour réduire le poids et les interférences avec le traqueur électromagnétique. Les mesures se font par des codeurs angulaires optiques ou potentiométriques. L'exosquelette mesure 3 ddl de la tête, 2 ddl pour chaque épaule, l'inclinaison du torse, 2 ddl pour chaque coude et 2 ddl pour chaque poignet (Figure 6.7).

Cette entreprise propose plusieurs systèmes adaptés aux applications des clients : pour un seul bras, pour les jambes, pour le visage (voir plus loin), etc. Le système pour deux bras repose sur un trépied et mesure 3 ddl pour chaque épaule, 1 ddl pour chaque coude et 2 ddl pour chaque poignet (Figure 6.8).

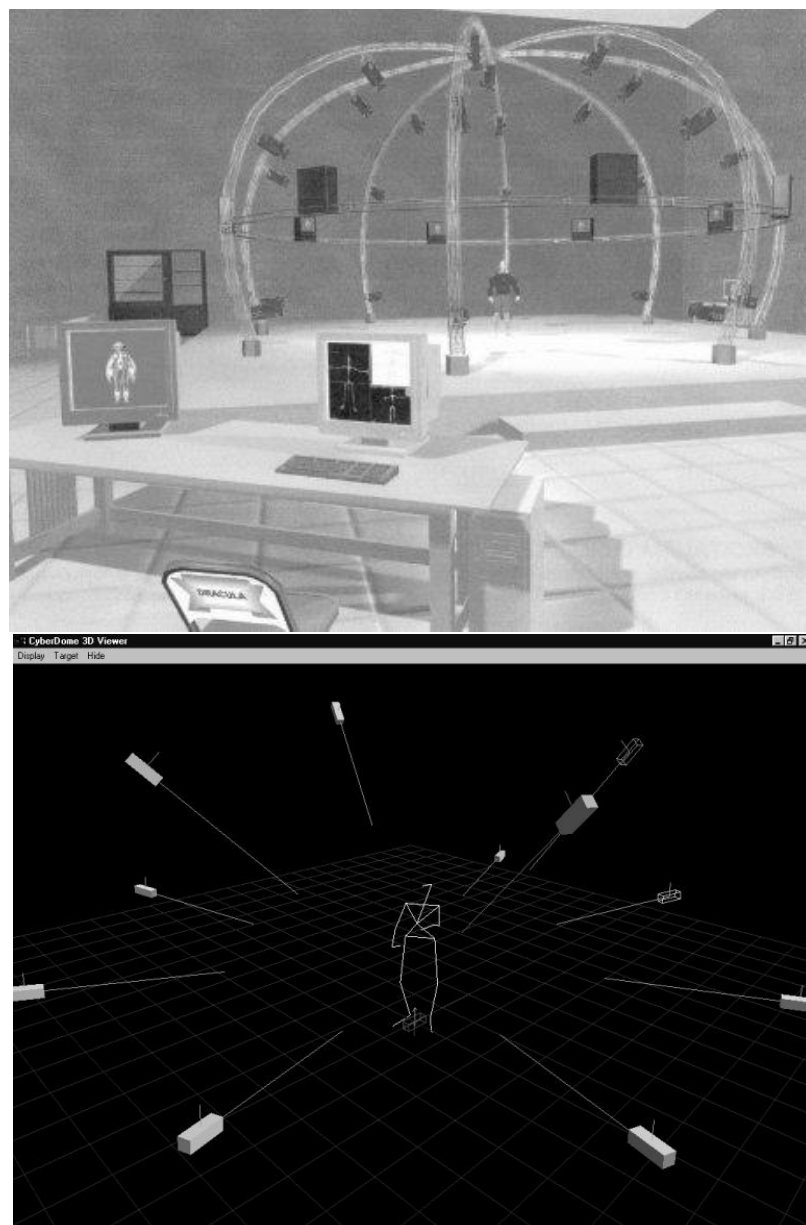


Figure 6.5 : La motion capture avec le Cyberdôme/CNBDI de XD Productions, copyright photos XD Productions



Figure 6.6: L'interface de localisation des membres supérieurs de Exos, copyright photo Exos

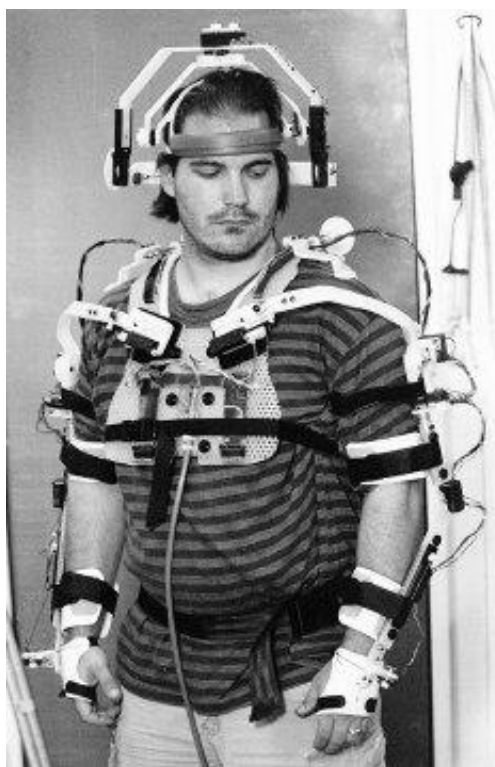


Figure 6.7 : Exosquelette Upper Body Waldo de Character Shop, copyright photo Character Shop



Figure 6.8 : Exosquelette Dual Arm Waldo de Character Shop, copyright photo Character Shop

Il n'existe pas non plus beaucoup d'interfaces spécifiques de détection des membres inférieurs. Character Shop a un produit du même type que ses autres produits. Pour analyser la position agenouillée d'une personne, un système original a été développé pour l'ETAS d'Angers pour étudier la position d'un tireur. Le repérage des contacts des membres inférieurs se fait par un tapis qui relève en 2D les pressions en chaque point (Figure 6.9).

6.4 INTERFACES SPÉCIFIQUES DE LOCOMOTION

Dans toute immersion d'une personne dans un environnement virtuel, quatre catégories de PCV (Primitives Comportementales Virtuelles) existent, dont celle du déplacement du sujet dans ce monde virtuel. La sensation de déplacement, chez l'homme, est donnée principalement par le système visuel et le système vestibulaire. On peut donc facilement donner une impression de déplacement en stimulant la vue : c'est le principe de base de bien des dispositifs d'affichage. Le système vestibulaire, grâce à l'oreille interne, apporte les informations relatives aux accélérations, vitesses angulaires, et forces d'inertie subies par la tête. En stimulant ces deux systèmes, on peut recréer une sensation plus forte de mouvement, mais dans ce cas c'est plus difficile à réaliser.

Dans un système de réalité virtuelle, en général, le sujet change de lieu virtuel sans avoir à déplacer son corps réel. Par contre, il peut virtuellement avoir différents types de déplacement, dont ceux qui sont impossibles dans un espace réel, comme le déplacement en «vol libre». Lorsque l'on s'immerge dans un espace virtuel, la pesanteur n'y est pas forcément représentée. Les déplacements de l'individu immergé sont à priori

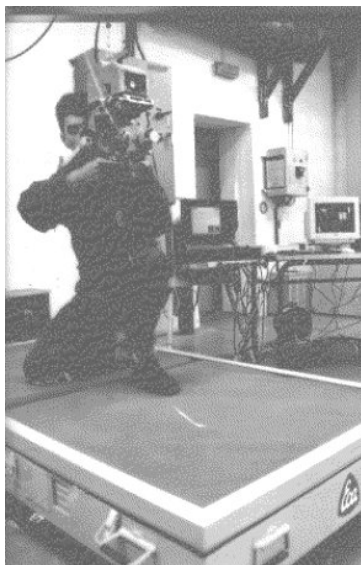


Figure 6.9 : Dispositif d'analyse de la posture d'une personne, photo J.P. Papin

libres de toute relation vis-à-vis du sol virtuel.

Nous ne développerons pas dans ce chapitre toutes les solutions possibles de déplacement virtuel, mais seulement celles qui exploitent les mouvements, approchés ou similaires, de la marche humaine. Mais le lecteur doit bien analyser cette PCV, très souvent présente dans l'application, ainsi que les niveaux I^2 mentales associées. Il n'est pas toujours obligé de faire concrètement «marcher» l'utilisateur. Des déplacements par le regard ou par des mouvements de main sont aussi à considérer et à comparer aux interfaces spécifiques de locomotion. Nous pouvons distinguer ainsi les mouvements exécutés réellement, de ceux exécutés seulement virtuellement. Le concepteur peut aussi réaliser une interface de locomotion avec, par exemple, des traqueurs classiques fixés aux pieds du sujet, ce dernier indiquant ses intentions de déplacement par des mouvements spécifiques des pieds.

La locomotion est accomplie le plus naturellement par la marche dans le monde réel. On peut penser qu'il suffit de créer une interface de marche pour résoudre le problème. En pratique, il n'est pas facile de réaliser et d'utiliser ce type d'interface, sauf si on limite le déplacement de la marche à une seule direction. Dans ce cas, un tapis roulant peut convenir en asservissant sa vitesse à la marche de l'utilisateur. Et pour permettre à celui-ci de changer de direction, on propose comme interface un tapis roulant équipé d'un guidon de direction ou d'un dispositif similaire pour les changements de direction. Les tapis expérimentaux à deux directions de marche (tapis 2D) sont par contre très contraignants dans leur utilisation. S'il existe bien des interfaces permettant de marcher, voire de courir, il n'y a pas ou peu d'interfaces pour ramper, nager, grimper ou monter des escaliers.

Vu les grandes difficultés techniques de faire une interface de marche dans deux directions (tout en restant sur place dans le monde réel), on peut souhaiter transformer

le mouvement de marche en un autre type de mouvement (roulement par patins à roulettes, par un fauteuil roulant ou par un vélo). Ces déplacements palliatifs ont le mérite de simplifier la réalisation de l'interface tout en donnant à la personne la possibilité d'activer ses jambes. Ce compromis permet une meilleure immersion sensori-motrice dans le monde virtuel que le simple déplacement virtuel commandé par une manette.

Certaines de ces interfaces motrices peuvent être analysées comme des interfaces sensori-motrices, car l'utilisateur peut ressentir sur ses jambes des efforts variables et commandables par l'ordinateur. Ceux-ci peuvent être modifiés grâce à différents dispositifs employés : inclinaison variable d'un tapis roulant, couple de contre-réaction variable sur un pédalier de vélo, etc. Mais nous présentons ces interfaces dans ce chapitre, car il s'agit d'abord pour celles-ci de détecter la locomotion et en complémentarité d'imposer des retours d'effort.

6.4.1 INTERFACES DE LOCOMOTION PAR PATINS

Les patins permettent des déplacements sur un plan virtuel infini, mais le sol réel est fixe : seuls les pieds glissent. Le mouvement relatif des pieds par rapport au sol est détecté et engendre un mouvement dans le monde virtuel qui peut être affiché sur un visiocasque. Un capteur de position donne l'orientation du marcheur. Fixée par des sangles à un portique ou se tenant à un anneau rigide l'entourant au niveau de la taille, la personne équipée de patins à roulettes peut se mouvoir virtuellement, tout en restant sur place dans le monde réel. Dans un système, développé par des chercheurs japonais [Iwata, 1994], les mouvements des pieds sont détectés par des capteurs à ultrasons suivant le principe de mesure de temps de vol (deux transmetteurs aux pieds et trois récepteurs fixes). Les données sont relevées à la fréquence de 20 Hz et avec une précision d'un centimètre. Le dispositif demande un apprentissage d'utilisation et présente un côté ludique, exploitable dans les jeux virtuels (Figure 6.10). A noter aussi l'utilisation de chausseurs avec des billes insérées dans les semelles pour réaliser un même type d'interface de locomotion.



Figure 6.10 : Interface de locomotion à patins à roulettes

6.4.2 INTERFACES DE LOCOMOTION PAR TAPIS

6.4.2.1 Tapis 1D

À l'inverse des patins, le tapis roulant fait défiler le sol réel pour compenser le mouvement de translation engendré par la marche. C'est une solution efficace pour obtenir l'interface de marche la plus naturelle qui soit. Cependant, le problème majeur rencontré avec cette solution est l'unique direction imposée par le tapis (tapis 1D), qui doit être complétée par une autre interface (volant, manette, boutons, etc). Comment, au niveau mental, la personne va t-elle assimiler ce type d'usage ?

Nous avons adapté et exploité ce type d'interface pour le projet Sofi de la SNCF, auquel nous collaborons. Nous pensons que les biais³ sensori-moteurs introduits dans le schème de la marche sur un tapis 1D sont assimilables par les utilisateurs. Dans le cas de l'exploitation d'un tapis 1D, il faut utiliser un autre schème pour permettre de tourner. Nous avons exploité le schème du changement de direction par un guidon. Nous avons fait des tests qui ont validé ce principe (Figure 6.11). Mais face à un tapis asservi, les utilisateurs ne sont pas toujours à l'aise. Pourtant, pour marcher le plus aisément sur un tapis, il suffit de ne pas se préoccuper de son fonctionnement. Mais l'appréhension de la chute en bout de tapis est assez forte pour certains utilisateurs. Dans ce cas, ils essayent d'asservir la vitesse du tapis et s'opposent au contrôleur informatique qui a le même objectif. Ceci crée des instabilités dans le mouvement du tapis [Burkhardt, 1999].



Figure 6.11 : Interface de locomotion : un tapis roulant asservi et un «guidon» pour se déplacer, photo Ecole des Mines de Paris

Dans le même ordre de solution, l'Institute for Simulation and Trainig a mis au point un système comprenant un tapis 1D avec une barre à l'avant. Celle-ci est utilisée pour donner la direction du mouvement dans le monde virtuel défilant dans un visiocasque.

Sarcos a réalisé plusieurs variantes de tapis. Le tapis le plus simple permet le déplacement dans une direction et la présence d'un élastique rattaché au marcheur crée une force de rappel pour simuler une impression d'effort. Mais celui-ci n'est pas réglable. La notion de distance dans le monde virtuel s'en trouve accrue. Une seconde génération du produit permet l'inclinaison commandable du tapis et fait gagner l'application

³Les biais sensori-moteurs principaux sont l'absence de mouvement réel détecté par le système vestibulaire et l'orientation de la trajectoire commandée manuellement et non par les jambes.

en réalisme (effets de la gravité, Figure 6.12) :

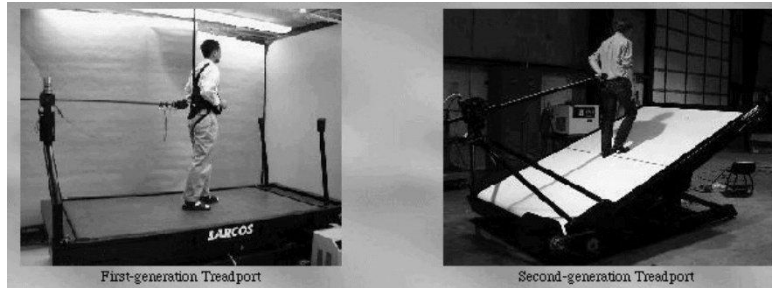


Figure 6.12 : Interface de locomotion par tapis 1D de Sarcos, copyright photo Sarcos

Ces systèmes peuvent utiliser des visiocasques, mais Sarcos a mis au point un système d'affichage à 3 écrans. Le marcheur dispose devant lui d'un panorama lui permettant d'aller dans la direction qu'il veut par l'intermédiaire d'un système de harnais (Figure 6.13).



Figure 6.13 : Tapis 1D avec écrans panoramiques de Sarcos, copyright photo Sarcos

6.4.2.2 Tapis 2D

Il est bien évident que le problème de fond des tapis 2D est un problème mécanique. Comment faire marcher une personne dans toutes les directions tout en la laissant sur place ? Certains ont pensé à exploiter une grande sphère libre de rotation par rapport à son centre, le marcheur étant soit en bas à l'intérieur ou sur la sphère à l'extérieur. Pour que cela fonctionne d'une façon satisfaisante, le rayon de la sphère doit être assez

grand. Le marcheur a alors une impression de se déplacer sur une surface plane et non sur une calotte sphérique. Cela implique une interface de très grande dimension et donc d'un coût élevé. Les autres solutions consistent dans des roulements perpendiculaires combinés entre eux mécaniquement.

Le tapis 2D Treadmil [Tsukuba, 1999] du laboratoire de Réalité Virtuelle de l'Université de Tsukuba permet à une personne de se déplacer dans toutes les directions d'un plan. Le principe utilisé est celui d'un tore composé de 12 tapis simples reliés côte à côte qui peuvent défiler suivant une direction perpendiculaire à leur axe de déroulement. Les mouvements des pieds sont mesurés par des capteurs électromagnétiques placés sur les genoux du marcheur. Le sol bouge dans le sens opposé de la marche par asservissement afin que l'utilisateur conserve en permanence la position centrale au centre du système mécanique. La distance entre le marcheur et le cercle du tapis détermine la puissance des moteurs qui augmente proportionnellement avec la distance. Le sol actif se comporte comme un ressort virtuel qui ramène continuellement le marcheur au centre du système (Figure 6.14).



Figure 6.14: Interface de locomotion par tapis 2D Treadmill, copyright photo Professor Iwata

Le plan infini est obtenu par la composition de deux translations orthogonales suivant les axes de déplacement X et Y observables sur la Figure 6.15. Les 12 tapis sont alimentés chacun par un moteur à courant alternatif de 80W. Dans la direction orthogonale, un moteur à courant alternatif de 200W entraîne l'ensemble des tapis. La vitesse de défilement des tapis est de $1,2 \text{ m.s}^{-1}$ en X et en Y. La longueur maximum de pas est de 30 cm sur une aire active de 1 m x 1 m.

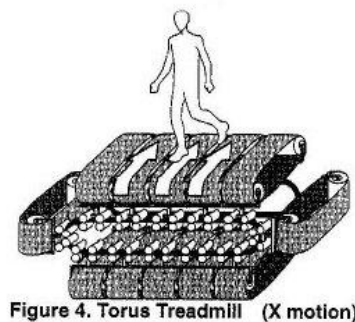


Figure 4. Torus Treadmill (X motion)

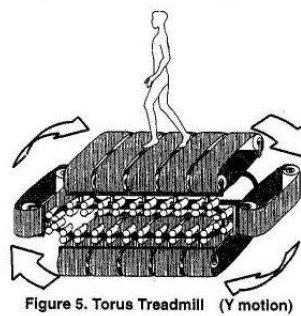


Figure 5. Torus Treadmill (Y motion)

Figure 6.15 : Schéma des mouvements sur le tapis 2D Treadmill, copyright photos Professor Iwata

6.4.3 INTERFACES DE LOCOMOTION EN POSITION ASSISE

Le VRBike de la société Sarcos est composé d'une selle et d'un pédalier ordinaires. La personne, assise sur la selle, pédale pour se déplacer dans le monde virtuel. Par rotation légère de son bassin, elle peut faire pivoter la selle et ainsi indiquer un changement de direction de sa trajectoire. Des retours d'effort peuvent être incorporés au pédalier en fonction de la difficulté virtuelle de déplacement (en côte par exemple). Ainsi, le pédalier permet de reproduire les efforts similaires à ceux d'un déplacement à pieds (Figure 6.4.3).



Figure 6.16 : Pédalier Stepper de Sarcos, copyright photo Sarcos

L'asservissement des images virtuelles à un vélo permet également un déplacement continu en environnement virtuel. Certains ont envisagé l'emploi d'un fauteuil roulant,

posé sur des rouleaux, comme interface de locomotion. Celle-ci peut être exploitée pour étudier en simulation les accès d'un handicapé en fauteuil roulant dans un bâtiment futur.

En conclusion, il n'y a pas d'interface de locomotion universelle. Tous les cas précédemment cités n'ont été réalisés que pour des applications spécifiques, en transformant le mouvement de marche en un mouvement plus techniquement réalisable (roulement, voire glissement en chaussant des skis sur une piste virtuelle).

6.5 INTERFACES DE DETECTION DE L'EXPRESSION DU VISAGE OU DES LÈVRES

Outre le mouvement de la tête, nous pouvons souhaiter aussi déterminer l'expression du visage et le mouvement des lèvres, principalement en réalité virtuelle pour animer en temps réel un clone de la personne observée. L'autre application potentielle de cette technique concerne l'animation de personnages pour la création de dessins animés [INA, 1995]. Les dispositifs adaptés à la détection de l'expression du visage ou des lèvres sont en général les mêmes systèmes. Deux grandes familles sont commercialisées : les systèmes optiques, sans contrainte d'utilisation mais pas toujours robuste et les systèmes mécaniques, très contraignant à l'usage. D'où la présentation suivante en fonction des deux principes exploités.

6.5.1 DÉTECTION OPTIQUE DE L'EXPRESSION DU VISAGE OU DES LÈVRES

La détection du mouvement des lèvres par traitement d'images peut se faire de différentes manières. En premier lieu, il faut d'abord reconnaître et positionner le visage. Quand la tête est filmée à partir d'une caméra posée sur l'écran de l'ordinateur, la première étape est d'extraire la tête du fond quelconque de l'image. Cette étape se fait maintenant sans trop de problèmes et en temps réel [Valente, 2000]. Ensuite, différentes méthodes sont envisageables. Des chercheurs préfèrent observer des pastilles collées sur la paire de lunettes que porte l'opérateur. Mais maintenant, pour simplifier l'utilisation des dispositifs, les algorithmes de traitement d'images détectent certains contours caractéristiques du visage sans poser des marques sur sa peau. Après le relevé du visage, la détection des lèvres peut se faire plus facilement si elles sont vivement colorées [INA, 1995]. Les lèvres sont maquillées en bleu (transformé en noir par un système chromakey) et sont filmées de face et de profil par deux caméras.

La simple reconnaissance des lèvres ouvertes ou fermées permet de synchroniser le mouvement du visage virtuel avec la parole. Mais la reconnaissance des expressions transmises par les lèvres est nettement plus difficile à réaliser en temps réel. Bien des recherches sont encore à poursuivre dans cette voie, ainsi que dans la reconnaissance d'expression de tout le visage.

Certains systèmes font partie intégrante d'un système de capture optique du mouvement (motion capture) (FaceTracker de X-IST Realtime Technologies, MoCap de Northern Digital), dont nous avons explicité les principes précédemment. Dans ce cas, il ne s'agit pas de retrouver les mouvements du corps et du squelette, mais d'analyser finement des petites déformations du visage pour en extraire une expression, ce qui est plus complexe à résoudre.

La société Northern Digital Inc (NDI) propose un système complet de capture de mouvement temps réel pour le visage, comprenant 256 petites pastilles. Il est capable d'acquérir jusqu'à 150 images par seconde.

Le produit FaceTracker de X-IST Realtime Technologies fait partie de l'ensemble complet proposé par la société allemande pour la capture de mouvement. Le principe de fonctionnement est basé sur une caméra infrarouge et un éclairage infrarouge, portés sur la tête de l'utilisateur. La caméra observe constamment le visage sur lequel sont placés 36 différents marqueurs autocollants de réflexion. Les mouvements des marqueurs sont suivis en temps réel. Des algorithmes, basés sur les déformations musculaires du visage, fonctionnent en temps réel sur des cartes de traitement d'images spécifiques.

SimGraphics a un système de détection du visage V-Actor pour l'animation de personnage en temps réel. Il est composé d'un « casque » avec une caméra par devant qui détecte le mouvement de petits marqueurs placés sur le visage (Figure 6.17).



Figure 6.17 : Le système V-Actor de SimGraphics, copyright photo SimGraphics

6.5.2 DÉTECTION MECANIQUE DE L'EXPRESSION DU VISAGE OU DES LÈVRES

Les méthodes mécaniques sont plus invasives. Elles sont employées pour animer des personnages de dessins animés. Le marionnettiste est équipé d'un système mécanique sur le visage, qui relève le mouvement du visage et éventuellement des lèvres. Mais ces systèmes ne sont pas très ergonomiques par rapport à l'acquisition optique. Un système d'acquisition des données du visage pour l'animation de personnages virtuels est proposé par Character Shop, en association possible de leur exosquelette Body Waldo (Figure 6.18).

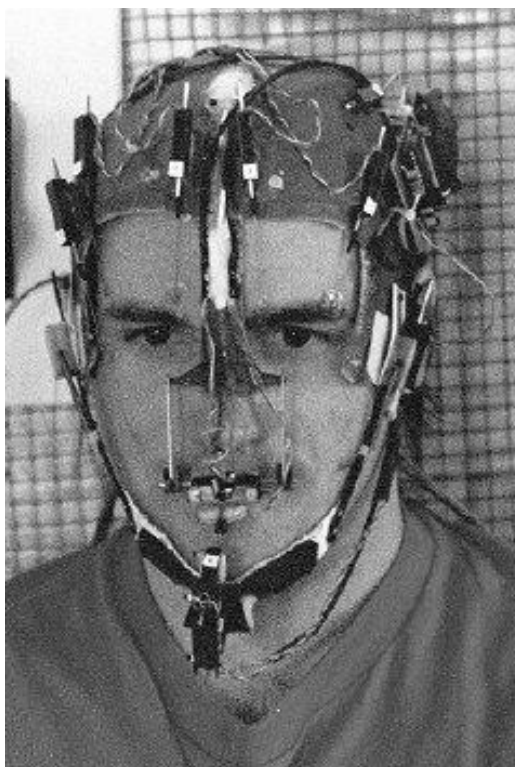


Figure 6.18 : Système de capture du visage Facial Waldo de Character Shop, copyright photo Character Shop

6.5.3 CONCLUSION

Nous pouvons dire que si différents systèmes sont commercialisés, il reste encore des travaux de recherche et de développement pour avoir des systèmes optiques entièrement transparents. A noter le cas particulier de l'entreprise belge «Neurones». Celle-ci propose une méthode originale pour animer les lèvres d'un acteur virtuel. Plutôt que de détecter les mouvements des lèvres du marionnettiste, le dispositif analyse sa voix et en déduit en temps réel ses mouvements de lèvres correspondants pour l'acteur virtuel. Le marionnettiste peut en plus choisir différentes expressions du visage pour donner une plus grande véracité de dialogue entre l'acteur virtuel et une personne.

6.6 OCULOMÈTRES

6.6.1 INTRODUCTION

Plusieurs systèmes d'oculomètres ont été développés pour l'ophtalmologie et pour l'aide aux handicapés. Les appareils sont basés sur différents principes : mesure du potentiel sur la surface de la peau près des yeux (voir la fin du chapitre sur les capteurs physiologiques), mesure électromagnétique d'une bobine fixée à une lentille et mesure optique de réflexion sur l'œil. Ces derniers systèmes, les seuls non invasifs, exploitent le traitement d'images actif (un faisceau infrarouge participe à la localisation de l'œil). Des systèmes sont utilisables à la condition restrictive que la tête soit immobile ou que le dispositif soit porté par la tête. Ces dispositifs ne sont guère transparents et il est donc difficile de les exploiter dans une application de réalité virtuelle.

6.6.2 OCULOMÈTRES OPTIQUES

Les techniques habituellement utilisées pour détecter le mouvement des yeux reposent sur la vidéo. On peut trouver des lunettes avec des miroirs placés au-dessous des yeux avec deux caméras pointant sur les miroirs. Les caméras déduisent le mouvement des yeux en fonction du déplacement du reflet de l'iris dans le miroir. Les problèmes liés à la détection du mouvement des yeux sont le temps de traitement et d'acquisition des informations (latence), les erreurs d'interprétation. La précision est de quelques minutes d'arc. Le dispositif utilisé implique un encombrement important.

Le système Eyeputer du LETI mesure les mouvements d'un œil par traitement d'images en temps réel (jusqu'à 480 images par secondes). La position de l'œil est déduite du traitement de l'image qui extrait la pupille et le reflet d'une source infrarouge dirigée sur la cornée. La précision obtenue est de $0,5^\circ$ [Charbonnier, 1994].

Le système Eygaze System de LC Technologies est utilisé en souris que l'on guide avec les yeux. La solution qu'ils ont adoptée est une caméra située sous l'écran qui émet une lumière de faible intensité et peut suivre son reflet sur l'œil.

L'oculomètre de SensoMotoric Instruments mesure les mouvements d'un œil ou des deux yeux. Il fonctionne avec une caméra vidéo à haute résolution et un éclairage infrarouge (Figure 6.19). SensoMotoric Instruments a développé un oculomètre pour une utilisation en souris 2D (Figure 6.20).

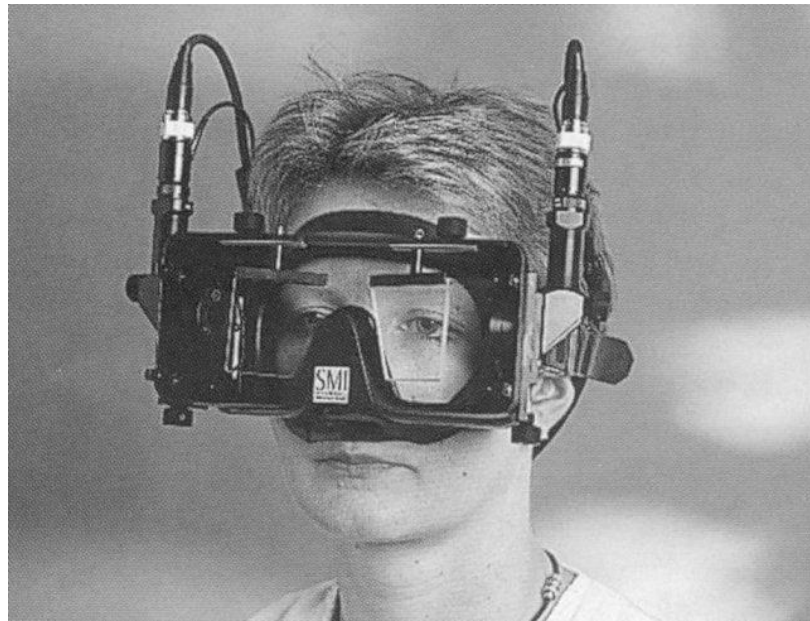


Figure 6.19 : L'oculomètre de SensoMotoric Instruments, copyright photo SensoMotoric Instruments



Figure 6.20 : L'oculomètre de SensoMotoric Instruments, copyright photo SensoMotoric Instruments

Le NAC eye marker recorder est un matériel largement utilisé pour l'analyse des stratégies visuelles (INRIA, INSEP, CERMA, DGA, INRS, etc.) [Papin, 1987].

6.6.3 OCULOMÈTRES ÉLECTROMAGNÉTIQUES

La détection des mouvements des yeux se fait par un système de lentilles de contact où des spires de cuivre sont incrustées. Le dispositif mesure des courants induits dans un environnement électromagnétique.

6.7 CAPTEURS PHYSIOLOGIQUES

Des laboratoires proposent des dispositifs mesurant de faibles différences de potentiel à la surface de la peau. Ces biocapteurs permettent de détecter directement les signaux biologiques commandant le mouvement des muscles oculaires ou squelettiques. Ils peuvent aussi relever les activités du cœur ou du cerveau, mais dans ce cas ils n'ont pas a priori d'utilité pour des applications de Réalité Virtuelle.

L'entreprise américaine BioControl Systems conçoit des biocapteurs permettant de transmettre leurs mesures à un ordinateur. Ces capteurs détectent quatre sources d'activités électriques, révélatrices des activités nerveuses efférentes : l'activité musculaire (EMG, électromyogramme), l'activité du mouvement d'un œil (EOG, électrooculogramme), l'activité du cerveau (EEG, électroencéphalogramme) ou l'activité du cœur (EKG, électrocardiogramme). Ces signaux de faibles voltages sont captés par des électrodes non invasives placées sur la peau. Ils sont ensuite reconnus et traités pour en extraire des caractéristiques pertinentes par une bibliothèque d'algorithmes programmables. En troisième étape, les caractéristiques relevées sont codées et transmises à l'ordinateur via une liaison RS232. Le système BioMuse peut acquérir huit signaux simultanément à fréquence élevée (20 KHz), à résolution sur 12 bits et les exploiter en temps réel. Le prix de base est d'environ 20000\$.

Les électromyogrammes peuvent servir à capter les mouvements de la main en analysant les signaux d'un groupe de muscles. Les senseurs sont placés de telle façon qu'ils laissent la main libre de tout dispositif (Figure 6.21). Dans ce cas, ils peuvent être considérés comme une interface «gant de données». Mais on peut aussi les utiliser comme oculomètre. Les électrooculogrammes sont capables de mesurer les déplacements des yeux dans leurs orbites. Les caractéristiques recueillies permettent, par exemple, de commander un curseur affiché sur un écran pour être exploité comme une souris. En mesurant sur les deux yeux leur convergence, on peut aussi diriger le curseur dans la direction de la profondeur de l'écran. On obtient dans ce cas une souris tridimensionnelle.

6.8 CONCLUSION

Des interfaces spécifiques de localisation corporelle sont disponibles commercialement, même s'il n'y a pas beaucoup de choix, sauf dans le domaine du «motion capture» où le marché des jeux et des animations est florissant. La gêne occasionnée par le port de combinaisons ou de dispositifs mécaniques est toujours un facteur limitant. Les interfaces de localisation sur un principe optique sont encore à développer pour



Figure 6.21 : Dispositifs électrophysiologiques Biomuse, copyright photo Moto Development Group

améliorer leur robustesse et leur facilité d'exploitation. Le prix en général élevé du matériel ne permet pas une ouverture vers le grand public avec des applications exploitant à grande échelle la localisation corporelle.

6.9 RÉFÉRENCES BIBLIOGRAPHIQUES

[Burkhardt, 1999]	J.M. Burkhardt, D. Lourdeaux, P. Fuchs, (1999) « <i>Conception d'un système de RV pour la formation des agents de conduite aux opérations en milieu ferroviaire</i> », Actes des journées réalité virtuelle et cognition, 14-15 Décembre, ENST, Paris, 123-132.
[Charbonnier, 1994]	C. Charbonnier, D. Masse, « <i>écriture par commande visuelle</i> », p185-191, congrès «Interface des mondes réels et virtuels», Montpellier, 1994.
[INA, 1995]	Projet Télévirtualité, Direction de la Recherche, INA, 94360 Bry sur Marne.
[Iwata, 1994]	H. Iwata, K. Matsuda, haptic walkthrough simulator, institute of engineering mechanics, Tsukuba, 305 Japan, 1994.
[Papin, 1987]	J.P. Papin, " <i>Utilisation des enregistrements du regard dans l'apprentissage de la prise d'information</i> ", Thèse de doctorat d'état es lettres, Paris V, 119 P, juin 1987.
[Tsukuba, 1999]	Laboratoire de réalité virtuelle de l'université de Tsukuba, http://intron.kz.tsukuba.ac.jp . « <i>Walking about virtual environment on an infinite floor</i> », Hiroo Iwata, IEEE99, 1999.
[Valente, 2000]	S. Valente and J.-L. Dugelay. " <i>Face tracking and realistic animations for telecommunicant clones</i> ". IEEE Multimedia, IEEE Computer Society, Spring 2000.

Adresses Web des entreprises :

Capture de mouvement

Principe optique

Actisystem	http://www.actisystem.fr
Character Shop	http://www.Character-Shop.com
Charnwood Dynamics	http://www.charndyn.com
Qualitys	http://www.qualisys.com
Mikromak	http://www.mikromak.com
Northern Digital	http://www.ndigital.com
Motion Analysis	http://www.motionanalysis.com
Motek	http://www.motek.org
Vicon	http://www.vicon.com
XD Productions	http://www.xdprod.com

Principe électromagnétique

Ascension Technology Corp.	http://www.ascension-tech.com
Polhemus	http://www.polhemus.com
Simulation	http://www.simulationfx.com/mocap.html

Principe mécanique

Capture du mouvement des levres	http://herens.idiap.ch/~luettin
Character Shop	http://www.Character-Shop.com
INA	http://www.ina.fr
Northern Digital Inc.	http://www.ndigital.com
SimGraphics	http://www.simg.com
X-IST	http://www.x-ist.de

Oculomètres

SensoMotoric Instruments	http://www.smi.de
LC Technologies Inc.	http://www.eyegaze.com

Capteurs physiologiques

Moto Development Group	http://www.moto.com

7 LES INTERFACES MANUELLES MOTRICES

Philippe Fuchs

7.1 INTRODUCTION

7.1.1 CAPTEUR DE LOCALISATION ET GANT DE DONNÉES

Nous abordons dans ce chapitre les interfaces motrices spécifiquement dédiées à la main. Dans le monde réel, la préhension d'un objet par la main est un phénomène mettant en jeu bien des processus complexes, au niveau sensori-moteur et au niveau mental, même si cela est inconscient. L'interface motrice qui semble la plus appropriée en théorie est le gant de données, capable de mesurer les mouvements de la main et de ses doigts. Combien de concepteurs se sont-ils précipités sans réflexion sur l'exploitation de gant de données pour manipuler des objets virtuels ? La manipulation est l'une des quatre catégories des PCV (Primitives Comportementales Virtuelles), autant dire que le problème du choix de l'interface de manipulation est fréquent. Si nous réfléchissons à cette question, nous pouvons observer que souvent il s'agit de pouvoir déplacer aisément en translation et en rotation l'objet (de 2 à 6 ddl). La préhension effective de l'objet est souvent simpliste : l'objet est pris ou l'objet est lâché. Dans ces cas *fréquents*, il n'est pas nécessaire d'utiliser un gant de données, un simple traqueur suffit pour indiquer les mouvements de l'objet et sa prise et son lâcher se font par une simple commande binaire. Cet ordre peut être donné par un bouton ou par le simple rapprochement visuel à l'écran entre un curseur virtuel et l'objet à manipuler. Nous profitons ainsi de la simplicité des lois du monde virtuel, comparativement au monde réel. Ceci explique que peu d'applications ont vraiment besoin de gant de données et que leur développement commercial est relativement lent. Le concepteur d'application de réalité virtuelle doit donc toujours se poser la question : l'utilisation d'un gant de données est-elle nécessaire ? Un traqueur ne suffit-il pas ?

7.1.2 CAPTEUR DE LOCALISATION ET INTERFACE DE COMMANDE

Mais il ne s'agit pas toujours de manipuler un objet dans un environnement virtuel. Nous avons indiqué dans le chapitre sur les traqueurs que ceux-ci peuvent servir :

- soit à manipuler un objet virtuel ;
- soit à diriger un curseur dans un environnement virtuel.

Dans le premier cas, l'utilisateur agit «*naturellement*» sur le monde virtuel, le traqueur se contentant de suivre le mouvement de sa main. Dans le deuxième cas, il s'agit en général de commander une action spécifique : par exemple, se déplacer dans le monde virtuel en indiquant une direction, pointer sur un menu déroulant, commander symboliquement l'ouverture d'une porte, etc. Dans ce deuxième cas, les traqueurs peuvent être remplacés par *des interfaces de commande* classiques (joysticks, souris 3D, etc.). Celles-ci, posées en général sur une table, permettent à l'utilisateur de commander manuellement le déplacement d'un curseur grâce aux mouvements relatifs de l'interface

par rapport à une position neutre. Ces mouvements ont de faibles courses (quelques millimètres à quelques centimètres). Il ne s'agit donc pas de localisation dans l'espace de la main pour ces interfaces de commande et elles sont basées évidemment sur d'autres principes que les traqueurs. A noter qu'il s'agit dans le cas d'un traqueur de mouvoir *naturellement* un curseur, donc sans grand effort intellectuel. A la différence des souris 3D et joysticks avec lesquels le curseur est mû grâce à une action manuelle, réfléchie et peu naturelle de l'utilisateur, sauf pour les habitués de ce type d'interface. Si un traqueur peut donc être utilisé en curseur 3D, les joysticks et les souris 3D ne peuvent être employés en capteur de localisation de la main de l'utilisateur.

Le chapitre est donc découpé en deux grandes parties, puisque les traqueurs ont déjà été présentés : en premier, les gants de données et ensuite les différentes interfaces de commande. Ces dernières sont présentées dans ce chapitre, mais elles sont donc à comparer aux traqueurs, dans le cas d'actions de commande dans le monde virtuel.

7.2 GANTS DE DONNÉES

Les gants de données (ou «gants numériques»), permettent de mesurer les mouvements relatifs des doigts par rapport à la main. Pour déterminer complètement les mouvements absolus, il faut aussi repérer la position de la main avec un traqueur. Les types de traqueur employés pour cette fonction sont principalement les capteurs électromagnétiques et parfois les capteurs à ultrasons. La difficulté principale est de mesurer les rotations des phalanges des doigts avec des dispositifs légers et peu encombrants. Plusieurs principes sont en concurrence : mesure par déformation de fibres optiques, par effet Hall ou par variation de résistance. Les problèmes techniques importants rencontrés sur ces dispositifs sont la répétabilité, la précision et la fiabilité des dispositifs.

7.2.1 GANTS À FIBRES OPTIQUES

Ce fut le premier principe utilisé pour un gant, qui fut conçu par la société américaine VPL, pionnière dans le domaine de la réalité virtuelle : le fameux DataGlove à fibre optique, inventé par Thomas Zimmerman en 1987, qui n'est plus commercialisé. A chaque doigt du gant est fixé deux (ou trois) fibres optiques. Celles-ci sont dans l'alignement des doigts et plus ou moins longues pour détecter les mouvements aux différentes articulations [VPL, 1987] (Figure 7.1).

Les fibres optiques sont reliées à un dispositif optoélectronique. Une des extrémités de chaque boucle optique est connectée à une LED et l'autre extrémité à un phototransistor. La fibre étant droite, il n'y a pas diminution de la lumière transmise. Par contre, si la fibre est pliée, l'intensité lumineuse reçue est atténuée, car on n'est plus dans les conditions de réflexion parfaite et la lumière recueillie est atténuée. La fibre optique est traitée à la pliure pour augmenter la perte de lumière quand le câble est plié. La variation d'intensité de la lumière réfléchie par la fibre est fonction de l'angle articulaire, ce qui permet d'en déduire sa valeur. La relation entre l'intensité et l'angle n'est pas linéaire. Elle n'est pas non plus constante dans le temps, surtout si le gant est employé par des personnes ayant des morphologies de la main différentes. La calibration du gant est donc indispensable et doit être faite pour chaque nouvelle utilisation. La précision des angles aux articulations est de l'ordre de 5 à 10 degrés. Ce gant n'est plus fabriqué, mais le partenaire japonais de VPL, Nissho Electronics Corporation, a conçu un nouveau gant, «Super Glove», sur un principe différent breveté avec dix senseurs

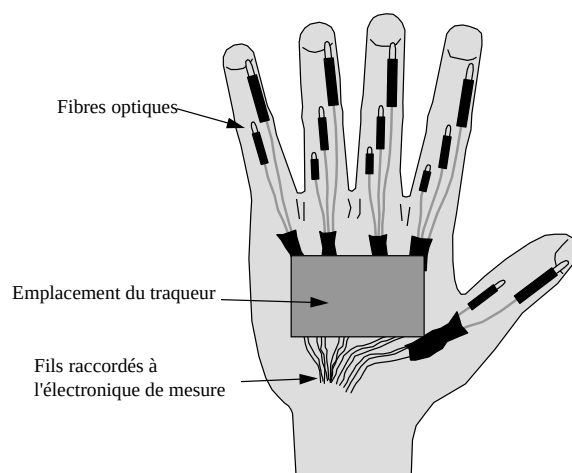


Figure 7.1 : Gant de données à fibres optiques

ayant une résolution de $0,3^\circ$ et une vitesse de lecture pour chaque senseur de 0,2 ms (Figure 7.2). Une version « sans fil » existe pour ce gant.

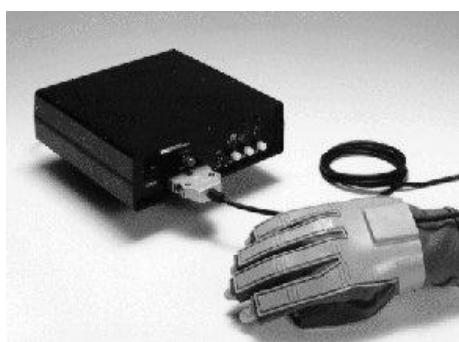


Figure 7.2 : Le gant de données SuperGlove de Nissho Electronics Corporation, copyright photo Nissho Electronics Corporation

Au lieu de détecter des flexions de fibres, on peut mesurer les translations relatives des fibres, une des deux extrémités glissant dans un manchon équipé d'une diode LED et recouvert sur sa paroi intérieure d'une surface réfléchissante. Selon l'angle de l'articulation du doigt, l'extrémité de la fibre s'éloigne plus ou moins de la LED et le phototransistor à l'autre bout de la fibre recueille une intensité lumineuse proportionnelle au déplacement relatif de la fibre dans le manchon, donc à l'angle de flexion [Wright, 1990] (Figure 7.3).

L'avantage de ces interfaces réside dans leur faible volume et leur légèreté, ceci procurant un certain confort. Des entreprises proposent actuellement des gants de données à fibres optiques à prix modique pour les ordinateurs compatibles PC. Un gant peut posséder un senseur par doigt, ayant une résolution sur huit bits, et les mesures sont transmises 25 fois par seconde.

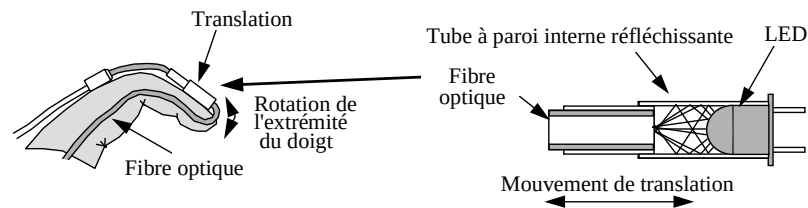


Figure 7.3 : Mesure de la déformation d'un doigt par translation d'une fibre optique

Les gants de la société Fifth Dimension possèdent l'avantage d'être abordables en terme de prix (environ 500 \$ US). Le principe de fonctionnement repose sur une technologie fibre optique (pour la flexion des doigts) couplée à 2 inclinomètres relatifs (série 5 uniquement). Le gant en lycra permet une grande souplesse. L'évolution de ces produits apporte une version sans fil (wireless). La version 5DT DataGlove 16/16W a deux capteurs optiques par doigt plus pour l'abduction, ayant deux inclinomètres pour le roulis et le tangage. La résolution pour la déformation des doigts est faible : sur 8 bits. La version sans fil par liaison radio a une portée jusqu'à 20 m (Figure 7.4).

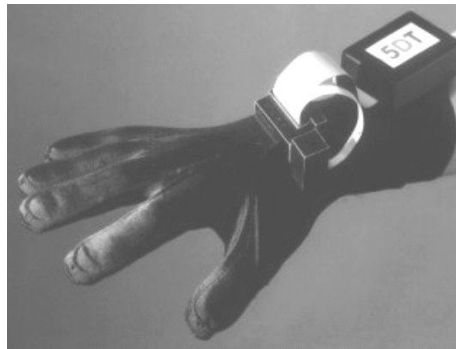


Figure 7.4 : Le gant de données optique 5DT DataGlove, copyright photo SimTeam

Essential Reality vient de développer un gant de données le «P5», à bas prix et aux caractéristiques intéressantes : les mesures des déformations des doigts et de la position de la main se font à 60 Hz. La précision du tracking est de 1° pour les rotations et de 13 mm pour les translations. L'étendue de mesure est de 3 à 4 pieds. Une version sans câble est prévue.

7.2.2 GANTS À VARIATION DE RÉSISTANCES

La variation d'une résistance, en fonction de l'allongement du support sur lequel elle est fixée, est exploitée pour des gants de données. Ce principe a été employé pour des gants bon marché et pour des gants élaborés utilisant des jauges de contraintes.

Le gant à bas prix se compose d'une couche d'encre conductrice contenant des particules de carbone. Quand le matériau se plie, la distance entre les particules de carbone

s'accroît et la résistivité du capteur augmente. Grâce à une procédure de calibration, on peut déterminer approximativement les mouvements des doigts. Ce système était employé dans le PowerGlove, mais la production de ce gant est arrêtée.

Les gants à jauges de contraintes font appel à des jauges minces englobées dans un matériau élastique souple (à base de Nylon et de Lycra). Ce type de gant est nettement plus performant, mais aussi plus onéreux. Les mesures sont codées sur huit bits et sont transmises à la fréquence de 100 Hz. La résolution est de l'ordre du degré (Figure 7.5).



Figure 7.5: Gant CyberGlove de Virtual Technologies, copyright photo Immersion S.A.

La société Virtual Technologies propose le CyberGlove avec plusieurs options (retour tactile CyberTouch et retour d'effort CyberGrasp, qui sont exposés dans le chapitre sur les interfaces manuelles sensori-motrices). Le gant est constitué d'un mélange Nylon (80%) / lycra (20%) qui le rend très flexible. Les extrémités des doigts ne sont pas recouvertes par le gant, ce qui laisse l'utilisateur libre de taper sur un clavier, par exemple, avec la même dextérité que s'il n'avait pas de gant. Le principe de fonctionnement du

CyberGlove repose sur 18 ou 22 capteurs de mesure angulaire disposés dans le gant (en particulier trois par doigt et l'abduction du pouce). Il n'y a pas de version sans câble proposée. Les caractéristiques sont les suivantes : la fréquence de mesure est de 150 Hz (sans filtre) et de 112 Hz avec filtre pour 18 capteurs ; le temps de latence est de 5 ms pour 18 capteurs ; le poids est de 85 g.

7.2.3 GANTS À EFFET HALL

Le Dextrous Hand Master de l'entreprise Sarcos a une structure métallique (exosquelette) portée sur le dos de la main. Il a quatre capteurs de position par doigt, donc vingt pour l'ensemble du gant. Chaque rotation d'une phalange est mesurée par un capteur à effet Hall placé à l'articulation mécanique du gant. La résolution (12 bits) des mesures est donnée au dixième de degré. Ces caractéristiques sont liées aux rotations des articulations de la structure mécanique, mais pas à celles des phalanges. Les mesures sont lues à la fréquence de 75 Hz. Une adaptation aux différentes tailles de la main de l'utilisateur est prévue. Le poids (400 gr environ) et l'encombrement sont des inconvénients importants de ce type d'interface (Figure 7.6).

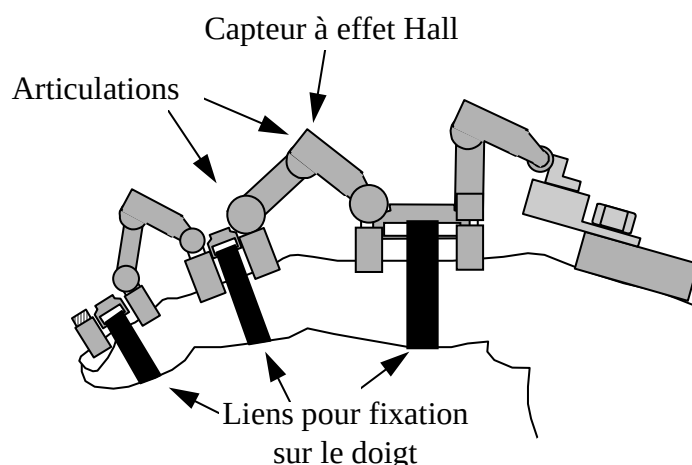


Figure 7.6 : Gant de données à effet Hall

7.2.4 CAS PARTICULIER : GANT DE COMMANDES BINAIRES

Les gants que nous venons d'exposer permettent de mesurer les déformations des doigts. Ce n'est pas le cas de la paire de gant «Pinch Glove», qui elle se «contente» de proposer à l'utilisateur d'envoyer une série de *commandes binaires*, associées à la position de la main. Car, comme pour les autres gants de données, il y a un traqueur sur chaque gant. Le principe de fonctionnement de ce gant repose sur le phénomène suivant : le contact entre deux ou plusieurs doigts crée une boucle conductrice (circuit électrique fermé), permettant une série de gestes virtuels de pincement (pinch) qui peuvent être utilisés dans les applications (Figure 7.7). Dès que le contact électrique est validé, les informations sont envoyées au PC de contrôle. Plus de 1000 gestes sont théoriquement commandables, mais il faut que l'utilisateur les connaisse !



Figure 7.7 : Le gant de commandes binaires Pinch Glove de Fakespace Labs, copyright photos Fakespace Labs

7.2.5 CONCLUSION

Nous venons de voir les trois principales catégories de gants de données, mais n'oublions pas qu'il est toujours envisageable par traitement d'images de suivre les mouvements des mains et des doigts. Des chercheurs ont entrepris de détecter les mouvements de main de personnes dans un lieu public. Celles-ci, pour leur information, doivent indiquer des menus affichés en pointant avec la main sur un écran géant. Dans ce cas, il n'est pas souhaitable de demander aux personnes de s'équiper d'un gant pour seulement pointer sur l'écran. La solution par traitement d'images (10 par seconde) est justifiée et permet une première approche de commande naturelle et transparente, malgré les difficultés algorithmiques. Les deux caméras sont judicieusement placées, au-dessus et de côté, permettant de localiser le doigt pointeur et sa direction. D'autres chercheurs détectent les contours de la main, permettant son suivi et la reconnaissance de la disposition de ses doigts. Ces travaux devraient permettre une communication gestuelle entre l'homme et la machine [Heap, 1995].

En conclusion, il n'existe pas encore un grand choix de gants de données. Très peu d'entreprises se sont lancées dans la commercialisation de ces gants, comparativement aux autres produits classiques de la réalité virtuelle que sont les visiocasques et les traqueurs. L'étendue modeste du marché doit freiner le développement de ce produit, plus que les problèmes techniques, le principal étant de trouver un principe de mesure permettant d'assurer la stabilité dans le temps des mesures et la fiabilité des résultats. Ces questions techniques semblent a priori moins complexes à résoudre pour les gants de données que pour d'autres interfaces comportementales.

7.3 INTERFACES DE COMMANDE

Il est possible de rencontrer en réalité virtuelle tous les dispositifs de commande classique, utilisés pour envoyer des ordres au monde virtuel ou au système qui le gère. Nous employons le terme d'*interface (manuelle¹) de commande* pour les désigner. Par rapport à nos définitions, nous ne les incluons pas dans les *Interfaces Comportementales*, développées pour la réalité virtuelle, car elles n'exploitent pas a priori un *comportement sensori-moteur naturel* de l'homme. C'est plutôt des ordres symboliques qui sont envoyés. Parmi ces interfaces de commande, il y a les souris 2D, les souris 3D, les stylets, les manettes (ou joystick), etc. Certaines de ces interfaces de commande ont été spécialement développées pour les environnements virtuels. C'est le cas des souris 3D, utilisées d'abord en CAO pour manipuler en environnement virtuel. Certaines des interfaces de commande ont été améliorées par des systèmes à retour tactile ou d'effort. Nous en parlerons dans le chapitre correspondant. La plupart de ces interfaces sont très simples et très connues. Nous ne ferons donc pas un développement plus long à leur sujet, sauf pour les souris 3D.

7.3.1 SOURIS 3D

Nous pouvons définir une souris 3D comme étant une interface permettant à l'utilisateur de *commander* en général trois translations perpendiculaires et trois rotations perpendiculaires dans l'espace. Comme nous l'avons déjà indiqué, rappelons qu'un

¹Nous présentons succinctement à la fin du chapitre quelques interfaces non-manuelles de commande

traqueur peut faire la même fonction, mais une souris 3D ne peut relever la localisation absolue de la main. A l'avantage des souris 3D, il est plus facile techniquement de les réaliser en jouant sur les mouvements volontaires de la main. La conséquence en est généralement un prix d'achat relativement faible par rapport aux capteurs de localisation.

Rappelons que s'il n'est pas trop difficile de translater tout objet virtuel (curseur pièce, etc.) dans n'importe quelle direction à partir des trois commandes de translation, l'orientation du curseur est délicate. Car les rotations suivant trois axes perpendiculaires ne sont pas des opérations géométriques commutatives. L'orientation d'un objet virtuel dans une direction souhaitée (et non orientation au hasard) ne peut être réalisée que par des habitués de ce type d'interface, tels que les ingénieurs travaillant quotidiennement sur les stations de travail de CAO.

Les souris 3D peuvent être basées sur le principe d'une manette 2D (manche à balai) à laquelle on peut appliquer une force verticale complémentaire pour la troisième dimension. Dans ce cas, on ne maîtrise que les trois degrés de liberté en translation. Un autre principe voisin est basé sur la manipulation d'une boule sur laquelle on applique manuellement des forces et des couples qui sont mesurés. Ces efforts variables commandent proportionnellement en général les vitesses de rotation et de translation de l'objet affiché à l'écran. A noter que la boule se déplace peu par rapport au socle auquel elle est rattachée, ce qui permet à l'opérateur de garder sa main pratiquement immobile. Ce type d'interface mesure des efforts, mais seulement pour une commande de mouvement, il n'est donc pas employé comme capteur d'effort.

Le principe technique est basé sur une boule dotée de capteurs (trois pour les forces et trois pour les couples) mesurant les efforts de la main de l'opérateur sur un élément peu déformable. Les forces et les couples sont enregistrés indirectement par des capteurs optiques qui mesurent des déplacements faibles dus aux déformations. Ces dispositifs souffrent du couplage des efforts. Il est difficile d'appliquer des forces sans des couples et inversement. On peut néanmoins par logiciel ne commander que des translations ou des rotations de l'objet (Figure 7.8).

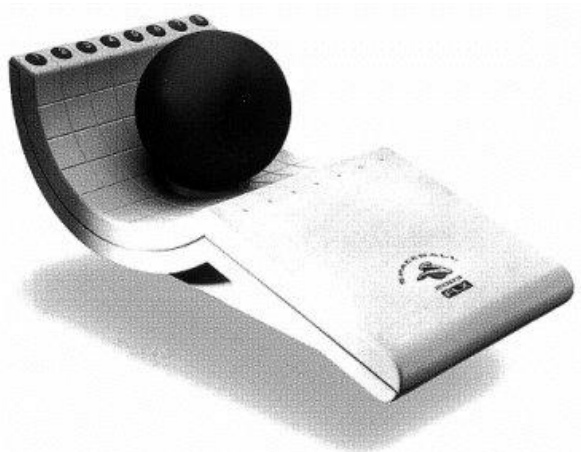


Figure 7.8 : Souris 3D SpaceBall, copyright photo Immersion S.A.

La souris « Space mouse » de LogiCad 3D fonctionne à partir d'un système optoélectronique de mesure des mouvements. La force maximum applicable est de 4,5 N et le couple maximum est de 0,1 N.m. La résolution est sur 600 incréments (Figure 7.9).



Figure 7.9 : La souris 3D Space mouse de 3Dconnexion, copyright photo 3Dconnexion

A noter qu'un constructeur propose une souris 3D exploitant le même principe que les capteurs de localisation à ultrasons. Elle est composée d'une base fixe avec trois émetteurs en triangle et d'une poignée mobile avec trois récepteurs en position identique, de petite taille (quelques centimètres). L'opérateur peut déplacer la main dans l'espace de mesure, même si ce dernier est limité à une zone proche de la base (utilisation sur un bureau). La frontière est étroite dans ce cas entre capteur de localisation et souris 3D. Toutes les souris 3D possèdent évidemment plusieurs boutons de commande. Pour toutes les souris 3D, l'opérateur ne peut pas oublier totalement la manipulation de ce type d'interface au cours de son utilisation, d'où un manque de transparence préjudiciable à toute immersion naturelle dans un monde virtuel.

7.3.2 INTERFACES NON-MANUELLES DE COMMANDE

Au lieu de donner des ordres de commande à la main, nous pouvons le faire de plusieurs façons. La commande par la voix est possible, surtout pour donner des ordres binaires. Il y a bien sûr les interfaces «pédestres» de commande, comme les pédales de véhicule. On peut toujours mettre des commandes au pied, même s'il ne s'agit pas de véhicule. Les interfaces sont simples à concevoir pour donner un simple ordre progressif.

Une idée plus originale a été développée pour faire des commandes avec la tête de l'utilisateur face à un écran. Les webcams devenant courantes sur les ordinateurs PC, l'entreprise Cybernet Systems les utilisent pour faire une commande par la tête. Par

traitement d'images, l'ordinateur détecte les mouvements latéraux (horizontaux et verticaux) de la tête. Ces commandes servent, par exemple, à déplacer le point de vue de la scène du jeu vidéo. C'est sûrement ludique, par contre une certaine fatigue doit apparaître après un certain temps.

7.4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Heap, 1995] T. Heap, «*real-time hand tracking and gesture recognition using smart snakes*», p261-271, congrès «Interface des mondes réels et virtuels», AJDMM, Montpellier, 1995.
- [VPL, 1987] VPL, «*Dataglove Model 2 users manual*», Reedwood City, CA 94063, 1987.
- [Wright, 1990] S. Wright, «*Optical attenuator movement detection system*», US patent N° 4972074, 1990.

Adresses Web des entreprises :

Gant de données

Ascension	http://www.ascension-tech.com
Essential Reality	http://www.essentialreality.com
Fifth Dimension	http://www.5dt.com
Fakespace Labs	http://www.fakespacelabs.com
Nissho Electronic Corp	http://www.tradepia.or.jp/nevc/advanced/vr
Sarcos	http://www.sarcos.com
Virtual Technologies	http://www.virtex.com

Commande avec la tête

Cybernet	http://www.cybernet.com

Souris 3D

3Dconnexion	http://www.3dconnexion.com
Logitech	http://www.vrdepot.com/vrteclg.htm
Spacetec	http://www.immersion.com

Quatrième partie

Les interfaces sensori-motrices

8 LES INTERFACES MANUELLES SENSORIMOTRICES, INTERFACES À RETOUR D'EFFORT

Philippe Fuchs

8.1 INTRODUCTION

Si on désire immerger le corps, totalement ou partiellement, dans un monde virtuel, il faut être capable au minimum de détecter les mouvements du corps, grâce aux interfaces spécifiques de localisation corporelle. Mais avec ces seules interfaces, la personne ne ressent rien en retour et perçoit un monde virtuel impalpable. Pour que l'utilisateur ait des sensations mécaniques en retour de ses actions, il faut exploiter les interfaces à simulation de mouvement, les interfaces à sensibilité cutanée (à retour tactile et thermique) ou les interfaces à retour d'effort. Les interfaces à sensibilité cutanée et à simulation de mouvement sont des interfaces purement sensorielles : elles ne transmettent des stimuli que vers l'utilisateur. Par contre, les interfaces à retour d'effort ont pour objectif d'appliquer des forces sur le corps mais aussi, toujours en pratique, elles mesurent la position de la partie du corps sur laquelle elles agissent. Cela principalement pour deux raisons :

- parce que les retours d'effort ne sont en général que la conséquence de l'action (du mouvement) de la main de l'utilisateur. Il faut donc fondamentalement connaître cette action, d'où les mesures de position, faites sur la même interface ;
- parce que les efforts à appliquer sont fonction, entre autres, de la position de la main, qu'il faut donc connaître.

Les interfaces à retour d'effort sont donc des *interfaces sensori-motrices* en pratique¹. La partie du corps concernée par ces interfaces est très souvent la main ou l'ensemble du membre supérieur, à part des interfaces à retour d'effort pédestres pour la simulation de pédales de véhicule, pédalier de vélo, etc. Il existe par exemple le VRBike : l'utilisateur en pédalant avance dans un environnement virtuel et reçoit en retour des efforts par l'intermédiaire du pédalier. Nous avons donc classé ces interfaces à retour d'effort dans l'ensemble des interfaces manuelles : les interfaces à retour d'effort sont donc en général des *interfaces manuelles sensori-motrices*. A noter que nous n'utilisons pas dans le livre le terme d'interface haptique qui regroupe les interfaces manuelles tactiles et à retour d'effort. Certaines interfaces à retour d'effort agissent sur la main et aussi sur le bras et l'avant bras. Ceci montre bien les grandes difficultés et les limites de ce type d'interface. Lorsqu'il s'agit d'appliquer des efforts sur la main par l'intermédiaire d'un outil, les problèmes techniques ne sont pas trop complexes et il existe commercialement des interfaces exploitables. Par contre, quand les retours d'effort doivent être simulés sans l'intermédiaire d'un outil, l'interfaçage est très délicat.

Il est à noter aussi que c'est l'interface qui fournit techniquement des efforts sur la main de l'utilisateur. Mais, suivant le principe mécanique d'action et de contre réaction, l'utilisateur doit fournir des efforts, s'il ne veut pas, par exemple, que l'interface

¹Nous ne connaissons pas d'interface à retour d'effort qui ne mesurerait pas en même temps la position de la zone d'action des efforts. Si vous en connaissez, informez-nous !

se déplace : en exemple, avec un gant à retour, on peut simuler à l'intérieur de la main les forces dues à une balle virtuelle qui se gonfle au creux de la main (forces du système sur l'homme). Inversement, on peut simuler la résistance de cette balle virtuelle à son écrasement par la main (forces de l'homme sur le système). Mais les efforts de l'homme sur l'interface ne sont pas directement mesurés. Car il existe très peu de *capteurs d'effort spécifiques* en Réalité Virtuelle, des interfaces qui seraient «motrices», inverses des retours d'effort, interfaces «sensori-motrices». Il y a des souris tridimensionnelles à boule qui captent les efforts de l'utilisateur sur la boule, mais en général, ceux-ci servent indirectement pour manipuler des objets et non pour mesurer intrinsèquement des efforts de la main, qui seraient appliqués sur les objets virtuels. Il existe qu'en même quelques manettes pour la Réalité Virtuelle dotées de capteurs de force piézoélectriques. Ceux-ci mesurent les pressions des doigts exercées par la personne. Mais dans cette interface également, on emploie plutôt ces capteurs d'effort pour manipuler une main virtuelle plus que pour appliquer des efforts sur un objet virtuel. C'est un dispositif rudimentaire qui peut se substituer à un gant de données. L'utilisateur, en agissant plus ou moins sur les capteurs de pression, fait ouvrir ou fermer les doigts de sa main virtuelle, positionnée dans l'espace virtuel grâce à un traqueur à 6 ddl incorporé dans la poignée (Figure 8.1).

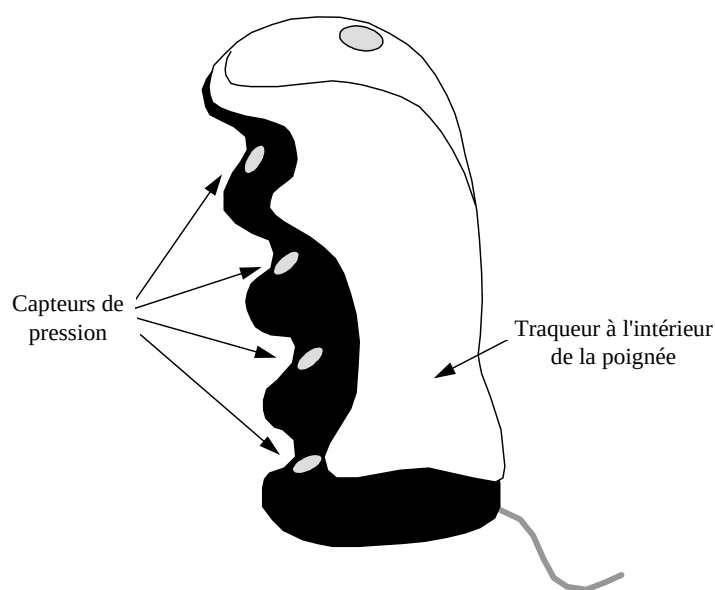


Figure 8.1 : Poignée avec capteurs de pression pour la commande d'une main virtuelle

Avant de présenter les différents types d'interfaces, nous explicitons la problématique de la création des retours d'effort qui est mécaniquement délicate ainsi que les différents actionneurs exploités. Si les retours d'effort apportent une dimension complémentaire majeure dans l'immersion d'un sujet, il faut savoir que ce n'est qu'au prix d'efforts techniques très importants.

8.2 ETUDE DE LA CRÉATION DES RETOURS D'EFFORT

8.2.1 RÉALISATION PHYSIQUE DES EFFORTS

Une interface à retour d'effort, exercé sur les muscles, transmet une information globale à la personne sur les forces de contact entre elle et l'objet virtuel. Mais ces dispositifs peuvent aussi s'opposer activement aux mouvements, en général des doigts ou de la main. Elles ont en conséquence plusieurs fonctions exploitables :

- appliquer des efforts variables sur la personne ;
- résister partiellement au mouvement de la personne ;
- empêcher un mouvement de la personne.

Ces conditions d'exploitation découlent simplement des lois de la statique et de la dynamique. En exemple, dans le premier cas, simuler le poids d'un objet virtuel tenu par la main ; dans le deuxième cas, simuler l'enfoncement de la main dans un corps déformable ; dans le troisième cas, simuler un obstacle fixe virtuel sur le parcours de la main.

Mécaniquement, les seules forces utilisables sur un espace assez grand sont les forces par contact. Les forces à distance sont inexploitable, car soient très faibles à distance (électromagnétiques) ou soient non maîtrisables (gravitationnelles). Il existe néanmoins des interfaces électromagnétiques à retour d'effort, conçue par IBM et aussi par le Robotics Institute de Carnegie Mellon University. Mais ce principe n'est exploitable que pour les manettes à retour d'effort où l'espace de travail en déplacement est très faible (Figure 8.2). L'interface de ce laboratoire a les caractéristiques suivantes :

- espace de travail : une sphère de 25 mm et plus ou moins 12° en rotation ;
- six degrés de liberté en mouvement et six actions de retour d'effort ;
- résolution en position de 10 microns ;
- force maximum de 60 N et couple maximum de 3 m.N ;
- bande passante de 75 Hz ;
- raideur de 2500N/m et de 50 Nm/rad.

La manette Magic Wrist d'IBM [Berkelman, 1997] propose une fonctionnalité similaire avec six ddl à retour d'effort. Son espace de travail est très faible : plus ou moins 5 mm et $3,5^\circ$. Il s'agit de créer des efforts sur une main pratiquement immobile qui donne des ordres de commande. La force maximum en pic est de 20 N et pour le couple de 1,7 m.N. La raideur est de 2300 N/m et de 0,8 m.N/deg. La bande passante est de 50 Hz (Figure 8.3).

8.2.2 ARCHITECTURE MÉCANIQUE DE L'INTERFACE

Sauf dans les cas précités, les efforts sont donc transmis par contact. La loi de la statique entre l'action et la réaction impose que tout solide en contact avec le corps de



Figure 8.2 : L'interface électromagnétique à retour d'effort : Haptic Magnetic Levitation device

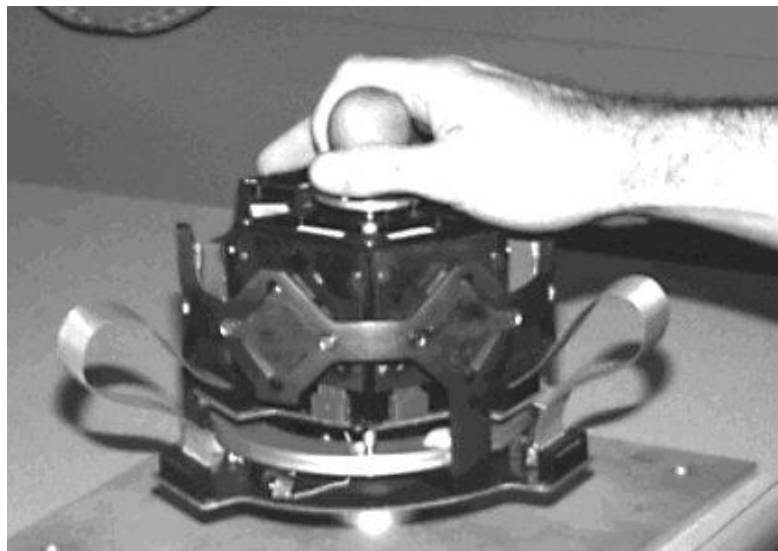


Figure 8.3 : La manette Magic Wrist d'IBM à six ddl à retour d'effort, copyright photo IBM

l'utilisateur soit aussi en contact ailleurs. Il y a donc deux types théoriques de solutions : le mécanisme créant des efforts s'appuie soit sur une autre partie du corps (gant à retour d'effort par exemple), soit sur un socle fixe (manette sur table de travail par exemple). La première solution, **retour d'effort à réaction interne**, permet une plus grande liberté de mouvement, mais a pour inconvénients le poids du mécanisme non compensable et l'encombrement autour du corps. La deuxième solution, **retour d'effort à réaction externe**, a les avantages et inconvénients inverses (Figure 8.4). Il faut noter aussi qu'il y a quelques interfaces qui sont entre ces deux solutions, tels que certains exosquelettes relativement volumineux : pour que l'utilisateur ne supporte pas sur son membre supérieur tout le poids du mécanisme de l'interface à réaction interne, ce dernier est aussi fixé à une de ses extrémités à un portique fixe.

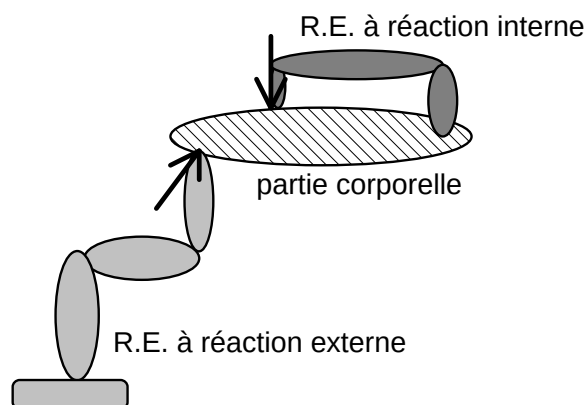


Figure 8.4 : Les deux architectures d'une interface à retour d'effort

Nous en déduisons techniquement qu'il y aura toujours des pièces mécaniques autour du corps, que certaines configurations du corps seront impossibles à cause du mécanisme et donc que *la réalisation d'une interface à retour d'effort universelle ne sera jamais possible* : l'encombrement des pièces ou la limitation des mouvements impliquera toujours des contraintes de fonctionnement pour l'utilisateur. Ceci explique en partie que les interfaces sont bien spécifiques à certains types d'application et *jamais universelles*.

Si les retours d'effort exigent, comme tous les autres canaux sensori-moteurs, des caractéristiques métrologiques classiques, ils ont d'autres spécificités. La première est la sécurité lorsque les retours de force virtuels appliquent des forces réelles à la main. Ces forces doivent être grandes pour impliquer un mouvement déterminé, mais limitées pour ne pas blesser la personne. Les caractéristiques métrologiques des interfaces à retour d'effort sont importantes pour une bonne immersion proprioceptive, en particulier les caractéristiques dynamiques. Mais la qualité de l'immersion dépend aussi des algorithmes de contrôle des actionneurs, implantés dans l'ordinateur. Ces algorithmes doivent exploiter au mieux les possibilités mécaniques (statiques et dynamiques) de l'interface en fonction de l'application souhaitée. L'appareil doit être précis et avoir une bande passante suffisante : plus les fréquences de transmission des efforts sont élevées, plus la simulation est réaliste. Il a été démontré que si le sens du toucher est sensible à des vibrations pouvant atteindre des fréquences de 10 KHz et la sensibilité pour les forces jusqu'à 300 Hz [Brooks, 1990], les gestes sont par contre assez lents et il est possible de les échantillonner de manière convenable autour de 10 Hz. Les forces

appliquées au dispositif à retour d'effort doivent donc être remises à jour à une cadence élevée. Mais, puisqu'elles sont directement reliées aux mouvements de l'opérateur, ces forces vont évoluer lentement. L'idée est donc de se contenter de calculer les forces à une fréquence relativement basse (autour de 30 Hz) et d'estimer la force que l'on va appliquer à l'interface à retour d'effort entre deux calculs successifs de la force. Cette caractéristique de bande passante est importante et dépend de celles du système mécanique et des actionneurs. Un des challenges technologiques est de concevoir de mini actionneurs ayant des bandes passantes élevées. Une résolution suffisante permet des actions progressives et évite les vibrations parasites. Le manque de raideur du mécanisme empêche de transmettre des efforts brefs et nets ainsi que le blocage total et instantané de la main.

8.3 LES DIFFÉRENTES INTERFACES À RETOUR D'EFFORT

8.3.1 RETOUR D'EFFORT À RÉACTION EXTERNE

8.3.1.1 Du bras maître à l'interface à retour d'effort

Les premiers travaux sur le retour d'effort faisaient appel à de grands manipulateurs conçus pour la téléopération. Cette technique est exploitable sans retour d'effort, mais dans ces conditions l'opérateur ne ressent pas physiquement les efforts pendant ses travaux et a du mal à les exécuter. La téléopération avec retour d'effort permet à l'opérateur de manipuler l'extrémité du *bras maître*, sur lequel le monde téléopéré retourne en réaction des efforts. Les bras maîtres sont donc des organes de commande pour piloter un robot à distance (Figure 8.3.1.1). La plupart de ceux-ci utilisent une structure robotique relativement classique de type série ou parallèle. Une présentation des bras maîtres se trouve dans un rapport du CEA, institution ayant conçu et développé des bras maîtres pour la téléopération [Gosselin, 2001].

A ses débuts, la réalité virtuelle a exploité des bras maîtres. Ensuite les techniques employées sur les bras maîtres ont été transposées pour la réalisation d'interface à retour d'effort spécifique pour la réalité virtuelle. A la différence des bras maîtres, les interfaces à retour d'effort privilégient en général le confort au détriment d'un espace de travail important et la sensibilité au détriment d'une forte capacité en effort.

Les algorithmes de commande de ces dispositifs ont autant d'importance que les qualités intrinsèques du mécanisme pour obtenir une interface performante (voir le chapitre correspondant). En général, l'interface à retour d'effort fournit six degrés de liberté de mouvement, qui sont normalement mesurés (ce n'est pas difficile à faire !), et trois ou six degrés de liberté pour le retour d'effort. Les interfaces à trois degrés de liberté pour le retour d'effort (trois en force ou trois en couple) sont courantes et assez facilement réalisables. Par contre, la conception et la réalisation d'une interface à six degrés de liberté pour le retour d'effort sont très délicates à entreprendre. Il y a peu d'offres commerciales correspondantes. Mais des recherches sont menées dans les laboratoires pour trouver de nouvelles structures mécaniques ayant des avantages pour tel type d'application. Si en théorie, nous pouvons définir l'interface visuelle idéale (voir le chapitre correspondant), il n'existe pas pour les interfaces à retour d'effort de structure mécanique idéale qui permettrait de faire tout type de retour d'effort. S'il n'existe pas d'interface à retour d'effort commercialisée et adaptée à votre application, il faut se

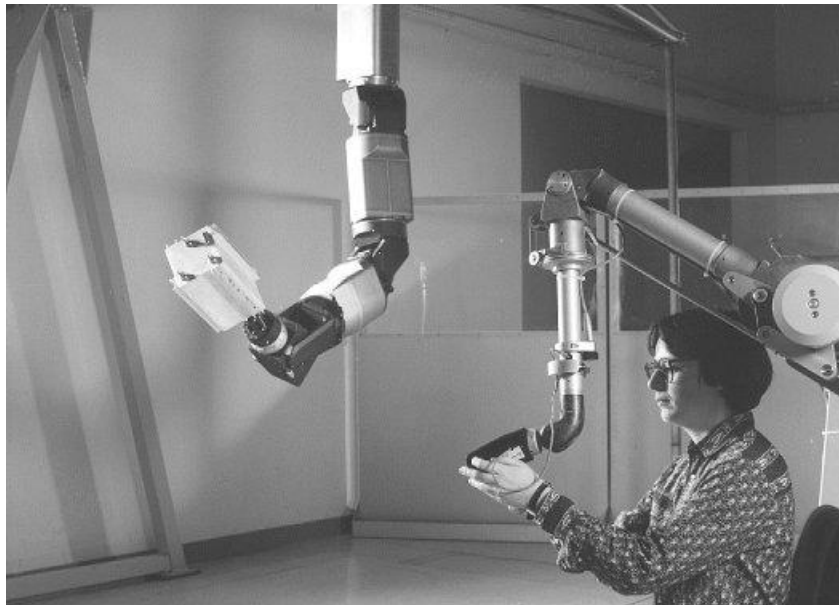


Figure 8.5 : Le bras maître MA 23 M de Cybernétix à 7 ddl à retour d'effort (position et pince), copyright photo Cybernétix

résoudre à en concevoir une, faite sur mesure. Ceci est évidemment nettement plus difficile à entreprendre que de faire une interface visuelle adaptée à votre application, la mécanique et le génie mécanique étant des techniques lourdes en investissements et en moyens.

La structure de l'interface peut être du type série, de la forme identique à un bras robot manipulateur ou du type parallèle, de la forme identique aux robots parallèles. Dans cette dernière configuration, les avantages principaux sont la rigidité plus grande de l'interface, donc une meilleure précision ainsi que la possibilité de mettre les actionneurs sur le support fixe. Les inconvénients principaux sont la plus faible étendue de travail, relativement aux dimensions de l'interface, et les limites absolues des rotations qui sont faibles.

8.3.1.2 Les interfaces classiques à réaction externe

Interface à structure série

Nous pouvons d'abord présenter la série d'interfaces à retour d'effort qui est actuellement la plus connue et exploitée : les Phantoms de la société SensAble Technologies. Le Phantom est directement issu de travaux au MIT. Le Phantom fournit six degrés de liberté de mouvement, qui sont mesurés, et trois (option classique) ou six degrés de liberté pour le retour d'effort. La conception du système Phantom permet à l'utilisateur d'interagir avec l'ordinateur soit en tenant un stylet à la main, soit une poignée ou soit en insérant son doigt dans un «dé à coudre». Son intégration lui permet de disposer des

points fort suivants :

- bonne fidélité du retour d'effort ;
- possibilité de travailler dans un environnement bureau ;
- compatible aux standard PC et UNIX.

Il existe deux gammes différentes : le Phantom premium et le Phantom Desktop. La première étant divisée en trois groupes : le Phantom 1.0, le Phantom 1.5 et le Phantom 3.0. pour lesquels seule la gamme d'espace de travail est différente. Ces interfaces sont vendues avec un logiciel de mise en œuvre, Ghost (General Haptics Open Toolkit), qui est un outil de développement orienté objet (C++). Il inclut les fonctions nécessaires à une prise en main et une intégration rapide des fonctions de retour d'effort aux logiciels existants. Le kit de développement GHOST SDK inclut des exemples de code, des aides au débogage, une librairie de primitives 3D et des objets aux caractéristiques haptiques.

Le Phantom Premium 1.0 permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 3 degrés de liberté (mode point). Son espace de travail couvre une zone de 13 cm x 18 cm x 25 cm, correspondant approximativement à la couverture de mouvement de poignet. Ses caractéristiques techniques sont :

- résolution en position : 0,03 mm ;
- espace de travail² : 13 x 18 x 25 cm ;
- friction : 0,04 N ;
- raideur : 3,5N/mm ;
- force maximum instantanée³ : 8,5 N ;
- force maximum en continue : 1,4 N ;
- inertie : < 75g.

Le Phantom Premium 1.5 permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 3 degrés de liberté (mode point). Son espace de travail plus grand couvre une zone de 19,5 x 27 x 31,5 cm, correspondant approximativement aux mouvements de l'avant-bras autour de l'axe du coude (Figure 8.6). Ses caractéristiques techniques sont :

- résolution en position : 0,03 mm ;
- espace de travail : 19,5 x 27 x 31,5 cm ;
- friction : 0,04 N ;
- raideur : 3,5 N/mm ;
- force maximum instantanée : 8,5 N ;
- force maximum en continue : 1,4 N ;
- inertie : < 75g

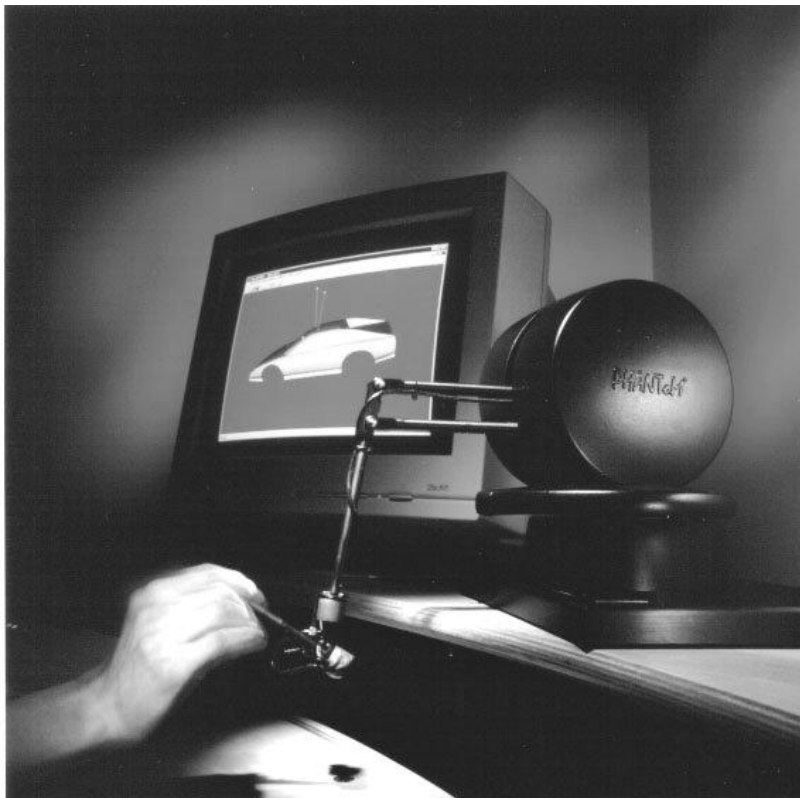


Figure 8.6 : Le Phantom Premium, version 1.5 de SensAble, copyright photo SimTeam

Le Phantom Premium 3.0 permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 3 degrés de liberté (mode point). Son espace de travail plus grand couvre une zone de 41 cm x 58 cm x 84 cm, correspondant approximativement aux mouvement du bras autour de l'axe de l'épaule (Figure 8.7). Ses caractéristiques techniques sont :

- résolution en position : 0,02 mm ;
- espace de travail : 41 cm x 58 cm x 84 cm ;
- friction : 0,2 N ;
- raideur : 1 N/mm ;
- force maximum instantanée : 22 N ;
- force maximum en continue : 3 N ;
- inertie : < 150g.

Le Phantom Desktop est une interface compacte qui se place sur une petite zone de bureau. Il permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 3 degrés de liberté (mode point). Son espace de travail plus grand couvre une zone de 16 cm x 13 cm x 13 cm (Figure 8.8). Ses caractéristiques techniques sont :

- résolution en position : 0,02 mm ;
- espace de travail : 16 cm x 13 cm x 13 cm ;
- friction : 0,06 N ;
- raideur : 3,2 N/mm ;
- force maximum instantanée : 6,4 N ;
- force maximum en continue : 1,7 N ;
- inertie : < 75g.

Le Phantom 1.5/6DOF de SensAble a six degrés de liberté et restitue un retour d'effort de 6 degrés de liberté (3 forces et 3 couples) sur une poignée (Figure 8.9). Ses caractéristiques techniques sont :

- résolution en position : 0,03 mm ;
- résolution en orientation : 0,0023°, sauf pour le roulis : 0,008° ;
- espace de travail : 19,5 cm x 27 cm x 37,5 cm et en rotation : 335° et 260° ;
- friction : 0,04 N en translation ;
- force maximum instantanée : 8,5 N ;
- force maximum en continu : 1,4 N ;

²L'espace de travail n'étant pas un simple parallélépipède, il faut se renseigner sur la définition des dimensions de l'espace de travail : est-ce les dimensions du parallélépipède inclus dans l'espace de travail ou est-ce les dimensions maximums de l'espace de travail, de forme quelconque ?

³Il faut distinguer la force maximum instantanée qui est applicable pendant un court instant (les moteurs électriques chauffent car ils ne sont pas en fonctionnement nominal) de celle qui est applicable en continue (pendant plusieurs heures, 24 H pour les Phantoms).



Figure 8.7: Le Phantom premium, version 3.0 de SensAble, copyright photo SimTeam



Figure 8.8: Le Phantom Desktop de SensAble, copyright photo SimTeam

- inertie : environ 100g ;
- raideur : 3,5 N/mm en translation et 5873 Nm²/rad en rotation ;
- couple maximum instantanée : 515 m.N et 170 m.N ;
- couple maximum en continu : 188 m.N et 48 m.N.

Une interface similaire et concurrente est le Virtuose du Service Robotique et Systèmes Interactifs du CEA-LIST. Suite à leur expérience de conception de bras maîtres, ils ont conçu et commercialisé (par l'entreprise Haption) une interface à retour d'effort avec une structure mécanique série de type robot manipulateur. Cette interface a deux versions :

- Virtuose 3D, qui permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 3 degrés de liberté ;
- Virtuose 6D, qui permet l'acquisition des six degrés de liberté et restitue un retour d'effort de 6 degrés de liberté.

Par rapport aux versions du Phantom, le Virtuose a de meilleures caractéristiques en étendue de travail, en force maximum et en raideur. Les moteurs électriques sont des moteurs «brushless».

Les caractéristiques du Virtuose 6D sont :

- résolution en position : 0.02 mm ;



Figure 8.9: Le Phantom 1.5/6DOF de SensAble, copyright photo SimTeam

- résolution en orientation : 0.006° ;
- espace de travail : 40 x 40 x 40 cm ;
- frottements secs : 2 à 3 N ;
- force maximum en continu : 40 N ;
- couple maximum : 2 m.N ;
- inertie apparente : environ 3 Kg ;
- poids : 35 Kg.



Figure 8.10 : Le Virtuose 6D de Haption, copyright photo Haption

Les caractéristiques du Virtuose 3D (Figure 8.11) sont :

- résolution en position : 0.01 mm ;
- espace de travail : 42 x 49 x 92 cm ;
- frottements secs : 0,3 à 0,6 N ;
- force maximum en continu : 34 N crête (11N continu) ;
- raideur apparente : 2300N/m ;

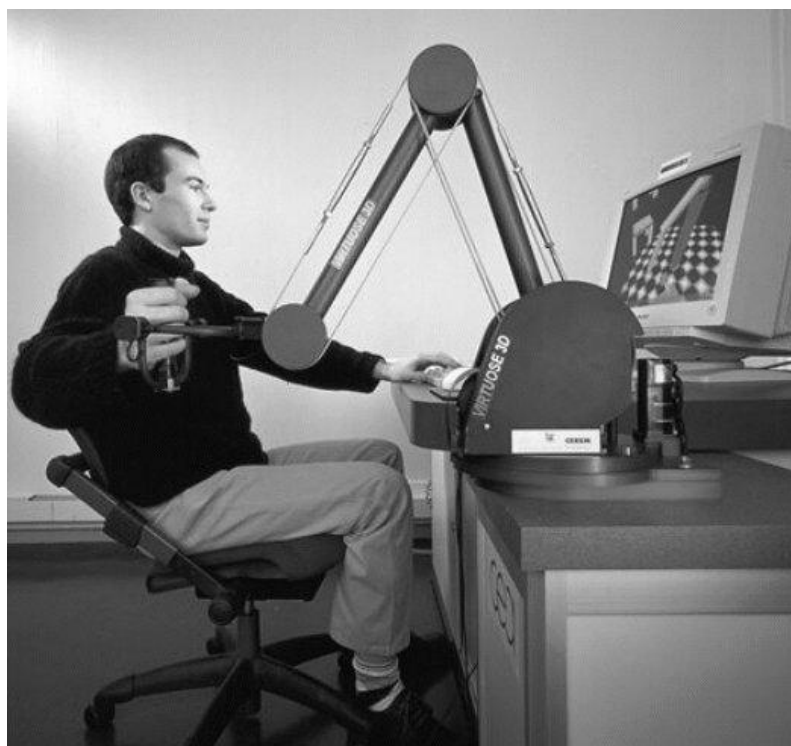


Figure 8.11 : Le Virtuose 3D du CEA-LIST et de Haption, copyright photo Haption

Toujours en structure série, mais basée sur des mouvements de translations pour trois degrés de liberté, l'entreprise canadienne MPB Technologies a conçu l'interface Freedom. La partie de translation est équilibrée dynamiquement. Les efforts pour la partie rotation sont transmis par câbles (Figure 8.12). Ses caractéristiques sont les suivantes :

- l'espace de travail : 22 x 24 x 22 cm en translation et 320°, 100° et 100° en rotation,
- la résolution en position est de 0,02 mm et de 0,005 rad en rotation ;
- la force maximum est de 2,5 N en pic et de 0,6 N en continu ;
- le couple maximum est de 125 m.N en pic et de 80 m.N en continu ;
- la bande passante est 80 Hz ;
- la raideur est de 3000 N/m ;
- l'inertie apparente est de 150 g.

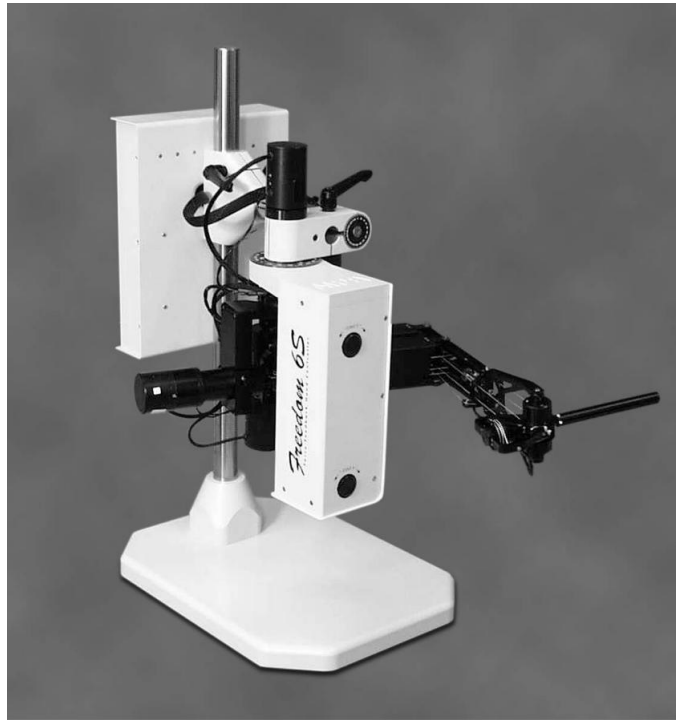


Figure 8.12: L'interface à retour d'effort Freedom 6S de MPB Communications Inc., copyright photo MBP Communications Inc.

Il existe aussi une interface originale à six ddl à retour d'effort en utilisant deux interfaces reliées aux deux extrémités d'une petite barre tenue entre les doigts (dimensions d'un stylo) [Berkelman, 1997], réalisée par l'Université de Tsukuba. Mais cela est toujours possible de relier une barre ou toute autre pièce à deux interfaces à 3 ddl à retour d'effort pour avoir 6 ddl à retour d'effort sur toute pièce.

interface à structure parallèle

Au lieu d'une structure série, des concepteurs d'interfaces à retour d'effort ont exploité les structures parallèles. Dans ce cas, l'interface possède a priori une meilleure précision, mais son volume de travail est plus réduit, relativement par rapport à ses dimensions, qu'une interface série. Accessoirement, ces interfaces à structure parallèle sont moins encombrantes et plus facilement transportables.

Des dispositifs ont été conçus pour des applications spécifiques. Le HapticMaster, de l'entreprise Nissho Electronics [Nissho, 2000] est un mécanisme basé sur trois pantographes, chacun commandé par deux moteurs à courant continu (Figure 8.13 et Figure 8.14). L'interface est un système à 6 degrés de liberté à retour d'effort. Son espace de travail est une sphère de rayon de 20 cm. La résolution est de $400\text{ }\mu\text{m}$ et les efforts maximums sont pour les forces de 21 N en pic, de 12 N en continu et pour les couples de 0,56 N.m en continu. La bande passante est supérieure à 200 Hz.

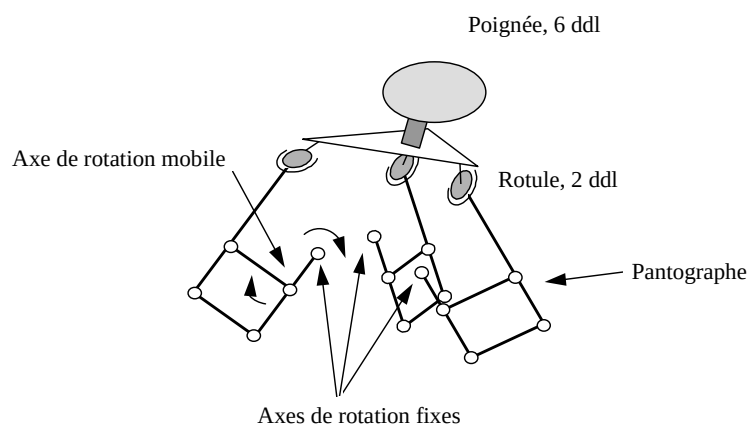


Figure 8.13 : Schéma de l'interface à retour d'effort à structure parallèle Haptic Master

L'entreprise suisse ForceDimension du centre de recherche VRAI Group, issue de l'EPFL de Lausanne, a développé une interface, le Delta haptic Device, basée sur une structure parallèle pour les rotations et pour les translations. L'interface est à six ou trois ddl à retour d'effort (Figure 8.15). Ses caractéristiques sont les suivantes :

- l'espace de travail : 30 cm de diamètre et 30 cm de longueur (cylindre) et plus ou moins 20° pour les rotations ;
- la force maximum est de 25 N en continu dans tout le cylindre et le couple maximum est de 0,2 m.N.

Le centre «Laboratory for Intelligent Mechanical Systems» de l'Université de Northwestern, a développé une interface, MWU Stewart Platform, basée sur une structure parallèle plane pour deux dimensions et une liaison glissière pour la troisième direction avec des moteurs brushless [Youngblut,96]. L'interface est à quatre ddl à retour d'effort : trois translations et une rotation d'axe verticale. Ses caractéristiques sont les suivantes :

- l'espace de travail : 20 x 20 x 9 cm en translation et 90° en rotation ;

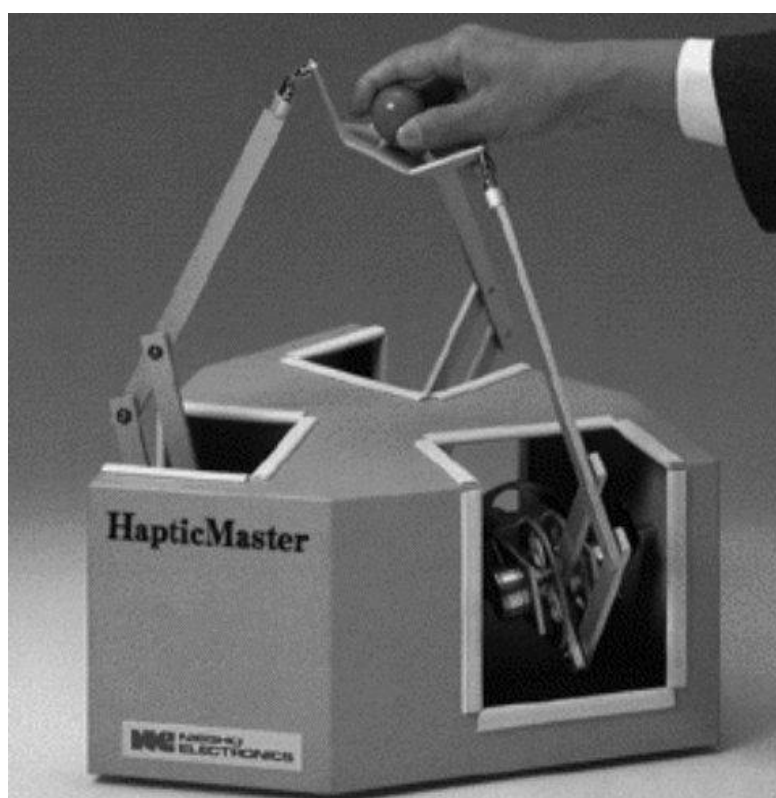


Figure 8.14 : L'interface à retour d'effort à structure parallèle HapticMaster, copyright photo Professor Iwata



Figure 8.15: L'interface à retour d'effort Delta haptic Device de ForceDimension, copyright photos ForceDimension

- la force maximum est de 90 N en pic et de 45 N en continu ;
- le couple maximum est de 1,35 m.N en continu ;
- la bande passante est de 100 Hz.

L'entreprise japonaise, Mitsubishi Precision Company, a conçu et réalisé l'interface MPC Haptic Device à six ddl à retour d'effort, basée sur une structure parallèle. Ses caractéristiques sont les suivantes :

- l'espace de travail : 15 x 15 x 10 cm en translation ;
- la force maximum est de 10 N ;
- l'inertie apparente est de 170 g.

8.3.1.3 Interface à structure particulière

A côté des deux structures classiques, série et parallèle, des interfaces à réaction externe basées sur les structures des robots, certains concepteurs ont réalisé des dispositifs originaux.

L'interface à retour d'effort de l'entreprise américaine Immersion Corporation sert à simuler des opérations chirurgicales, telle que l'endoscopie. Elle permet au chirurgien de simuler la manipulation de l'outil chirurgical. Grâce à l'ordinateur couplé au dispositif, des forces réalistes sont appliquées à la main de l'opérateur (Figure 8.16). Celui-ci sent en retour de la compliance, de la viscosité et même la texture de surface. Cet outil permet au chirurgien de se former aux gestes opératoires ou des les évaluer avant une intervention : le simulateur MIST de formation et d'évaluation en chirurgie laparoscopique.

Le Laparoscope virtuel permet de mesurer les déplacements d'une paire d'outils chirurgicaux, chacun d'entre eux pouvant évoluer sur 5 degrés de liberté. L'interface est basée sur un mécanisme sphérique à cinq barres avec des réducteurs à cabestans et à câbles. Deux degrés de liberté permettent de pivoter autour du point d'insertion. Un degré supplémentaire permet la translation de l'outil le long de l'axe d'insertion sur 20 cm. Un quatrième degré permet à chaque outil de pivoter sur son axe longitudinal. Le cinquième degré de liberté est fournie par la variation d'ouverture de l'outil. Seuls, les trois premiers degrés de liberté sont à retour d'effort. C'est donc une interface à retour d'effort sur 3 ddl. Les cinq degrés de liberté mesurés permettent de transmettre à l'ordinateur une description complète et réaliste de la procédure chirurgicale. Ses caractéristiques sont les suivantes :

- les capteurs fournissent une résolution angulaire de $0,012^\circ$ sauf pour le pivotement sur son axe : $0,35^\circ$; La résolution linéaire est de 0,13 mm ;
- le temps de latence est inférieur à 1 ms ;
- la bande passante est de 650 Hz pour la translation et de 120 Hz pour les rotations ;
- l'espace de travail est de plus ou moins 30° pour les angles et de 100 mm en translation, ce qui correspond approximativement à un volume de 10 x 23 x 23 cm ;
- la force maximum est de 8 N et les couples de 0,43 N.m ;



Figure 8.16 : L'interface à retour d'effort Impulse Engine de Immersion Corporation, copyright photo Immersion Corporation

- la résolution en effort est de 0,14 N ;
- le Laparoscope virtuel peut être facilement modifié pour changer d'outils en cours d'opération. Ces outils standards peuvent provenir de différents fabricants.

Développé par Virtual présence, MIST est un outil de formation et d'évaluation pour la chirurgie laparoscopique. Le logiciel MIST crée un volume opératoire sous la forme d'un graphisme 3D et visualise la position et l'orientation des instruments à l'intérieur de cet espace virtuel en fonction des données envoyées par l'interface. Le logiciel propose une série d'exercices pour développer et tester les aptitudes psychomotrices spécifiquement requises pour des opérations chirurgicales. Les exercices peuvent être configurés selon différents niveaux de difficultés.

L'ACROE de Grenoble a conçu et exploite une interface de communication instrumentale avec retour d'effort. Ce clavier rétroactif permet de capter les mouvements des doigts et d'appliquer des efforts sur ceux-ci. Les touches du clavier peuvent s'enfoncer sur une longueur de trois centimètres. Elles sont munies de capteurs de position ayant une résolution de quatre micromètres. Les retours d'effort sont fournis sur chaque touche par un moteur associé à hautes performances : il fournit une force de 80 N avec un temps de réponse de 0,2 ms. En reliant par des leviers une, deux ou trois touches, on obtient des manettes à un, deux ou trois degrés de liberté. Ce dispositif permet de parfaitement simuler des actions manuelles. Associé à un système de modélisation et de simulation temps réel, tout objet virtuel est considéré comme un ensemble de masses ponctuelles et de liaisons mécaniques entre elles. Suivant les caractéristiques mécaniques des liaisons (viscoélastiques, linéaires ou non), le dispositif peut simuler toutes sortes d'objets : parties d'instrument de musique (archet, corde), objets articulés, fluides, sables, etc. [Cadoz, 1994].

Nous avons présenté dans le chapitre sur les traqueurs, un capteur de localisation (Spidar II) basé sur des mesures par câble. Mais l'intérêt de ces derniers est d'offrir simultanément des retours d'effort sur un ou deux doigts [Sato, 2001]. L'extrémité d'un doigt est attachée à quatre fils reliés à quatre moteurs-encodeurs qui maintiennent une tension sur les fils. Le doigt peut se mouvoir dans un volume d'un mètre cube environ (Figure 8.17).

Grâce aux encodeurs, on relève à chaque instant la longueur des parties déroulées des fils et on détermine géométriquement la position du doigt. Les moteurs peuvent appliquer une tension variable aux fils, ce qui donne les retours d'effort (Figure 8.18). Le dispositif peut être utilisé avec deux doigts, en doublant les fils et les moteurs-encodeurs. Mais dans ce cas, il y a évidemment des incompatibilités de mouvement entre les deux doigts.

Avec cette interface, on peut faire de la «poterie virtuelle» en déplaçant le doigt pour façonner le pot virtuel en rotation ou saisir un objet en sentant les efforts sur deux doigts. Les caractéristiques actuelles sont :

- espace de travail : 90 x 90 x 60 cm ;
- résolution en position : 0,025 mm ;
- effort maximum : 4 N en continu et 16 N en pic ;
- raideur apparente : 10000N/m.

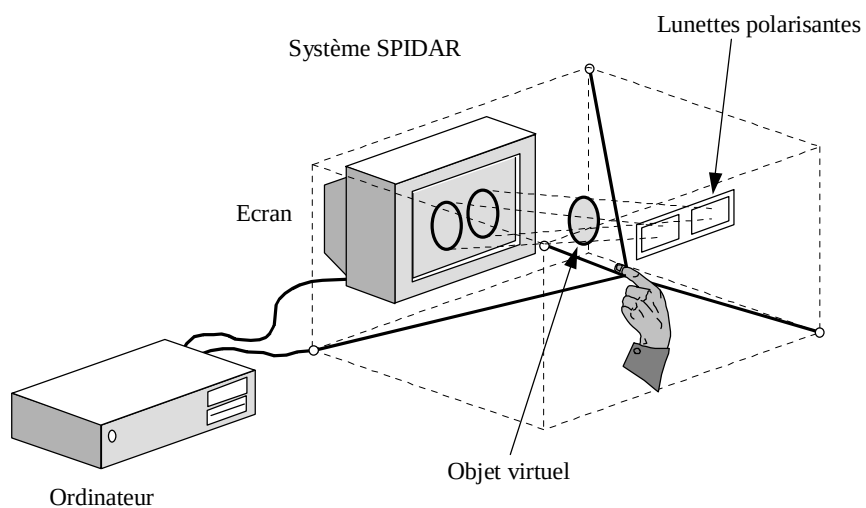


Figure 8.17 : Mesure de position d'un doigt et retour d'effort sur celui-ci

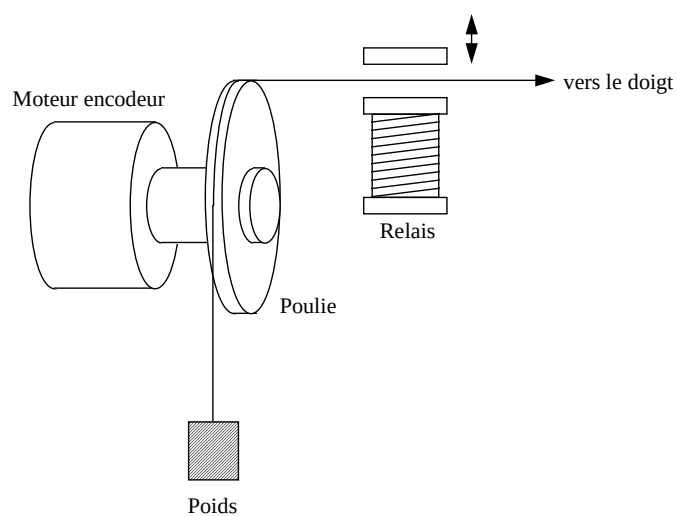


Figure 8.18 : Mécanisme de mesure et de tension d'un fil par moteur-encodeur

8.3.1.4 Manettes à retour d'effort

Il existe des manettes ou joystick à retour d'effort grand public, mais ceux-ci ont des qualités médiocres. Nous ne présentons que des interfaces professionnelles. Les manettes à retour d'effort sont moins encombrantes et peuvent se poser sur une table. Elles ont au moins 3 degrés de liberté et un tout petit espace de travail, quelques centimètres. Les moteurs électriques qui appliquent les efforts sur la manette doivent avoir des caractéristiques dynamiques élevées.

Un manche de commande à performances élevées, le Robotop de Matra Marconi Space, a deux ou trois ddl à retour d'effort et travaille sur un domaine en X/Y de plus ou moins 60 mm et en option en Z de plus ou moins 30 mm. La poignée intègre les capteurs d'effort. L'effort maximum est de 25 N et une raideur de 8000 N/m. Il peut servir à commander le mouvement de l'extrémité d'un robot et en retour informer l'opérateur des forces subies par le robot. Le Robotop a servi durant la mission franco-russe Cassiopée dans la station orbitale MIR en 1996 pour des expériences de perception de forces par le cerveau, en microgravité, mises au point par le LPPA du Collège de France [http1] (Figure 8.19).



Figure 8.19 : La manette Robotop de Matra Marconi Space, copyright Matra Marconi Space

Nous avons présenté au début du chapitre deux manettes basées sur un principe de lévitation magnétique, la manette Magic Wrist d'IBM et l'Haptic Haptic Magnetic Levitation du Robotics Institute de Carnegie Mellon University à espace de travail très faibles.

Pour augmenter l'espace de travail, le Biorobotics Laboratory de l'Université de Washington a réalisé une interface à trois ddl à retour d'effort, l'Excalibur, composée d'une poignée montée sur un plateau à trois translations perpendiculaires. La transmission des efforts se fait par câble. L'espace de travail est de 30 x 30 x 20 cm avec une résolution de 0,008 mm. La force maximum est de 100 N en continu et de 200 N en pic (Figure 8.20).

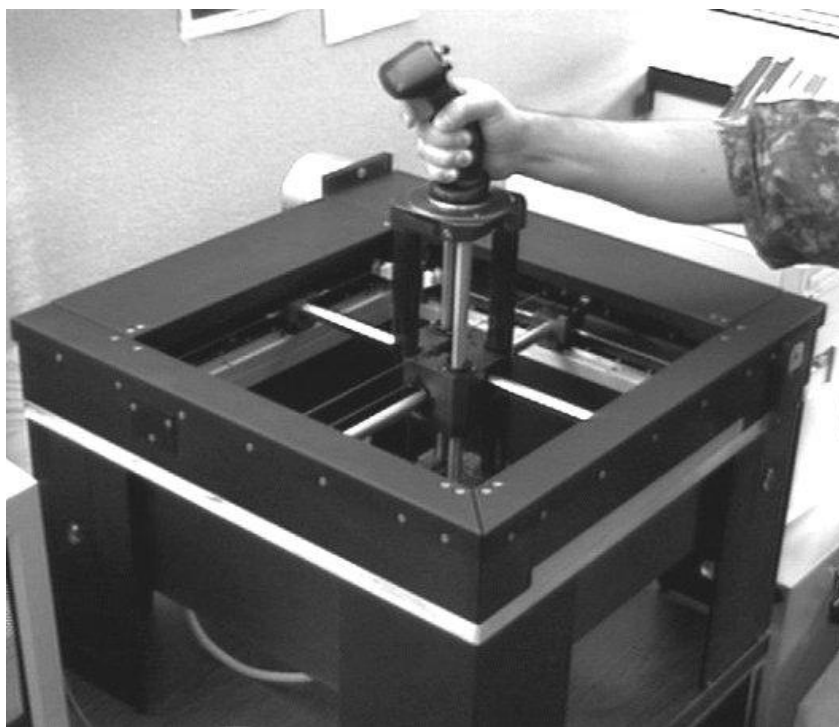


Figure 8.20 : L'interface Excalibur à trois ddl à retour d'effort du Biorobotics Laboratory, copyright photo Biorobotics Laboratory

8.3.1.5 Souris à retour d'effort

Il existe quelques souris 2D avec des retours d'effort. Ces derniers offrent des stimuli qui facilitent l'interaction dans un monde 2D. Elles n'ont pas a priori d'intérêt pour des applications de réalité virtuelle qui sont tridimensionnelles. La société Haptic technologies (Haptech), rachetée par Immersion Corporation, propose la Mouse Cat (Figure 8.21) et le Pen Cat. Elles ont un espace de travail de 7 par 10 cm avec des forces en pic de 10 N, sur évidemment 2ddl.



Figure 8.21 : La souris à deux ddl à retour d'effort Mouse Cat, copyright photo Immersion Corporation

8.3.2 RETOUR D'EFFORT À RÉACTION INTERNE

8.3.2.1 Exosquelette

Le problème des systèmes à réaction externe est de n'agir directement que sur la main (ou quelques fois sur le pied). Les efforts sont transmis indirectement sur les autres parties du corps. Si nous voulons appliquer des efforts contrôlables sur différentes parties du corps, la solution est de placer un mécanisme s'appuyant et agissant tout le long du bras et jusqu'à l'épaule, si besoin. Ces appareils portables permettent un retour d'effort sur le poignet, le bras et l'épaule grâce à un ensemble de tringles, avec l'inconvénient d'un poids non négligeable. Les couples appliqués sont produits par des moteurs à courant continu, trois pour l'épaule, un pour le coude et un pour l'avant-bras. En opposant une résistance variable en différents points du membre, l'interface reproduit plus fidèlement le poids et l'inertie d'un objet virtuel tenu en main. Elle permet également de contraindre les mouvements de l'opérateur, par exemple en cas de collisions virtuelles. Comme nous l'avons signalé auparavant, les difficultés techniques de réalisation sont importantes, surtout pour les actionneurs qui doivent avoir des performances élevées avec une contrainte supplémentaire dans ce cas : la légèreté pour ne pas trop diminuer la transparence du dispositif. La personne ne doit pas se sentir bloquer dans un carcan, mais être libre de mouvement dans un espace virtuel.

La société américaine Exos propose plusieurs exosquelettes agissant sur plusieurs parties du membre supérieur avec différents degrés de liberté. Le système le plus complet,

le Force ArmMaster, a cinq moteurs à courant continu pour cinq articulations : trois à l'épaule (120° de rotation), une au coude (100° de rotation) et une à l'avant bras (100° de rotation). Les dimensions des barres de l'exosquelette sont réglables suivant la morphologie de la personne. Le poids est de huit kilogrammes, les couples de friction sont de 4% des couples moteurs. Les couples maximums sont de 3,5 m.N à l'épaule, 0,8 m.N au coude et de 0,2 m.N à l'avant bras (voir Figure 6.6, page 203).

D'autres exosquelettes, comme le Dextrous Arm Master de l'entreprise américaine Sarcos, sont plus puissants et en conséquence semi-portables. Les parties appliquant les efforts sur les doigts, la main et l'avant bras sont liés à ces derniers tandis que les parties lourdes, correspondant à l'épaule et au bras, sont supportées par un portique fixe. Le Dextrous Arm Master a dix ddl à retour d'effort qui sont fournis par des actionneurs hydrauliques permettant des efforts importants (Figure 8.22).

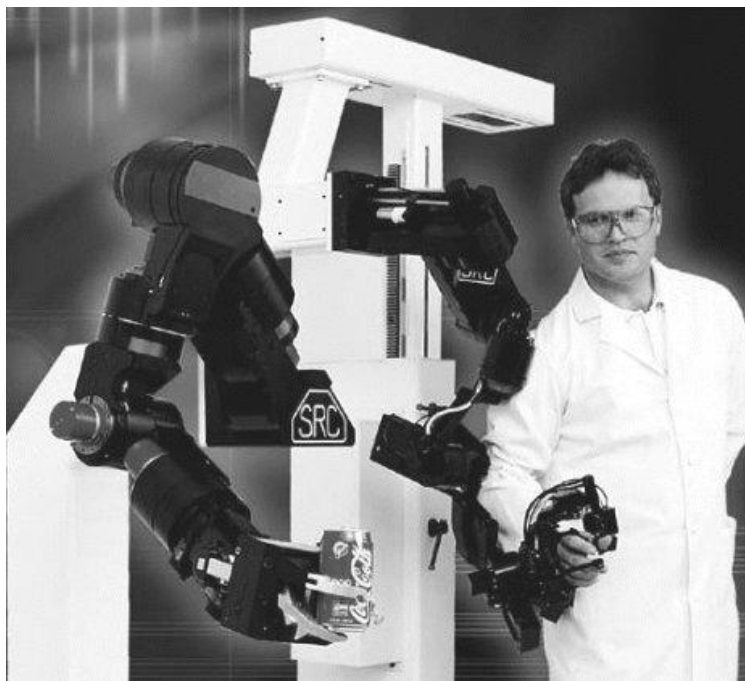


Figure 8.22 : L'exosquelette semi-portable Dextrous Arm Master de Sarcos, copyright photo Sarcos

Un autre exemple d'exosquelette semi-portable est celui développé par le Boston Dynamique Incorporation (BDI) pour des applications de la NASA (Figure 8.23).

8.3.2.2 Gant de données avec retour d'effort

Pour des tâches plus délicates, on peut avoir besoin de contrôler des forces appliquées indépendamment sur plusieurs doigts plutôt qu'au poignet comme pour les bras maîtres. Une interface, le Rutgers Master I, utilisant trois petits vérins pneumatiques



Figure 8.23 : L'exosquelette semi-portable de BDI, copyright photo Boston Dynamics

fixés sur la paume de la main, a été développée. Chaque vérin, monté sur une articulation rotoïde, peut exercer une force jusqu'à 4 N sur un doigt, avec une fréquence de l'ordre de 10 Hz (Figure 8.24). Ce système de poids relativement faible (45 grammes) peut être intégré sur un gant de données. Le dispositif exige un compresseur d'air (pression de 6 bars). L'encombrement de l'intérieur de la main ne permet pas de fermer entièrement celle-ci. L'interface autorise des simulations performantes de retour d'effort sur la main, comme la manipulation d'objets déformables.

Le successeur de ce système est le Rutgers Master II qui est composé de quatre vérins à faibles frottements, mesurant directement leur allongement. Cette solution permet de se dispenser d'employer un gant de données en complément de l'interface à retour d'effort [Burdea, 1996].

Si on veut agir sur tous les doigts, des dispositifs exploitent des actionneurs électriques qui agissent à distance sur les différentes parties des doigts par l'intermédiaire de câbles. Le Laboratoire de robotique de Paris (LRP) a réalisé une telle interface dont les avantages sont de diminuer l'encombrement autour des doigts et le poids sur la main. Des potentiomètres mesurent les rotations des phalanges avec une résolution de un degré pour déterminer la position des doigts par rapport à la paume de la main [Bouzit, 1993].

Sur le même principe, le CyberGrasp de Virtual technologies est un exosquelette léger qui s'adapte sur un CyberGlove pour ajouter un retour d'effort sur chacun des doigts de la main. Les forces sont fournies aux doigts par un réseau de câbles, dont les tensions sont réglables par cinq moteurs à courant continu. Les forces appliquées à chaque doigt sont individuellement programmables. L'exosquelette est conçu pour être adaptable à un grand nombre de mains. Un CyberGlove, gant de données à 18 ou 22 capteurs, est

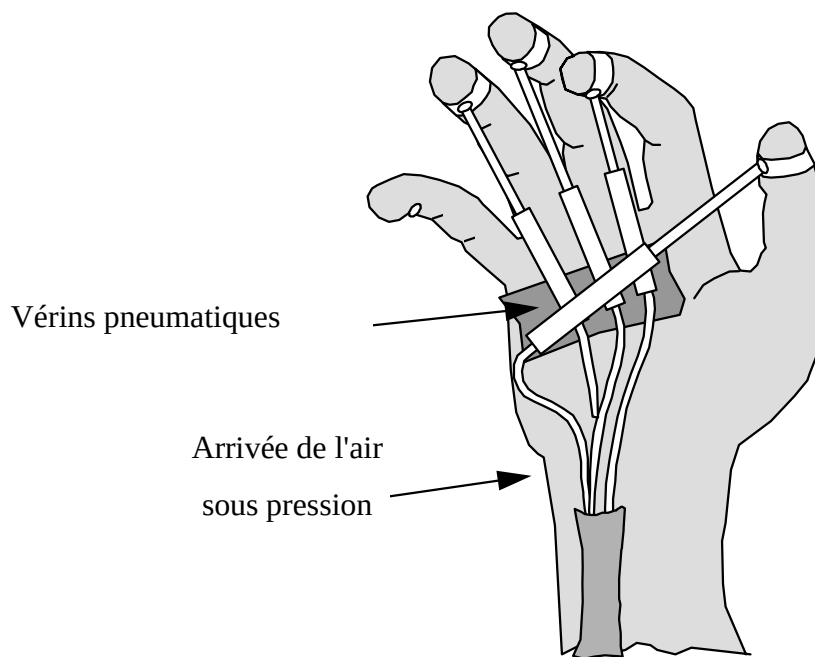


Figure 8.24 : Le gant à retour d'effort Rutgers Master I

requis pour utiliser le CyberGrasp (Figure 8.25). Les caractéristiques techniques sont les suivantes :

- force continue maximale : 12 N par doigt ;
- résolution en position des doigts : $0,5^\circ$;
- résolution de la force sur 12 bits ;
- poids (sans le CyberGlove) : 350 g.

Il existe une nouvelle option pour le CyberGrasp : le Sac à dos GraspPack permet un retour d'effort mobile. Il permet de porter le CyberGrasp, de se mouvoir dans un environnement virtuel avec un plus large espace de travail et éventuellement de travailler avec deux CyberGrasp sur les deux mains en même temps (Figure 8.26).

L'entreprise Exos propose deux gants de données à retour d'effort : Le Dextrous Hand Master et le Safire. Le premier, présenté dans le chapitre des interfaces manuelles motrices, a une structure métallique de type exosquelette, portée sur le dos de la main. Il a quatre capteurs de position à effet Hall par doigt, donc vingt pour l'ensemble du gant. Le gant Safire peut lui avoir jusqu'à onze retours d'effort : trois sur trois doigts (pouce, index et majeur) et deux sur le poignet. Les couples maximums sont de 0,35 m.N à 0,7 m.N. Le poids du gant est de 3,4 kg avec onze retours d'effort, ce qui n'est pas négligeable.

Pour terminer cette présentation des interfaces à retour d'effort, le CyberForce de Virtual Technologies, est un dispositif original : il est composé d'un gant à retour d'effort,



Figure 8.25 : Le gant de données à cinq retours d'effort CyberGrasp de Virtual Technologies, copyright Immersion S.A.



Figure 8.26 : le Sac à dos GraspPack, copyright photo Immersion S.A.

le CyberGrasp, qui est relié à un bras articulé à structure série. Le nombre de retour d'effort est de huit, trois pour le bras articulé et cinq pour le gant (Figure 8.27). Les caractéristiques du gant sont évidemment les mêmes que celles du CyberGrasp et du CyberGlove.

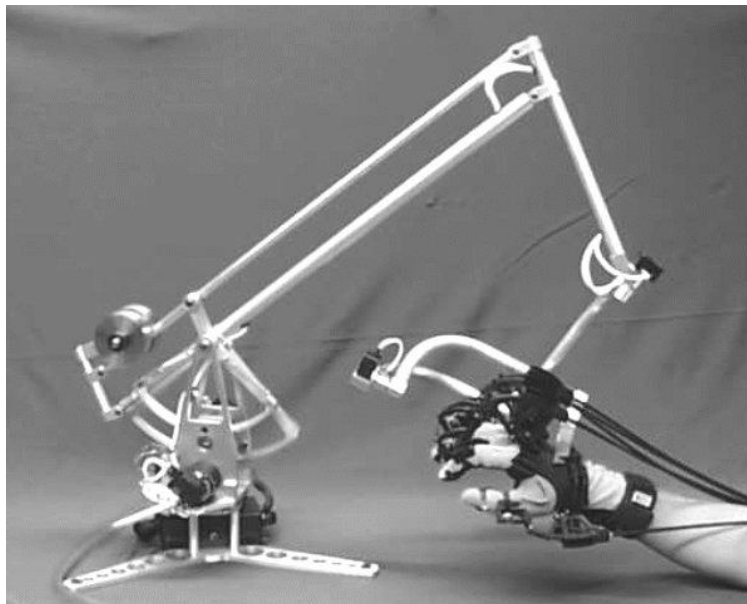


Figure 8.27 : L'interface à retour d'effort CyberForce de Virtual Technologies, copyright photo Immersion S.A.

8.4 CONCLUSION

Si les retours d'effort apportent une dimension complémentaire majeure dans l'immersion d'un sujet, il faut savoir que ce n'est qu'au prix d'efforts techniques très importants. Ce chapitre montre que bien des centres de recherche et des entreprises ont prospecté différentes voies pour fournir des retours d'efforts, principalement sur la main ou sur tout le membre supérieur. S'il s'agit de fournir des efforts uniquement par l'intermédiaire d'un outil, sans déplacement important de celui-ci, les réalisations sont relativement envisageables actuellement. Dans ce cas, le problème mécanique n'est pas trop complexe : il «suffit» de fournir un torseur d'effort mécanique sur une pièce (l'outil) qui est prise en main par l'utilisateur. Par contre, s'il faut créer des efforts sur différentes parties du membre supérieur, jusqu'à des efforts très grands (pour bloquer le mouvement corporel) et en des positions spatiales très variables, distantes de un mètre ou plus, les solutions sont extrêmement difficiles à réaliser. Rappelons aussi que ce chapitre ne présente que le côté mécanique des retours d'effort. Mais la réalisation de la commande de l'interface, fût-elle satisfaisante du point de vue mécanique, reste aussi un problème informatique temps réel (à 1000 Hz) difficile à résoudre.

8.5 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Berkelman, 1997] P. Berkelman, R. L. Hollis «*Dynamic Performance of a Magnetic Levitation Haptic Device*», proceedings of SPIE, volume 3206, Temanipulator and Telepresence Technologies, Pittsburgh, Pennsylvanie, 1997.
- [Bouzit, 1993] M. Bouzit, P. Richard, P. Coiffet P, «*Dextrous handmaster control system, technical report*», Laboratoire de Robotique de Paris, 21 pp, 1993.
- [Brooks, 1990] T. Brooks, «*telerobotics reponse requirements* », IEE International Conference on Systems, Man and Cybernetics, pp.113-120, 1990.
- [Burdea, 1996] G. Burdea, «*Force and touch feedback for virtual reality* », Edition New York : J. Wiley and sons, 339 p, 1996.
- [Cadoz, 1994] C. Cadoz, «*Le retour d'effort dans la communication gestuelle avec la machine*», ACROE, congrès «Interface des mondes réels et virtuels», Montpellier, 1994.
- [Gosselin, 2001] F. Gosselin, «*Etat de l'art des Organes de Commande pour la Réalité Virtuelle*», service robotique systèmes interactifs du LIST, CEA, Fontenay aux Roses, juin 2001.
- [Nissho, 2000] Nissho Electronic Corp. Advanced Electronics Systems Dept : <http://www.tradepia.or.jp/nevc/advanced/vr/vr-e.htm> - Page de présentation du Virtual Reality Group - Mars 2000.
- [Sato, 2001] M. Sato, «*Evolution of Spidar* », VRIC, Virtual Reality International Conference, proceedings, p 81-84, Laval Virtual, 2001.
- [Youngblut, 1996] C. Youngblut, R.E. Johnson, S.H. Nash, R.A. Wienclaw, C.A. Will, «*Review of Virtual Environment Interface Technology*» Kinesthetic Interfaces, Institute For Defense Analyses, IDA, 1996.

Adresse Web des entreprises et des centres de recherche :

[http1] <http://www.cnrs.fr/SDV>

Interfaces á retour d'effort à réaction externe

Interfaces classiques á réaction externe à structure série

Cybernetix	http://www.cybernetix.fr
Haption	http://www.haption.com
MPB Communications	http://www.mpb-technologies.ca/space
Sensable	http://www.sensable.com

Interfaces classiques á réaction externe à structure parallèle

ForceDimension	http://www.forcedimension.com
Nissho Electronics	http://www.tradepia.or.jp/nevc
Virtual Reality and Active Interfaces	http://imts7.epfl.ch

Interfaces á structure particulière

Immersion Corporation	http://www.immersion.com
Biorbtics Lab.	http://rscs.ee.washington.edu/BRL
Virtual presence	http://www.vrweb.com
Iwata Lab.	http://intron.kz.tsukuba.ac.jp/HM/txt.html

Manettes à retour d'effort

Biorobotics Laboratory Washington Univ.	http://www.ornl.gov/Library/biorobotic.htm
Matra Marconi Space	http://www.apgate.com/elec/company
Robotics Institute Carnegie Mellon Univ.	http://www.cs.cmu.edu
Hans Walischmiller GmbH	http://www.hwm.com/products/joyarm

Souris à retour d'effort

Immersion Corporation	http://www.immersion.com
Logicad3D	http://www.spacemouse.com

Interfaces à retour d'effort à réaction interne**Exosquelettes**

Incorporation	http://www.bdi.com
Sarcos	http://www.sarcos.com/teleprod.html
Virtual Technologies	http://www.virtex.com

Gant de données avec retour d'effort

LRP	http://www.robot.uvsq.fr/lrp
Rutgers University	http://vizlab.rutgers.edu/burdea
Virtual Technologies	http://www.virtex.com

Cinquième partie

Les interfaces sensorielles

9 LES INTERFACES MANUELLES SENSORIELLES, LES INTERFACES À SENSIBILITÉ CUTANÉE

Jean-Paul Papin et Philippe Fuchs

9.1 INTRODUCTION

Ce chapitre présente les interfaces à sensibilité cutanée *spécifiques*, c'est à dire celles qui ne procurent que des stimuli tactiles, avec des efforts répartis de faible intensité sur les parties du corps. Rappelons que toute interface à retour d'effort (à forte intensité) crée automatiquement des sensations cutanées en plus des efforts. Mais dans ces dernières interfaces, les sensations cutanées ne sont qu'une conséquence des efforts appliqués et ne sont pas contrôlables directement. Ces systèmes peuvent être utilisés de manière isolée ou en même temps que des systèmes à retour de force, mais dans ce dernier cas il y a la difficulté du mixage des efforts mécaniques. La partie du corps concernée par ces dispositifs est principalement la main, donc ils font aussi partie de l'ensemble des *interfaces manuelles*.

L'absence de sensibilité cutanée dans un monde virtuel ôte une source d'informations qui peut être gênante pour certaines simulations virtuelles. Malgré les difficultés et le peu d'offres commerciales pour ce type d'interface, la sensibilité cutanée de la personne peut être employée à bon escient. Nous avons vu que cette sensibilité est composée du toucher, de la sensibilité à la pression et aux vibrations. Pour ces sensations, nous expliquons les possibilités offertes par les interfaces à retour tactile («sens tactile» pris au sens large : toutes les actions mécaniques sur la peau). La sensibilité superficielle cutanée englobe aussi la sensibilité thermique qui est surtout une sensibilité à la variation de température.

Il est possible de distinguer deux fonctionnalités d'interfaces tactiles :

- les interfaces avec retour tactile informent l'utilisateur sur les effets mécaniques du contact : intensités des forces de contact et glissement à la surface de contact ;
- les interfaces avec retour tactile permettent à l'utilisateur de mieux interpréter l'aspect superficiel de l'objet : sa forme locale de la surface, sa matière et sa texture superficielle.

Ces senseurs pour être efficaces doivent stimuler localement des petites surfaces. Ils peuvent être légers et petits et ainsi ne pas entraîner de fatigue pour l'utilisateur. Ils peuvent être portés par le sujet ou posés sur une table.

Nous pouvons nous poser la question de l'utilité de telle interface. Dans quel type d'application, les interfaces à sensibilité cutanée sont elles utiles, voire indispensables ? Dans les applications de la réalité virtuelle, où au niveau des I² fonctionnelles, un des objectifs de l'application est de reconnaître un objet parmi d'autres. Dans ces conditions, il est très souhaitable d'exploiter de telles interfaces : reconnaissance de sa texture, de sa rugosité, de sa température ou du glissement de l'objet sur la peau :

- pour des applications ludiques ;
- pour de la formation à prédominance manuelle (médicale, artisanale, etc.) ;
- pour des applications de conception de produit, dont l'ergonomie du toucher est importante (ergonomie de conception des planches de bord et des boutons de commandes d'un véhicule, etc.).

Mais ce type d'application n'est pas courant. Pour des applications, où il n'est pas exigé l'utilisation des interfaces tactiles au niveau des I² fonctionnelles, l'emploi de telles interfaces est moins évident. Sûrement, en premier, parce que nous pouvons facilement nous «activer» dans un monde virtuel sans contact tactile et sans variation de température. Autant la vue est nécessaire, autant le toucher n'est pas indispensable pour «vivre» dans un monde virtuel. Nous pourrions proposer peut être la même règle pour l'ouïe, mais il est tellement facile de mettre du son dans un monde virtuel, avec des interfaces peu onéreuses et non invasives, que nous n'hésitons pas à le faire. Car il y a une contrainte majeure des interfaces à sensibilité cutanée : il faut équiper la peau de l'utilisateur.

Au niveau des I² mentales et sensori-motrices, nous pouvons nous poser la question autrement : quels avantages y a-t-il de faire des efforts pour rajouter des stimuli cutanés ? S'il s'agit simplement d'informer le sujet que le contact a lieu entre l'objet et sa main, les schèmes à substitution sensori-motrice sont en général efficaces : au moment du contact, changement de couleur, clignotement de l'objet, variation brusque de la force sur l'interface à retour d'effort ou bruit caractéristique. Mais si on veut que le sujet soit bien informé du contact ou attrape correctement un objet en main, ces interfaces ont leur utilité. Les informations de surface glissante aident la sensation individuelle de stabilité et de prise d'un objet. Sans une telle information, l'utilisateur doit appliquer une trop grande force de serrage (avec une interface à retour d'effort) pour être sûr que l'objet est fermement maintenu. L'utilisateur se fatigue du fait d'un effort accru.

Par rapport aux autres dispositifs, ces interfaces à sensibilité cutanée sont encore peu répandues. Elles sont encore souvent en cours de développement ou au début de leur commercialisation. Les recherches dans ce domaine ont commencé il y a environ une trentaine d'années. Dans la suite de ce chapitre, nous présentons ces interfaces suivant les deux phénomènes physiques :

- les interfaces à retour thermique qui stimulent les récepteurs thermiques de la peau ;
- les interfaces à retour tactile qui interagissent avec les récepteurs tactiles par des actions mécaniques.

9.2 INTERFACE À RETOUR THERMIQUE

Les sensations qui apparaissent lors d'un changement de température dépendent de la vitesse de changement, de l'étendue de la surface stimulée et de la température initiale. Peu d'interfaces à retour thermique ont été développées, en partie parce que les applications de ces interfaces sont peu nombreuses. L'ordinateur commande une petite pompe à chaleur à effet Peltier pour simuler des variations de température. La pompe à chaleur permet de transférer de la chaleur du côté froid (l'air ambiant) au côté chaud (en contact avec la peau). Un thermocouple associé à la pompe relève la température à chaque instant. L'ordinateur peut donc réguler en boucle fermée la température et d'une

façon efficace en élévation ou en baisse car la pompe à chaleur est réversible. Des interfaces permettent de monter à plus de 67° tandis que d'autres pour des questions de sécurité limitent la montée en température à 45°. Le temps de réponse est très petit, de l'ordre du centième de seconde et les petites dimensions de l'interface permettent de la placer à l'extrémité d'un doigt. La pompe à chaleur est commandée par l'ordinateur via une liaison série RS232 qui régule la température par une loi PID.

Le phénomène, exploité classiquement dans les capteurs de température à thermocouples, est basé sur deux jonctions entre deux métaux différents : si les deux jonctions sont à deux températures différentes, il se crée une différence de potentiel qui est fonction des températures. Connaissant une température, on peut mesurer l'autre. Mais si nous inversons les causes et les conséquences, nous obtenons l'effet Peltier [Bonnier, 1997]. Celui-ci a trait à la production ou à l'absorption de chaleur au niveau de la jonction de deux conducteurs quand un courant d'I les parcourt :

$$dQ = \Pi_{12} dI \quad (9.1)$$

avec dQ : quantité élémentaire de chaleur, Π_{12} : coefficient lié aux deux conducteurs de la jonction et dI : courant élémentaire. Un changement de sens de parcours du courant transforme le dégagement de chaleur en absorption de chaleur. C'est le principe de la réfrigération par effet Peltier.

Le X/10 est une interface à retour thermique conçue par CM Research, fonctionnant suivant l'effet Peltier. Elle possède 8 canaux, chacun pouvant être configuré comme une entrée (capteur de température) ou une sortie (interface cutanée thermique). Chaque canal possède une « thermode » qui peut éventuellement se placer au bout des doigts (Figure 9.1).

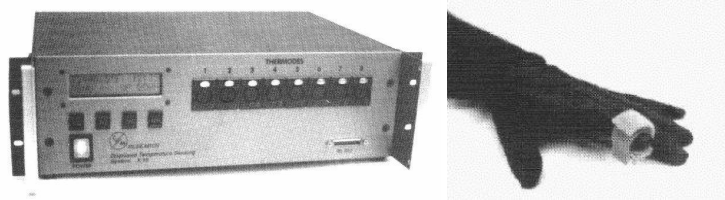


Figure 9.1 : L'interface thermique X/10 et une thermode, copyright photo CM Research

C'est la seule interface thermique commercialisée. Elle est utilisée actuellement :

- dans des applications de robotique à distance : des capteurs de température sont placés sur un manipulateur. Les signaux de température sont envoyés au X/10 et contrôlent les thermodes sur les doigts de l'opérateur ;
- dans les applications où interviennent des prothèses : des capteurs placés sur des prothèses sont utilisés pour déterminer la température. Le X/10 est utilisé pour transmettre cette température en un autre endroit du corps sensible à celle-ci ;
- dans les applications de Réalité Virtuelle : dans certains mondes virtuels, il est important de pouvoir ressentir la température.

9.3 INTERFACES À RETOUR TACTILE

Les interfaces à retour tactile sont des interfaces sensorielles qui permettent de stimuler la peau de la personne. Nous employons le terme «retour» car dans les applications de téléopération et de Réalité Virtuelle, les stimuli tactiles sont en général créés par l'ordinateur en retour à une action motrice de la personne. Les interfaces pour le sens tactile sont en cours de développement ou au début de leur commercialisation. Les principes physiques de ces interfaces sont basés sur la pression pneumatique, sur les vibrations de bobines, sur les matrices d'aiguilles ou sur les manettes à retour tactile. Des essais ont été aussi effectués avec des biocapteurs qui envoient des impulsions électriques sur la peau simulant le contact.

9.3.1 RETOUR TACTILE PNEUMATIQUE

Une méthode pour réaliser un retour tactile consiste à équiper un gant de micro ballons. Ce principe a été employé pour concevoir le gant «Teletact», de UK's Advanced Robotics Research Center, qui possède 20 ou 30 ballons à air situés à l'intérieur de la paume et des doigts. Le gonflement (jusqu'à 0,8 bar) et le dégonflement des ballons est piloté par une interface de contrôle proportionnel, qui gère des vannes électropneumatiques permettant l'entrée ou la sortie d'air dans les ballons. L'air sous pression est fourni par un petit compresseur. La pression n'étant pas très élevée, cette interface procure des sensations de toucher et de pression sur la peau. En gonflant le ballon dans la paume de la main à plus haute pression (2 bars environ), on peut bloquer le mouvement de la main. Dans ce cas, on se rapproche plus d'une interface à retour d'effort. Ce type d'interface présente pour principal inconvénient de n'avoir qu'une faible largeur de bande en fréquence.

9.3.2 RETOUR TACTILE À MATRICE D'AIGUILLES

Une autre méthode consiste à utiliser des plaquettes de matrices d'aiguilles qui sont actionnées par des fils en alliage à mémoire de forme et viennent à la rencontre des doigts. Ces fils en alliage de nickel et de titane, appelé Nitinol, ont la particularité de se raccourcir quand on les monte en température, ce qui leur permet de prendre une forme mémorisée. En les chauffant légèrement électriquement, leur variation de dimension permet d'appliquer de petites forces, non négligeables par rapport à la dimension de ces actionneurs. Les ordres de grandeur sont les suivants : avec un courant de 3 volts et de 0,2 ampère, le fil de 15 grammes passe d'une longueur de 10 à 9,6 cm et crée une force de 1,3 Newton.

On peut ainsi simuler plusieurs points de contact sur un seul doigt. Les aiguilles sont suffisamment fines pour stimuler une petite partie de la peau avec une faible force pour ne pas atteindre le seuil de douleur. Suivant la position des aiguilles actionnées sur la matrice, on peut simuler une surface en contact, une arête ou seulement un coin d'un objet. Une sensation vibratoire peut être obtenue par passage et extinction répétés du courant électrique. Un dispositif portable a été développé et a été intégré dans un gant (Figure 9.2).

Ces micro actionneurs sont légers, compacts et ont un rapport puissance/poids élevé, ils ont une efficacité énergétique faible (seulement 2%-5%). Un inconvénient concerne

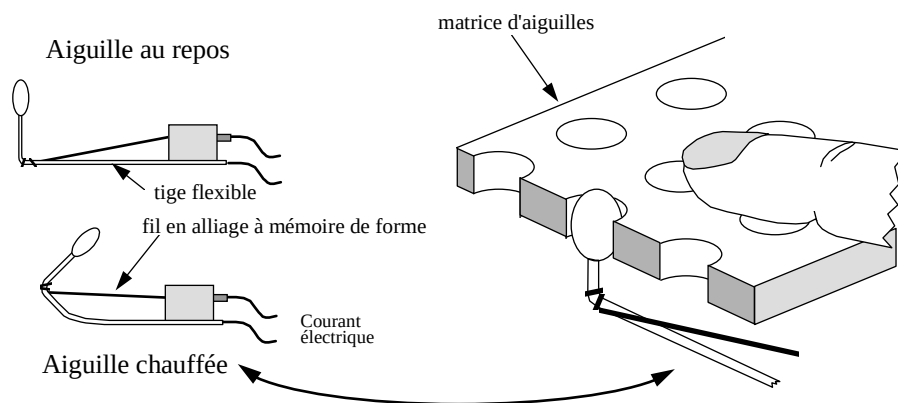


Figure 9.2 : Principe d'un retour tactile par une matrice d'aiguilles

la variation de température non désirée car la déperdition de puissance fournit de la chaleur parasite.

La société Transdimension a mis sur le marché le plus petit dispositif disponible sur le marché : le «tactor TDITAC8x9». Il s'agit d'un actionneur linéaire de 8 mm de diamètre et de 9.6 mm de long. Il peut transmettre une force jusqu'à 0.25 Newton et créer des déplacements jusqu'à 2.5 mm. Il existe en plusieurs dimensions : 8 par 9,8 mm et 12,1 par 14 mm. Les poids sont de 3 g et 3,5 g, le courant électrique est de 60 mA et 85 mA. Une autre version d'actionneur peut fournir des déplacements de 3 mm avec une force 1 N (Figure 9.3).



Figure 9.3 : Le tactor TDITAC8x9 de la société Transdimension, copyright photo Transdimension

Une des rares offres commerciales du marché, le contrôleur de matrice tactile (Programmable Tactile Array Controller) a une interface graphique et un langage facile à utiliser qui permettent à l'utilisateur de générer n'importe quel motif spatio-temporel

de stimulation tactile. Il permet de contrôler une matrice jusqu'à 256 tactors. Cette technique est la plus adaptée actuellement aux interfaces tactiles car les tactors ont une dimension réduite, sont légers et permettent une bande de fréquence acceptable. Ils permettent les plus grands déplacements et les forces les plus importantes pour ce genre d'interface.

Le VideoTact de Unitech Research n'a pas été développé pour la réalité virtuelle. C'est une interface à retour tactile pour l'interface graphique «Windows». Elle remplace l'interface graphique par une interface tactile et permet aux aveugles d'utiliser un ordinateur. L'affichage d'images à différents formats numériques et vidéo est possible. Le VideoTact consiste en une matrice de 768 électrodes en titane, rafraîchies 100 fois par seconde.

Le « Tactor Evaluation System » (TES) de Engineering Acoustics, Inc. n'a pas été développé pour la réalité virtuelle, lui aussi, mais pour montrer la viabilité d'utiliser des stimulations tactiles de la peau pour communiquer en temps réel des informations de navigation ou de danger à un plongeur en train de travailler dans des conditions de faibles visibilités. Les tactors n'utilisent pas des alliages à mémoire de forme mais des petits moteurs linéaires comprenant un aimant mobile et une bobine fixe enfermés dans une cage remplie d'huile.

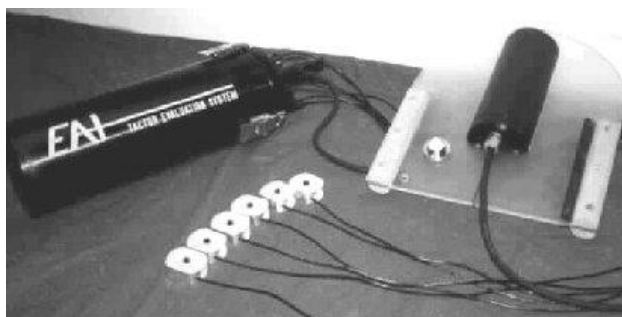


Figure 9.4 : L'interface tactile «Tactor Evaluation System» de Engineering Acoustics, copyright photo Engineering Acoustics

La peau est stimulée par un appareil qui transforme les impulsions électriques en un signal mécanique : un tactor émet des impulsions sur la peau à une fréquence donnée. Le TES fournit des informations sur la direction à suivre ou sur la profondeur à l'aide de 6 tactors. Les indications sont envoyées sur les différents tactors avec des impulsions qui correspondent aux ordres « haut, bas, droite, gauche ».

9.3.3 RETOUR TACTILE PAR VIBRATIONS

Pour un retour de vibrations, on peut employer des bobinages électriques, comme celles utilisées dans les haut-parleurs. En excitant ces bobines à une fréquence de l'ordre de 100 à 250 Hz avec des amplitudes variables, on applique des vibrations à la peau. Ces bobines sont normalement en contact avec l'extrémité des doigts et ne donnent qu'une seule information homogène sur la peau. Les retours de vibrations peuvent servir à la simulation d'une pièce virtuelle en train de vibrer mais peuvent servir comme infor-

mation symbolique. Le principal défaut de ce type d'interface est donc l'impossibilité d'appliquer spatialement des stimuli à amplitude différente. Les bobines audio sont petites, légères et donc confortables. Elles demandent peu de puissance électrique. Elles sont bruyantes et relativement coûteuses. Les haut-parleurs miniatures sont relativement gros et ne sont pas à même d'être placés directement sur le bout des doigts de l'utilisateur. Ils ont un meilleur usage mis sur un plus grand actionneur, les vibrations étant transmises par une interface non-portable plus grande. Nous pouvons aussi noter le cas particulier du siège vibrant, basé sur le même principe, dont l'objectif est de fournir des mouvements oscillants et des sensations tactiles sur une grande partie du corps et non plus sur la main. Cette interface est présentée dans le chapitre des interfaces à simulation de mouvement.

Le CyberTouch de Virtual technologies est une interface tactile pour le gant Cyberglove de la même entreprise. Des petits stimulateurs tactiles sont attachés à chaque doigt et sur la paume de l'utilisateur pour produire des vibrations soutenues. Ils peuvent être utilisés individuellement ou combinés pour produire des motifs synchronisés. La fréquence de vibration générée est contrôlable et peut varier de 0 à 125Hz (Figure 9.5).



Figure 9.5 : Le gant de données Cyberglove avec retour vibro-tactile Cybertouch de Virtual Technologies, copyright photo Immersion S.A.

Il existe d'autres interfaces sur le même principe comme le TouchMaster de EXOS, entreprise rachetée par Microsoft (fréquence maximum de 210 Hz et poids de 0,2 oz). Cette interface tactile s'incorpore dans le gant de données à retour d'effort, le Dextrous HandMaster, de la même entreprise (Figure 9.6).

Pour les stimulations vibratoires, des fréquences autour de 64 Hz sont recommandées pour activer les récepteurs FAII (Fast Adapting II ; les fibres à adaptation rapide avec un champ de réception diffus ; voir le chapitre sur les sens et réponses motrices de l'homme). Dans ce cas il est seulement nécessaire d'agir sur le bout d'un doigt par un seul actionneur car les récepteurs FAII ne fournissent pas d'information localisée.

D'autres procédés ont été employés pour appliquer des stimuli tactiles sur la peau : des solénoïdes, des électrodes ou des actionneurs piézoélectriques. Des recherches sont en cours pour développer de petites matrices à actions électromécaniques (MEMS), ayant une résolution ponctuelle de 20 par 20 sur une surface d'un centimètre carré. Les micro



Figure 9.6: L'interface tactile TouchMaster de EXOS, copyright photo Exos

actionneurs intégrés sur cette matrice sont adressables électroniquement. Ce dispositif peut être léger et peu onéreux.

9.3.4 RETOUR TACTILE PAR MANETTE

L'utilisateur, tenant une manette avec ses doigts ou sa main, ressent les mouvements qui sont imposés à la manette. Des sensations de rugosité d'une surface virtuelle sont obtenues en appliquant des forces de résistance pendant le déplacement de l'extrémité de la manette. Des simulations de fluide visqueux peuvent aussi être créées. La difficulté technique à surmonter est d'appliquer de faibles variations de stimuli tactiles sur la manette. Des manettes appliquant des stimuli plus importants sont classées dans les interfaces à retour d'effort. La séparation n'est pas toujours bien nette entre celles qui n'appliquent que des sensations tactiles et celles qui imposent de faibles efforts aux muscles.

Des constructeurs de joystick et de souris exploitent en retour des vibrations sur ces dispositifs. Les vibrations peuvent être variables. Sur une manette d'un constructeur, les vibrations sont codées sur 8 bits, 4 pour l'intensité et 4 pour la durée des vibrations. Elles peuvent indiquer une intensité de contact, une rugosité de surface ou toute autre information symbolique.

Le Tactile TrackPoint Touch Feedback d'IBM possède un actionneur au-dessous du périphérique de pointage qui peut stimuler le doigt de l'utilisateur. Il donne un choc contre le doigt pour indiquer la position du curseur. Mais cette interface est à employer dans les interfaces IHM classique plus que pour des applications de réalité virtuelle.

9.3.5 RETOUR TACTILE PAR ÉLECTRODES CUTANÉES

Elles sont spécialement conçues pour transmettre des courants électriques aux différents capteurs cutanés. Ces actionneurs demandent une consommation de puissance assez basse, ils sont légers, n'ont pas de mouvements propres et restent en contact constant avec la peau. Ils sont, pourtant, plus invasifs que les systèmes pneumatiques

ou vibrants et peuvent produire des douleurs si le courant électrique n'est pas approprié. De plus la résistivité de la peau varie avec des conditions de sécheresse, la peau glabre est meilleure que celle ayant une forte pilosité pour une stimulation électro-cutanée. Des dispositifs, qui mesurent des signaux électriques générés par l'activité musculaire (électromyogramme), sont commercialisés chez BioControl Systems.

9.3.6 LE CAS PARTICULIER DU MIXAGE DE PHÉNOMÈNES VIRTUEL ET RÉEL

Si le lecteur réfléchit sur ce que permettent les interfaces présentées, il voit facilement les limites de toute simulation tactile actuelle. Ces interfaces stimulent des petites zones de la peau d'un ou de plusieurs doigts avec des actions assez rudimentaires. Si nous souhaitons simuler assez parfaitement les sensations tactiles lors de l'utilisation de boutons de différentes formes, nous voyons les difficultés et les limites intrinsèques de ces dispositifs. Et en général, il faut aussi simuler les efforts entre la main et le bouton. Une autre solution originale a été proposée, mais dans ce cas les boutons sont réels et donc non modifiables. Ceci supprime une potentialité majeure de la réalité virtuelle. Le système TOPIT (Touched Objects Positioned In Time) a été construit pour fournir un retour d'effort et un retour tactile à l'utilisateur d'un cockpit virtuel. L'utilisateur porte un casque qui projette l'intérieur d'un cockpit, incluant le tableau de bord ainsi que la scène extérieure. Une représentation de la main de l'utilisateur est aussi rendue dans la scène virtuelle. Le pilote peut activer différentes sortes de boutons sur le tableau de bord et ressentir précisément les forces et les textures de ces boutons. L'objectif est de fournir un simulateur qui peut être configuré entièrement par logiciel pour représenter différents cockpits (Figure 9.7).

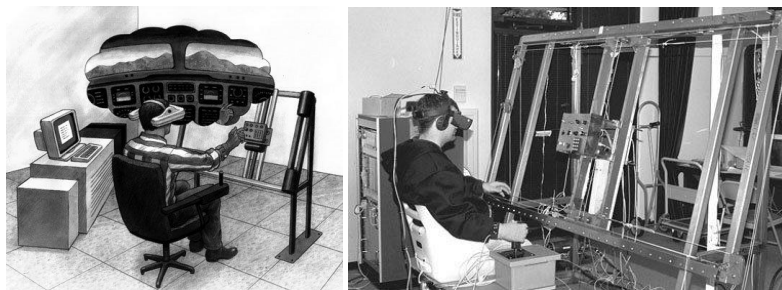


Figure 9.7 : Le système TOPIT pour cockpit virtuel, copyright photo CGSD Corporation

La sensation de toucher d'un bouton sur le tableau de bord est fournie par un mécanisme qui place de vrais boutons dans leur position correcte sous la main de l'utilisateur. Un traqueur et un gant de donnée suivent continuellement la position de la main et l'extrapolent quand l'utilisateur va toucher un bouton. En utilisant la position extrapolée, un ordinateur commande un mécanisme qui place le type correct de bouton sous le doigt de l'utilisateur. On arrive ainsi avec un tableau de quelques boutons à représenter les tableaux de bord complexes des avions modernes. Mais nous voyons la limite de cette simulation : il s'agit de travailler sur la conception ergonomique de cockpit et non sur celle de nouveaux boutons (Figure 9.8).



Figure 9.8 : Le gant et la platine mobile de boutons du système TOPIT, copyright photo CGSD Corporation

9.4 CONCLUSION

En conclusion, le choix commercial des interfaces à sensibilité cutanée est assez restreint. Certains produits sont encore en cours de développement dans des laboratoires de recherche. Peu d'applications en Réalité Virtuelle exploitent ce type d'interface comportementale pour deux principales raisons : ces dispositifs n'offrent que des sensations cutanées rudimentaires et celles-ci sont rarement indispensables dans bien des applications.

9.5 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Bonnier, 1997] G. BONNIER et E. DEVIN, « Thermocouples », Encyclopédie Techniques de l'Ingénieur, tome « Mesures et Contrôle », article R2590, Vol. R 3, 1997.

Adresses d'entreprises :

Interfaces à retour thermique

CM Research <http://www.cmresearch.com/haptic.html>

Interfaces à retour tactile

Retour tactile pneumatique

UK's Robotics Research Center <http://www.hitl.washington.edu/scivw/EVE>

Retour tactile à matrice d'aiguilles

Transdimension <http://www.transdimension.com>
 Unitech Research <http://www.execpc.com/%7Eunittech>
 Engineering Acoustics, Inc. <http://www.eaiinfo.com/page5.html>

Retour tactile par vibrations

Virtual Technologies <http://www.virtex.com>

Retour tactile par manette

IBM <http://www.almaden.ibm.com/cs/user>
Logitech <http://www.vrdepot.com/vrteclg.htm>

Retour tactile par électrodes cutanées

BioControl Systems <http://www.biocontrol.com>

Mixage virtuel et réel

CGSD Corporation <http://www.cgsd.com/TOPIT.html>

10 LES INTERFACES VISUELLES

Philippe Fuchs

10.1 INTRODUCTION AUX INTERFACES VISUELLES

L'utilisation du sens visuel est presque toujours indispensable en réalité virtuelle. Un dispositif RV n'exploitant pas la vision est extrêmement rare. Il pourrait s'agir de dispositif pour les non voyants ou pour représenter leur perception du monde aux voyants. Les logiciels de réalité virtuelle permettent de créer des images de synthèse de qualité de plus en plus grande. Au niveau du matériel informatique, on a des puissances de calculs suffisantes pour créer en temps réel des images tridimensionnelles de qualité, même s'il reste encore des verrous techniques. On ne peut actuellement calculer en temps réel des images photoréalistes avec des reflets suffisamment proches de la réalité. Nous pouvons exploiter d'autres interfaces visuelles qu'un simple écran monoscopique pour obtenir une immersion visuelle efficace.

L'interface idéale et universelle devrait avoir des caractéristiques métrologiques correspondant aux capacités maximales du système visuel humain pour bien exploiter ce canal sensoriel. Cette interface sensorielle doit donc offrir quatre capacités de plus que celles d'un écran ordinaire : de **grands champs de vision horizontaux et verticaux** correspondant à ceux des yeux, une **vision stéréoscopique dans tout le champ de vision binoculaire**, une **haute résolution** graphique exploitant toutes les performances des acuités monoscopique et stéréoscopique ainsi qu'une **immersion du regard**¹ dans le monde virtuel. Ce dernier point n'est possible que si l'interface visuelle est couplée à un capteur localisant l'orientation de la tête de l'opérateur. La puissance informatique doit bien sûr être aussi suffisante pour afficher en temps réel devant les yeux des images gauche et droite à haute résolution, correspondant à la direction de son regard. En pratique, nous verrons techniquement qu'il y a deux types d'immersion du regard proposés :

- l'immersion totale du regard dans un visiocasque ou dans un visiocube à six faces (voir le paragraphe 10.2.5.3). La tête de l'observateur peut se translater et tourner dans toutes les directions ;
- l'immersion partielle et concentrique du regard : devant un écran, souvent de grande taille, l'observateur peut regarder la scène virtuelle (ou l'objet) en translatant ou en tournant son regard sur plus ou moins 45° environ par rapport à la normale de l'écran. La translation de la tête, parallèlement à l'écran, est aussi limitée, de l'ordre de la demi largeur de l'écran.

La présentation des interfaces visuelles dans ce chapitre peut être basée sur différentes classifications, selon que l'on privilégie une des quatre capacités précédemment citées pour l'interface universelle. Nous prenons comme deux premières classes les interfaces qui permettent ou non l'immersion du regard, donc les interfaces portables et les interfaces à support fixe. Les sous catégories sont fonction des champs de vision que proposent les interfaces et ensuite fonction des possibilités de vision stéréoscopique.

¹L'immersion du regard : l'utilisateur voit toujours la scène virtuelle, même s'il translate ou tourne sa tête (son regard) dans n'importe quelle direction.

Les interfaces à support fixe sans ou avec peu d'immersion du regard :

- écran d'ordinateur monoscopique ou stéréoscopique ;
- fenêtre immersive ;
- visiobureau ou bureau immersif ;
- visiosalle et visiocube.

Les interfaces portables avec immersion totale du regard :

- visiocasque ;
- vidéolunette ;
- visiocasque semi-transparent.

Après la présentation des différentes classes d'interfaces, nous expliciterons les problèmes et les solutions de la création des images en relief qui, sous couvert de facilité d'utilisation, demeure une technique délicate à exploiter correctement.

Nous ne présentons dans les paragraphes suivants que les techniques commercialisées ou proche de l'être. Nous ne développerons pas les techniques holographiques, encore non exploitées commercialement pour les images animées, ainsi que les systèmes fournissant des images volumiques tridimensionnelles, spectaculaires mais peu immersives (présentées succinctement en annexe à la fin du chapitre). Les dispositifs commercialisés ne répondent pas toujours au problème d'immersion visuelle souhaitée. Le concepteur de dispositif RV ne doit pas a priori écarter (par facilité ou par ignorance) la possibilité de concevoir par lui-même son interface. Au centre de Robotique de l'Ecole des Mines de Paris, nous avons exploité plusieurs fois cette solution en employant des vidéoprojecteurs (monoscopiques ou stéréoscopiques) qui permettent une architecture et un dimensionnement sur mesure de l'interface visuelle.

10.2 INTERFACES VISUELLES À SUPPORT FIXE

10.2.1 ECRANS D'ORDINATEUR MONOSCOPIQUES

Un écran d'ordinateur classique (donc monoscopique) est souvent utilisé pour différents domaines proches ou inclus à celui de la réalité virtuelle. Les jeux vidéo, la CAO, la téléopération et la télévirtualité exploitent principalement les écrans monoscopiques d'ordinateur. Nous n'explicitons pas ce type de matériel, bien connu par le lecteur. Signalons de nouveau que ce type d'écran n'a pas encore de résolution suffisante comparée à l'acuité visuelle humaine, que ce soit pour les écrans à tube cathodique et encore plus pour les écrans plats LCD (Liquid Crystal Display). Il faut encore attendre les écrans ayant une résolution de 3000 à 5000 pixels horizontalement, suivant le champ de vision exploité.

A noter qu'un constructeur, Panoram Technologies, propose d'augmenter le champ de vision horizontal (140°) et la résolution (3072 par 768 pixels) de son interface visuelle qui est composée d'un ensemble de trois écrans monoscopiques d'ordinateur, placés jointivement et horizontalement à la hauteur des yeux de l'observateur. Une version renforcée est aussi proposée pour une utilisation en extérieure (Figure 10.1).

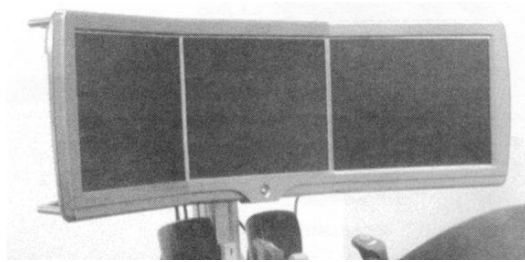


Figure 10.1 : Triple écran pour une immersion grand angle mono-utilisateur, copyright photo Panoram technologies

10.2.2 ECRANS D'ORDINATEUR STÉRÉOSCOPIQUES

Quelle que soit la technique employée en vision stéréoscopique, l'objectif est toujours le même : fournir à chaque œil une image différente. Pour ceci, il y a deux principes : soit on place près des yeux deux petits écrans (cas des visiocasques), soit on sépare techniquement les deux images affichées sur un seul écran, celui-ci étant plus ou moins éloigné des yeux. Dans cette dernière catégorie, le principe repose sur un dispositif technique permettant de séparer les images pour ne visualiser sur chaque œil que l'image désirée. La classification qui suit est donc fonction du lieu où se fait cette séparation et du type de dispositif employé. Nous ne présentons pas certains dispositifs particuliers comme la méthode à déplacement de fentes devant l'écran. Dans ce cas l'œil ne voit à un instant donné qu'une colonne de son image correspondante [Tilton, 1987].

10.2.2.1 Séparation au niveau de l'écran

Écran autostéréoscopique avec illumination :

L'écran est un ensemble de deux plans : l'un est un écran à cristaux liquides, l'autre juste derrière est composé de fines colonnes illuminées séparées entre elles par des zones sombres [Kaneko, 1987]. Il y a une colonne de lumière pour deux colonnes de pixels de l'écran à cristaux liquides. Chaque colonne de lumière est espacée de telle façon qu'un observateur dans une bonne position voit celle-ci à travers la colonne paire de pixels avec son œil gauche, et à travers la colonne impaire de pixels avec son œil droit. Les pixels ne sont visibles dans ce dispositif que s'ils sont éclairés (Figure 10.2).

En affichant l'image de gauche sur les colonnes paires, l'œil gauche de l'observateur perçoit la bonne image, de même pour les colonnes impaires avec l'œil droit. L'écran peut être monochrome ou en couleur et affiche 320 x 380 pixels par œil. Avec un capteur de localisation, l'écran peut afficher des images différentes suivant la position des yeux pour fournir un effet visuel correspondant mieux à une vision réelle [Eichenlaub, 1990].

Écran autostéréoscopique avec réseau lenticulaire :

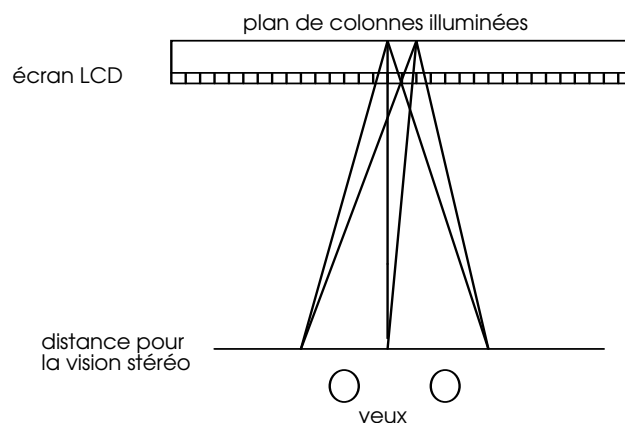


Figure 10.2 : Principe de la vision avec l'écran autostéréoscopique à illumination

Le principe est le même que celui bien connu des cartes postales en relief ou des photos en relief. A noter qu'un laboratoire français propose depuis peu le développement de photos en relief prises avec appareils jetables ou réutilisables. L'écran est recouvert d'un réseau lenticulaire : un plan composé de lentilles demi-cylindriques d'axes parallèles aux colonnes de l'écran. Celles-ci ont des caractéristiques optiques qui impliquent que chaque œil de l'utilisateur voit des colonnes différentes de pixels, correspondant à deux images différentes. Derrière chaque lentille cylindrique, on a pour chaque œil 1, 2 ou 4 colonnes de pixels. S'il y en a plusieurs, en déplaçant légèrement la tête, l'observateur voit la scène affichée en continu sous différents points de vue. Dans ce cas, la vision se rapproche plus de la vision naturelle. Mais ce changement de point de vue n'est possible que sur une petite portion d'espace (Figure 10.3).

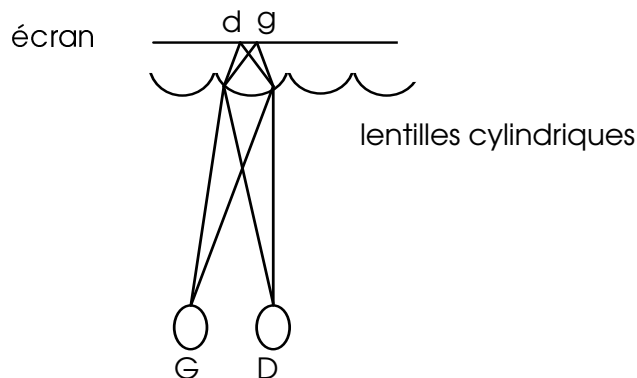


Figure 10.3 : Principe de la vision avec l'écran autostéréoscopique à réseau lenticulaire

Plusieurs types d'écrans ont été testés par des chercheurs japonais : écran PDP (Plasma Display Panels) (640 x 400 pixels en monochrome), écran EL (ElectroLuminescent) (640 x 400 pixels en monochrome), écran LCD (Liquid Crystal Display : écran à cristaux liquides) (720 x 480 pixels RGB à 1440 x 1024 pour des téléviseurs de grandes

dimensions : 50"). Pour alléger l'écran lenticulaire, celui-ci peut être séparé de l'écran d'affichage par une couche d'air, permettant de diminuer l'épaisseur des lentilles tout en gardant les mêmes caractéristiques optiques [Isono, 1990]. Les écrans plats sont préférentiellement utilisés, car il est plus facile dans ce cas d'aligner les colonnes de pixels avec les lentilles cylindriques. C'est la principale difficulté technique des écrans autostéréoscopiques, surtout pour ceux de grandes dimensions. Le principal avantage des écrans autostéréoscopiques est de permettre à l'observateur de ne porter aucune lunette (grand intérêt en aéronautique et surtout dans les lieux publics pour une vision immédiate). Mais, en conséquence de la possibilité de plusieurs points de vue, l'inconvénient est de diminuer la résolution des images au moins par 2 (ou par 4 ou 8). Comme on ne sait pas faire actuellement des écrans à résolution suffisante, ceci est un défaut important de ce principe.

Une invention française, l'alioscope, exploite le principe autostéréoscopique avec un réseau lenticulaire, qui est placé devant un écran plat. Un bloc optique spécifique (1 dioptre convergent et 4 dioptres divergents) est positionné devant le capteur CCD d'une caméra classique. Un ordinateur PC entre la caméra et l'écran est obligatoire pour l'inversion des groupes de 4 colonnes de pixels pour une vision correcte avec 4 points de vue. Ce procédé permet à son utilisateur de filmer avec une seule caméra et une seule optique de vue et de fournir des images vidéo en relief. Les intérêts de cette invention sont : son coût modeste, la perception de l'objet principal au niveau de l'écran, l'existence de plusieurs points de vue. Mais il y a des inconvénients liés à ce principe. La résolution des images est 4 fois inférieure par rapport à un écran de mêmes caractéristiques en résolution, ayant double fréquence pour faire un multiplexage temporel (voir paragraphe suivant). Le moiré de l'écran autostéréoscopique dégrade un peu la qualité des images et les points de vue stéréoscopiques ne sont pas variables [http1].

10.2.2.2 Séparation par des lunettes

Séparation par différenciation colorimétrique :

Le procédé ancien et bien connu avec des lunettes bicolores (procédé anaglyphe) permet facilement et à faible coût de créer sur une feuille ou sur un écran une vision stéréoscopique. Chaque œil ne voit qu'une des deux images de couleur différente, rouge et cyan (la couleur complémentaire du rouge), grâce à deux filtres sur la paire de lunettes. Des programmes existent pour compatibles PC à très bon marché, permettant une vision en relief en couleur, même si certaines couleurs sont difficiles à restituer et si cela procure une fatigue complémentaire due à la reconstitution des couleurs. Ce procédé a très peu d'attrait maintenant que des constructeurs proposent des lunettes actives pour compatibles PC à prix abordable.

Mais le procédé vient d'être remis à jour par l'entreprise TAN, exploitant un brevet de Daimler-Chrysler. Chaque œil de l'observateur reçoit une image filtrée différente sur les composantes Rouge, Vert et Bleue (RVB) du spectre des couleurs. Ceci est possible en mettant une paire de lunettes ayant deux filtres colorimétriques différents et en équipant de même les deux vidéoprojecteurs DLP (ou CRT), affichant les images sur un grand écran. Il est possible faire 4 différenciations sur les filtres et ainsi de proposer à deux observateurs de voir une scène avec deux points de vue différents. Il se pose alors le problème de la qualité de restitution des couleurs exactes des objets virtuels. Ceci sera gênant dans les applications où la qualité colorimétrique est importante. Les

couleurs variant légèrement suivant l'angle de réflexion sur l'écran, la vision stéréoscopique peut être perturbée (Figure 10.4).

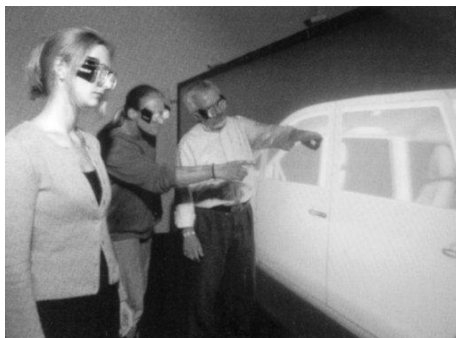


Figure 10.4 : Projection stéréoscopique avec lunettes filtrant par différenciation colorimétrique, photo TAN ProjektionsTechnologie GmbH

Séparation par lunettes à obturateur électronique (lunettes actives) :

Ces lunettes sont composées de deux écrans à cristaux liquides qui obturent alternativement 60 (ou 50) fois par seconde chaque œil, pendant que le moniteur affiche les images à la fréquence de 100 hertz. La fréquence normale du moniteur est donc doublée pour éviter le scintillement des images sur chaque œil (25 images par seconde). Le moniteur doit être prévu pour cette exploitation : la plupart des stations de travail en possèdent un en standard. Les lunettes étaient équipées dans le passé d'écrans PLZT moins transparents que les écrans LCD, mais à séparation parfaite entre les images gauches et droites. Ces écrans PLZT gardent leur intérêt pour des tests psychophysiques.

Pour synchroniser les obturateurs avec l'affichage du moniteur, deux procédés sont possibles : la synchronisation est transmise soit par un câble reliant les lunettes au moniteur (dispositif à faible coût pour PC), soit par des liaisons infrarouges (dispositif classique), solution plus souple d'emploi. Avec une liaison infrarouge, il faut que le récepteur IR sur les lunettes soit positionné dans le cône d'émission de l'émetteur IR. Si ce n'est pas le cas, il faut mettre plusieurs émetteurs IR, ce qui est indispensable dans les salles immersives ou quand l'observateur est très proche d'un grand écran. La difficulté technique majeure est de n'afficher au bon instant que l'image désirée sans empiètement d'une image sur l'autre : phénomène de chevauchement (crosstalk), [Lipton, 1991], [Yeh, 1990]. Les temps d'ouverture (2,8 ms) et de fermeture (0,2 ms) ne sont pas suffisamment petits pour empêcher totalement le chevauchement des images. Ceci procure des images fantômes qui perturbent et fatiguent l'observateur. Certains constructeurs proposent pour les lunettes actives un arrêt de l'obturation lorsque l'on porte son regard en dehors du cône infrarouge de synchronisation, ce qui permet de regarder des écrans monoscopiques. Un avantage des lunettes actives par rapport aux lunettes passives est d'être utilisables avec des très grands moniteurs ayant jusqu'à 37" de diagonale (Figure 10.5 et Figure 10.6) .

Séparation par lunettes polarisantes (lunettes passives) :

Il y a deux techniques pour la polarisation au niveau de l'écran :

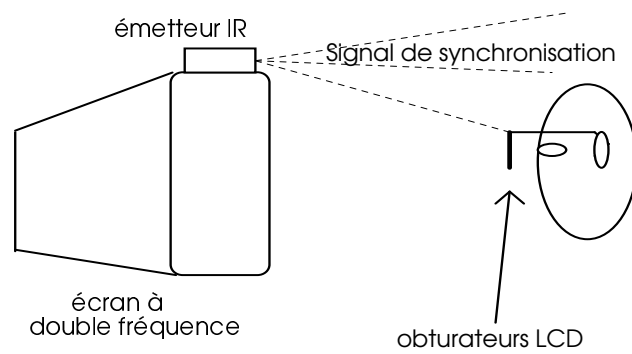


Figure 10.5 : Principe de la vision stéréoscopique avec lunettes actives

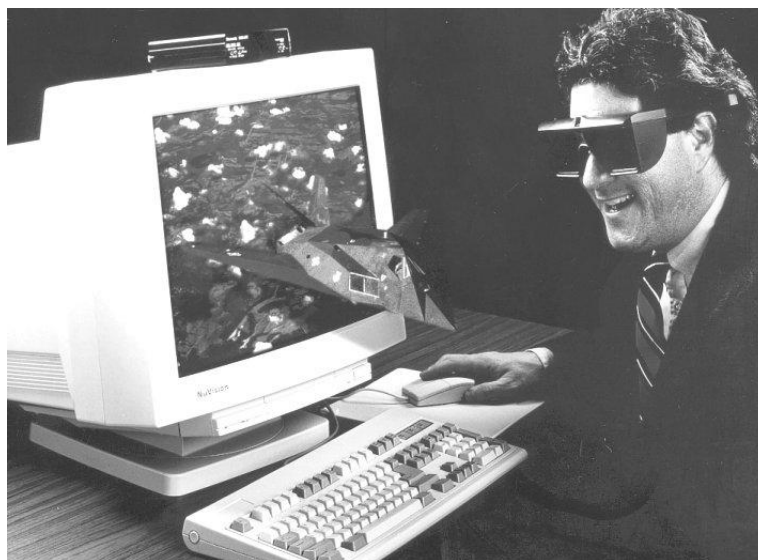


Figure 10.6 : Vision en relief par lunettes actives Stereotek de Nuvision, copyright photo Immersio S.A.

1- Multiplexage temporel :

Devant le moniteur est placé un écran à cristaux liquides permettant une polarisation de la lumière différente dans le temps pour chaque couple d'images stéréoscopiques. Avec des lunettes ayant deux filtres ne permettant que la transmission de l'image désirée pour chaque œil, l'observateur voit les images en relief. La fréquence du moniteur doit être aussi doublée pour la même raison que précédemment et le problème technique majeur est le même que pour les lunettes actives : chevauchement des images droite et gauche entre elles (crosstalk). Les temps d'ouverture (3,5 ms) et de fermeture (0,3 ms) ne sont pas suffisamment petits pour empêcher totalement le chevauchement des images. Les lunettes polarisantes permettent de voir en même temps hors de l'écran sans gêne. Elles ont l'avantage d'être plus légères que les lunettes actives et de ne pas exiger d'alimentation électrique par pile (Figure 10.7).

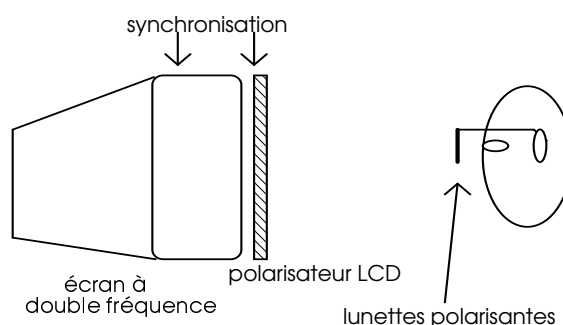


Figure 10.7: Principe de la vision stéréoscopique avec lunettes passive et multiplexage temporel

La polarisation des images est soit une polarisation croisée à 90° (la moins performante et la moins chère pour les lunettes) soit une polarisation «circulaire». Les principes de la polarisation de la lumière et des lunettes actives et passives sont explicités en annexe à la fin du chapitre. Les filtres polarisateurs, devant les écrans ou devant les vidéoprojecteurs, font perdre de l'intensité lumineuse.

Le chevauchement temporel des images est le principal problème technique qui n'est pas parfaitement résolu actuellement. Ceci crée des images fantômes perturbant la vision en relief, surtout pour les images composées de vert qui est la couleur la moins bien séparée. Il faut donc éviter le vert dans les images en couleur quand on le peut, comme pour les représentations symboliques par des couleurs.

Dans ces techniques stéréoscopiques, il est possible de remplacer le moniteur par un vidéoprojecteur à double fréquence permettant une projection sur grand écran (voir paragraphe 10.2.6). Ces deux techniques, les plus employées actuellement, posent les mêmes problèmes psychophysiques pour les images en relief. Il ne faut donc pas affirmer, comme nous l'entendons souvent, que l'on préfère les lunettes passives, ou inversement les lunettes actives, parce qu'on a moins ressenti de fatigue et de gêne oculaires. La gêne dépendant avant tout des caractéristiques des images proposées, la plupart du temps l'observateur des images en relief a eu mal aux yeux parce que celles-ci étaient mal réglées et non parce qu'il utilisait soit les lunettes actives, soit les lunettes passives.

2 - Multiplexage spatial :

Une technique, peu répandue actuellement, a été développée par l'entreprise VRex.

Au lieu de concevoir une polarisation temporelle, le procédé exploite une polarisation spatiale : alternativement une ligne sur deux de l’affichage est polarisée différemment. Avec une paire de lunettes polarisante, chaque œil ne voit qu’une ligne sur deux de l’écran correspondant à la bonne image stéréoscopique. Ce procédé est réalisable grâce à la fabrication de microscopiques polariseurs (μpol) de dimensions inférieures à $10\ \mu\text{m}$ (Figure 10.8).

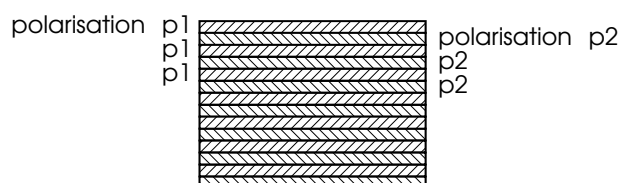


Figure 10.8 : Principe de la vision stéréoscopique avec lunettes passives et multiplexage spatial

Ce principe permet de créer des images stéréoscopiques sur ordinateur, fixe ou portable, et aussi avec rétroprojecteur pour une vision multi-utilisateur. La résolution de l’affichage est divisée par deux verticalement par rapport à la résolution de l’écran.

Remarque : ces trois techniques peuvent être exploitées pour la restitution d’images réelles ou synthétiques. Dans le cas d’images réelles, l’acquisition des deux images peut se faire avec deux caméras standard (synchronisées ou non selon le procédé) ou avec une seule caméra stéréo monobloc (composée de deux capteurs CCD et de deux objectifs à axes parallèles, Figure 10.9). Les possibilités de réglage mécanique ou logiciel pour fusionner les images sont à étudier pour comparer les différentes solutions (réglage des parallaxes, commande des zooms, etc.).

10.2.3 FENÊTRE IMMERSIVE

Si on propose la possibilité de créer des images en relief sur un écran d’ordinateur, on peut aussi rechercher avec un tel type d’écran à augmenter artificiellement l’immersion du regard sans utiliser un visiocasque. C’est ce que propose Virtual Research avec sa solution originale : «Window VR ». L’écran est orientable et accroché à un portique. L’observateur, manipulant l’écran, peut regarder la scène virtuelle dans toutes les directions, « par la fenêtre ». L’écran, de résolution XGA (1024 x 768) ou SXGA (1280 x 1024) et même en 21" (1600 x 1200), est bien sûr traqué : 3ddl (orientation seulement) ou 6ddl. Des boutons se trouvent en plus sur deux manettes pour différentes commandes (Figure 10.10).

10.2.4 VISIOBUREAUX OU BUREAUX IMMERSIFS

Ce type d’interface, apparu au milieu des années 90, permet d’afficher sur une grande surface des images stéréoscopiques (de la grandeur d’un bureau ou d’une planche à dessin). Ce type d’interface permet ainsi à un ou deux utilisateurs de visualiser des objets tridimensionnels. Ces bureaux immersifs (ou visio-bureaux) fonctionnent globalement comme les écrans stéréoscopiques avec séparation par lunettes, excepté que l’image est projetée sur un écran par l’intermédiaire de rétroprojecteur et éventuellement via

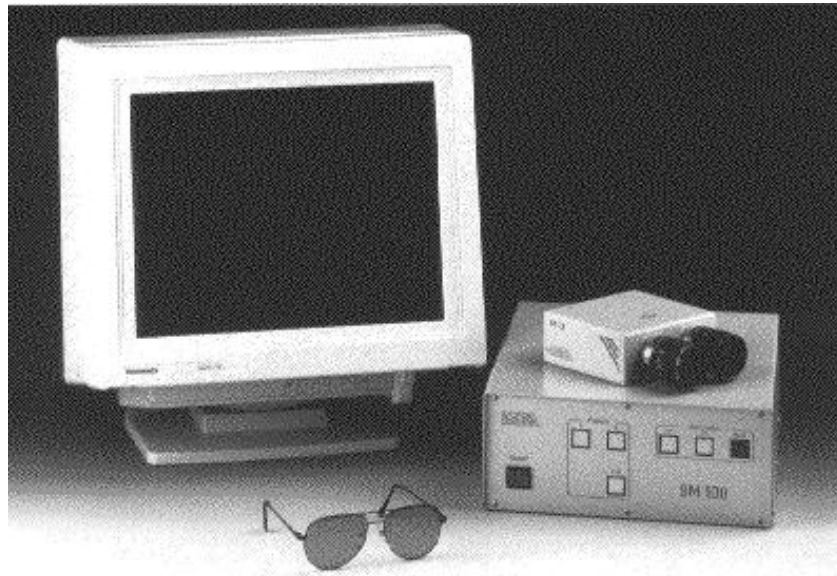


Figure 10.9 : Système Kappa de restitution stéréoscopique d'images réelles, copyright photo ThetaScan



Figure 10.10 : Window VR de Virtual Research, distribuée par Immersion S.A., copyright photo PSA

un ou deux miroirs. La séparation des images stéréoscopiques se fait soit par l'intermédiaire de lunettes actives ou soit par l'intermédiaire de lunettes passives. Deux cas de visualisation tridimensionnelle sont possibles :

- le point de vue des images projetées est fixe : l'observateur doit garder la tête fixe pour ne pas voir la déformation visuelle de la scène virtuelle ;
- la tête de l'observateur étant « traquée », le point de vue change en fonction de la position de la tête pour proposer dans ce cas l'immersion concentrique du regard. La perception tridimensionnelle de la scène virtuelle est alors plus performante (voir le paragraphe sur la perception de la profondeur au chapitre 4). L'observateur peut observer l'objet virtuel sur différents angles pour mieux le percevoir. Mais ces angles différents sont limités ! Il ne faut pas trop croire certaines publicités. Les contraintes des images en relief sont telles que l'on ne peut tourner autour de l'objet qu'à plus ou moins 30° (ordre de grandeur).

Remarques : dans le deuxième cas, avec la tête traquée, un *seul* observateur peut regarder correctement la scène virtuelle ou l'objet virtuel. Combien de démonstrations avec plusieurs observateurs, dont un seul avec la tête traquée, ont eu lieu, perturbant et fatiguant visuellement les autres observateurs ?

Si on veut à tout prix montrer la scène virtuelle à deux observateurs, il faut fournir quatre images différentes (et non deux) ainsi que les lunettes actives ou à différenciation colorimétrique séparant ces quatre images pour les deux observateurs. Mais dans ce cas, la station graphique doit avoir la puissance nécessaire pour générer le double d'images. Certains constructeurs proposent cette fonctionnalité.

Autre remarque : l'observateur, avec sa tête traquée, doit percevoir l'objet immobile, même en bougeant sa tête. Cela n'est bien sûr possible que si le capteur de localisation soit suffisamment précis et avec très peu de temps de latence. Bien des démonstrations mal réglées nous ont convaincus que ce n'est pas simple à résoudre. Un autre capteur de localisation ou un gant de données peut être fourni pour manipuler les objets virtuels.

Indépendamment du suivi de la tête de l'observateur, les dimensions, l'inclinaison et le nombre d'images du bureau immersif sont variables. Car on trouve des bureaux à deux écrans perpendiculaires, un vertical et l'autre horizontal.

Le bureau immersif « Consul » de Barco a deux écrans stéréoscopiques de dimensions 1,02 m par 1,36 m. L'encombrement du meuble est de 2,2 m par 2,2 m par 1,6 m, ce qui est relativement faible pour ce type de matériel (Figure 10.11). Le bureau immersif « Holobench » de TAN a deux écrans stéréoscopiques. L'encombrement du meuble est plus important que celui du Consul (Figure 10.12).

Néanmoins, ces interfaces s'avèrent relativement chères et nécessitent un espace relativement important. Mais la représentation d'objets en grand (pour certains à l'échelle 1) est à ce prix. Barco propose aussi le bureau immersif Baron de dimension 1,7 m en diagonale, avec vision stéréoscopique (Figure 10.13).

Fakespace, l'un des principaux constructeurs de ces bureaux virtuels propose ainsi plusieurs modèles utilisant l'une ou l'autre de ces techniques : Mini Workbench et Immersive Workbench, basés sur un projecteur CRT (tube cathodique) et utilisant des lunettes



Figure 10.11 : Le bureau immersif à deux écran Consul de Barco, copyright photo Barco

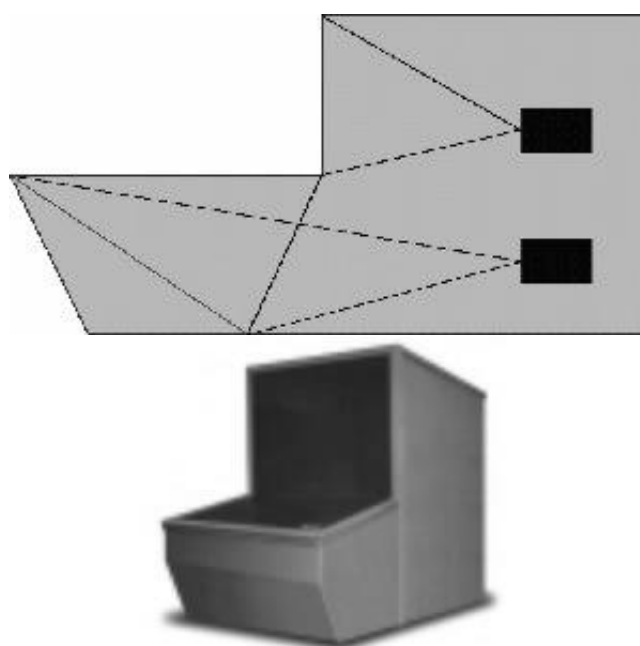


Figure 10.12 : Le bureau immersif Holobench de TAN, copyright photo Tan Projektionstechnologie GmbH

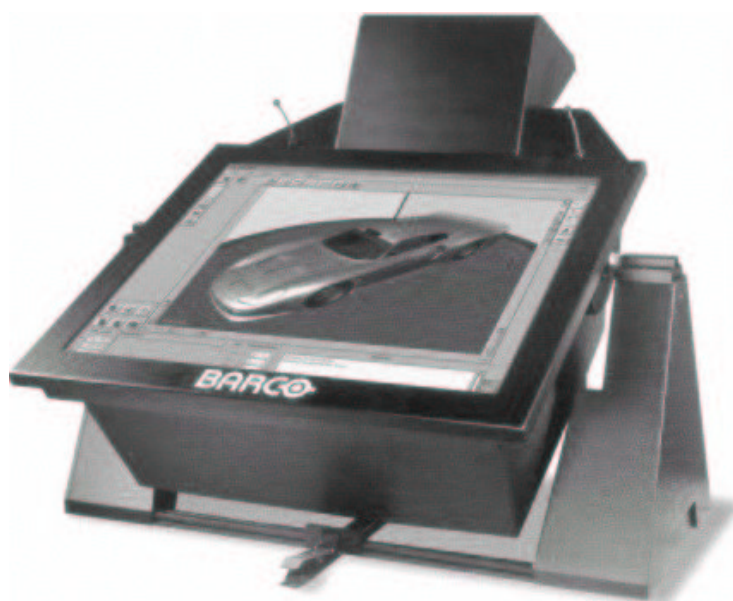


Figure 10.13: Le visiobureau Baron de Barco, copyright photo Barco

actives, offrant des surfaces de visualisation de 44" pour la diagonale du premier et de 90" (soit 2m28) pour le second, avec une résolution de 1280*1024 (Figure 10.14).

VersaBench utilise la polarisation par multiplexage spatial sur l'écran pour restituer la stéréoscopie. Deux projecteurs sont utilisés pour éclairer l'écran de 60" de diagonale.

Un cas particulier et original de «bureau immersif» est la «visionstation» de Elumens (Figure 10.15). L'image est projetée sur une surface sphérique devant l'observateur pour qu'elle recouvre une grande partie de son champ de vision. L'immersion visuelle physique est donc plus grande. La contre partie est que le vidéoprojecteur, ayant une optique grand angle adaptée, se trouve devant l'observateur entre celui-ci et la calotte sphérique. La déformation des images est calculée en soft, ce qui procure un retard dans l'affichage des images. La projection par vidéoprojecteur LCD de 1024 par 768 pixels se fait sur une calotte sphérique de 1,5 m environ de large. Mais Elumens propose plusieurs variantes de plus grandes dimensions : le visiondome V31 de 3 m de diamètre pour l'affichage à 1027x768, le visiondome V4 et le visiondome V5 de 5 m de diamètre pour l'affichage à 1365x1024 pour plus de 10 observateurs (Figure 10.16). Avec ce dernier type de matériel, on est plus dans le cas de salle immersive et non de bureau immersif. C'est le sujet du prochain paragraphe.



Figure 10.14 : Le bureau immersif Workbench de Fakespace Systems, copyright photo Fakespace Systems

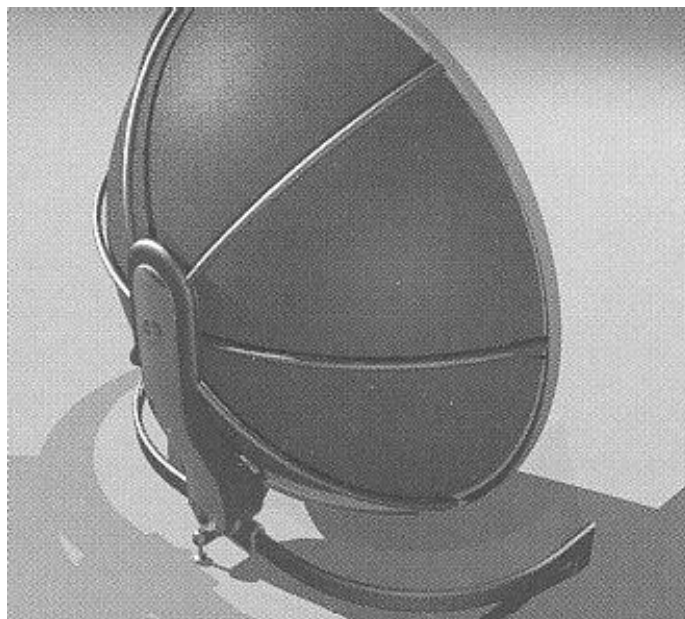


Figure 10.15 : La visionstation d'Elumens, copyright photo Immersion S.A.

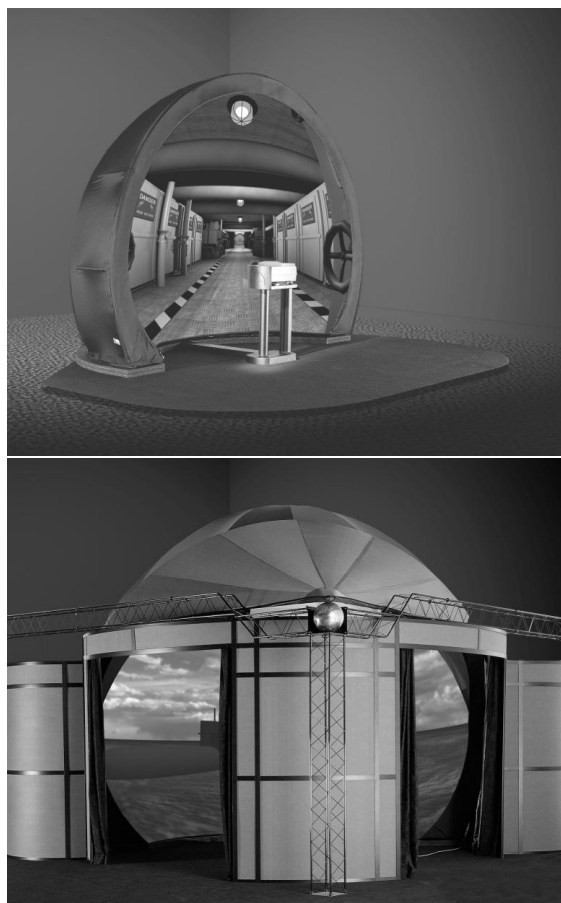


Figure 10.16 : Les visiondome V5 et V31 d'Elumens, copyright photo Immersion S.A.

10.2.5 INTERFACES VISUELLES À TAILLE HUMAINE : VISIOSALLE (SALLE IMMERSIVE) ET VISIOCUBE

10.2.5.1 Choix entre les dispositifs «prêt à voir» et les «sur mesure»

Dans cette partie, nous présentons les interfaces visuelles composées souvent de plusieurs vidéoprojecteurs et d'un ou plusieurs écrans suffisamment grands, à *taille humaine*, pour permettre *l'immersion dans un grand champ de vision* d'un ou de plusieurs observateurs. Ces dispositifs sont exploités en général pour représenter la scène virtuelle à l'échelle 1. Plusieurs vidéoprojecteurs sont souvent employés afin d'augmenter la résolution de l'image globale, ce qui est presque obligatoirement nécessaire si on veut couvrir un grand champ de vision. Notons que de même pour un écran d'ordinateur ou pour n'importe quelle salle immersive, on est loin d'atteindre le niveau de l'acuité humaine : environ 1° d'angle, donc pour 180°, il faudrait une image de plus de 10000 pixels horizontalement (actuellement environ 3000 pixels). S'il existe plusieurs types commercialisés, que nous présentons dans les paragraphes suivants, nous voudrions vous convaincre que nous pouvons aussi concevoir un tel dispositif répondant plus précisément au problème d'immersion visuelle exigée par l'application (suivant les I² sensori-motrices et mentales) : au lieu du « prêt à porter » (ou mieux dit du « prêt à voir »), il faut peut-être faire du « sur mesure ». C'est ce que nous avons entrepris dans le cas du magasin virtuel (écran plat de 2 m par 4 m avec six vidéoprojecteurs) ou dans notre projet IMAVE (Immersion Adaptée à un Véhicule) pour l'entreprise Sommer Allibert. Pour visualiser une planche de bord en relief, nous avons utilisé un écran suffisamment grand pour représenter la planche de bord à l'échelle 1. L'écran est très proche de l'observateur pour que celui-ci puisse voir les images en relief sans fatigue et avec les bonnes perspectives. Sa tête est traquée pour qu'il puisse voir la planche de bord sous différents angles en temps réel, [Moreau, 2001].

Pour résumer, avant de choisir ou concevoir une interface visuelle, il faut déterminer l'acuité visuelle, les champs de vision, la perception de profondeur (vision stéréoscopique ou non), l'immersion du regard ou non (visiocasque) et la fréquence d'affichage nécessaires à l'application suivant les I² sensori-motrices et mentales. Ceci est développé au paragraphe 10.4.

10.2.5.2 Visiosalles ou salles immersives

Une visiosalle (ou salle immersive) est une interface de visualisation composée principalement d'une station graphique puissante et d'un ensemble de vidéoprojecteurs (trois en général) permettant l'affichage d'une image à haute résolution *sur grand écran, plat ou semi-cylindrique*. Elle permet ainsi à quelques dizaines de personnes de regarder simultanément des images de synthèse générées en temps réel. De telles salles peuvent compter plusieurs dizaines de sièges et sont pour certaines démontables, ce qui permet de les installer, par exemple, sur les sites d'expositions. Elles sont utilisées pour visualiser un produit ou un ensemble complexe et permettent ainsi de prendre des décisions quant à la réalisation d'un projet à partir de sa maquette virtuelle. Par exemple, les observateurs peuvent se déplacer dans un complexe architectural comme dans la réalité, les images étant alors à l'échelle 1. L'emploi de plusieurs vidéoprojecteurs oblige d'utiliser la technique du « edge blending » pour créer une seule image. Cette technique consiste à superposer les bords des images connexes et à diminuer l'intensité



Figure 10.17 : Dispositif de visualisation pour l'Immersion Adaptée à un Véhicule, IMAVE (photo Ecole des Mines de Paris)

lumineuse des zones de recouvrement pour que l'intensité additionnelle des deux vidéoprojecteurs égale l'intensité des zones de non-recouvrements. Si ceci est bien réglé pour les trois couleurs RVB des vidéoprojecteurs tritubes, les recouvrements d'images sont imperceptibles, l'observateur ne voit qu'une image. Un test révélateur d'un « edge blending » bien réglé est le suivant : il faut obligatoirement éteindre un vidéoprojecteur pour que les observateurs s'aperçoivent que l'image est créée avec trois vidéoprojecteurs. Pour un réglage performant en temps réel, Barco place trois caméras dans chaque vidéoprojecteur pour ajuster les réglages du « edge blending ». Car le lecteur doit savoir que le principal défaut des vidéoprojecteurs tritubes est leur manque de stabilité dans le temps de leur réglage et donc ceci impose une maintenance assez contraignante.

Des premiers types d'interfaces visuelles à la limite entre les bureaux immersifs et les salles immersives sont le VRCOVE de VREX et le Reality Center 3300W de SGI. Le VRCOVE offre à un *seul* utilisateur une immersion sur 180° en utilisant 3 écrans stéréoscopiques entourant l'utilisateur (Figure 10.18).



Figure 10.18 : VRCOVE (Virtual Reality COmmand Visualization Environment),
copyright photo VREX

Le Reality Center 3300W offre une image plane à hauteur de la tête des observateurs. Vu sa structure, celui-ci est transportable et est nettement moins encombrant qu'une vraie salle immersive. L'image est de 2,7m par 1,8 pour le « two channel » et de 3,8m par 1,8 m pour le « three channel », (Figure 10.19).

Plusieurs constructeurs de salles immersives existent et proposent deux principaux types de configuration *sur grand écran, plat ou semi-cylindrique*. Dans le cas d'un écran semi-cylindrique, le champ de vision recouvert pour l'observateur central est de 160 à 180° en général (Figure 10.21). A noter que l'image n'est calculée et exacte que pour l'observateur central, les autres observateurs ont une vision légèrement déformée de la scène virtuelle. Cela n'est perceptible que dans certains cas et pour des yeux aguerris. L'écran plat se justifie plus qu'un écran courbe pour certaines scènes observées : par exemple, le cas classique d'une représentation d'un véhicule, où il vaut mieux un écran plat partant du niveau du sol. Les structures de salles immersives sont similaires. Leurs différences portent principalement sur leur taille et donc le nombre de sièges, la possibilité d'images stéréoscopiques, leur puissance lumineuse (peut-on avoir un éclairage complémentaire dans la salle pour prendre des notes ?) et bien sûr la puissance de la station graphique. Une autre différence technique est le type de projection : frontale ou rétroprojetée. Dans ce dernier cas, l'observateur peut être proche de l'écran sans projeter son ombre sur celui-ci. Pour les écrans plats, les deux types de

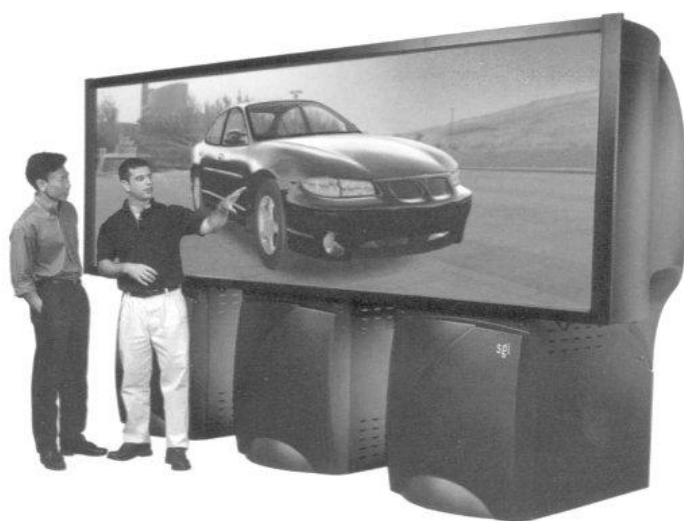


Figure 10.19 : Le Reality Center 3300W à trois canaux de SGI, copyright photo SGI

projection ne posent pas de contraintes. Par contre pour les écrans semi-cylindriques, il est rare et plus difficile de faire en rétroprojection. Cela n'est possible que si le rayon de courbure est assez grand.

Un exemple de salle immersive à écran plat : la solution se présente sous la forme d'un écran de très grande taille (de 6'*8' à 8'*24') sur l'arrière duquel on projette l'image à l'aide de 2 ou 3 projecteurs (3456 x 1024 en configuration à 3 projecteurs). Elle permet ainsi de simuler des situations ou l'utilisateur regarde dans une direction fixe par rapport à l'écran (Figure 10.20).

10.2.5.3 Visiocubes

Un autre type d'interface visuelle à taille humaine est le visiocube ou CAVE² : il s'agit ici de faire concurrence aux visiocasques, car on souhaite avec ce type d'interface proposer une immersion visuelle totale : sur 360° horizontalement avec un visiocube à 4 faces et sur 360° horizontalement et verticalement avec un visiocube à 6 faces. Les grandes différences avec les visiocasques sont le champ de vision humain totalement exploité et l'absence d'isolation de la personne sous un casque (intéressant pour les I² mentales). Le fait de permettre à 2 ou 3 personnes d'observer la scène virtuelle affichée n'est pas un atout majeur : si on veut correctement voir dans un visiocube, il faut que l'utilisateur ait la tête traquée pour lui fournir en temps réel le bon point de vue. Autrement, les raccordements entre les différentes images planes ne sont pas visuellement corrects. Un visiocube à 6 faces est donc un ensemble de 6 écrans formant ainsi un cube de 3 mètres de côté environ autour de l'utilisateur qui se retrouve ainsi totalement

²Cave étant un nom déposé, nous devrions employer un nom générique comme visiocube, en comparaison à visiocasque pour les casques RV (HMD), mais la coutume étant bien établie de parler de Cave (prononciation anglaise), il va être difficile de parler de visiocube !

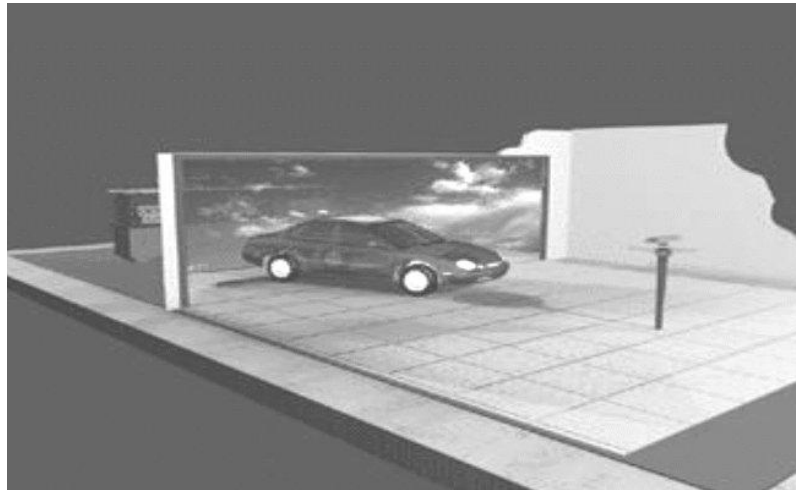


Figure 10.20 : Salle immersive à écran plat Immersive Workwall de Fakespace Systems, copyright photo Fakespace Systems

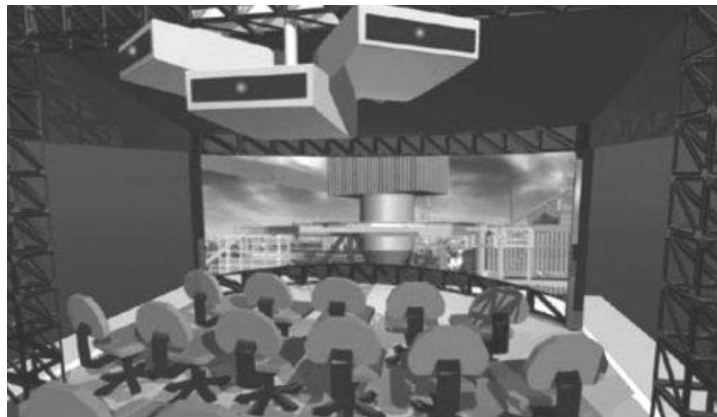


Figure 10.21 : Salle immersive avec écran semi-cylindrique et trois vidéoprojecteurs

immergé dans l'environnement virtuel. Ce système est équipé d'au moins 3 projecteurs (visiocube à 3 faces) qui projettent sur les faces extérieures par réflexion comme le montre la Figure 10.22. Ces images peuvent être projetées en relief, l'utilisateur étant alors équipé de lunettes actives (synchronisation par infrarouge).

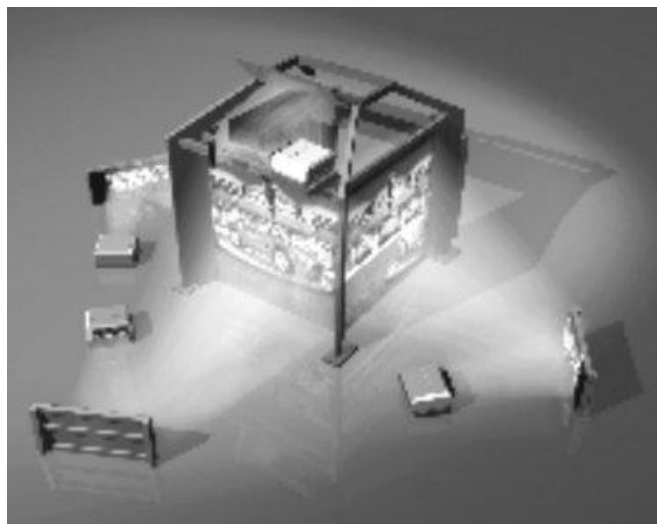


Figure 10.22: Illustration du principe du système CAVE

Jusqu'à présent, les visiocubes étaient reliés à des stations graphiques haut de gamme (Onyx de Silicon Graphics Inc.). Mais il commence à apparaître la technique des « clusters » de PC. En France, à Laval, le SAS³ (prononcer Sas Cube) vient d'être installé et a pour caractéristiques principales d'offrir un excellent rapport performance / prix, d'être mobile, léger et de fonctionner avec différentes plates-formes informatiques. Il peut fonctionner avec une grappe de 2, 4 ou 8 PC ou bien avec la station Onyx de SGI. Constitué d'une structure allégée utilisant des matériaux facilement démontables, le SAS³ est le premier du genre aisément transportable (Figure 10.23).

Le choix entre un visiocube et une visiosalle est un problème qui se pose à bien des industriels. Vu l'importance de l'investissement, ils souhaiteraient souvent avoir les deux configurations. C'est ce que propose Fakespace avec son RAVE (Figure 10.24).

Pour conclure sur ces types d'interfaces, leurs principaux défauts sont le prix très élevé ainsi que leur encombrement : il faut prévoir un local assez vaste, ce qui augmente le prix de l'installation.

10.2.5.4 Écrans de simulateur

Ces dispositifs existent depuis plusieurs dizaines d'années, depuis que l'on fait de la simulation en aéronautique. Cette activité est, comme nous l'avons signalé en introduction, devancière de la réalité virtuelle. Les grands moyens techniques et financiers consacrés à la simulation ont permis de faire des interfaces visuelles présentant une

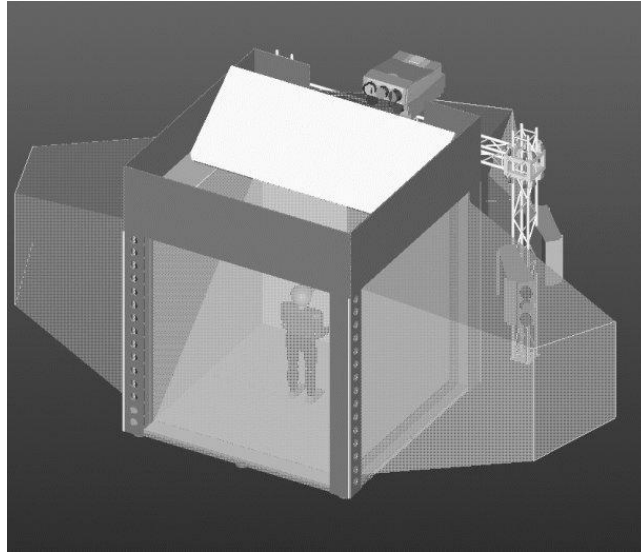


Figure 10.23 : Le SAS³, visiocube transportable et multi plate-forme de Clarté à Laval, copyright photo Clarté

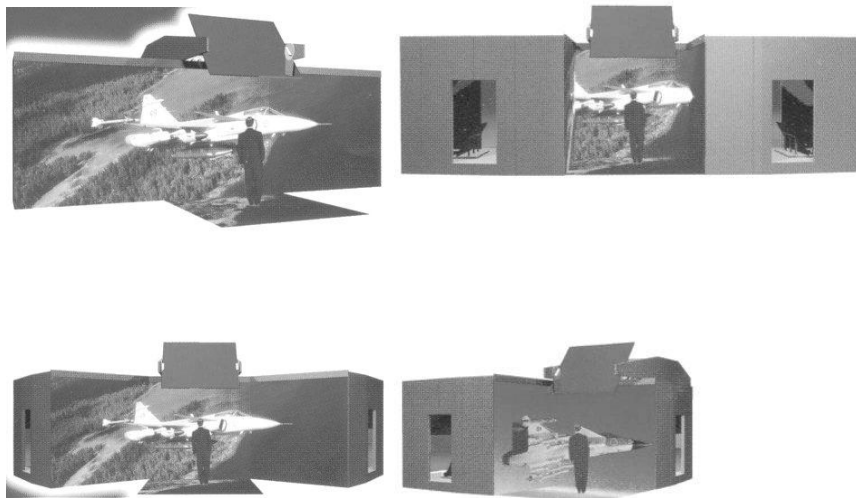


Figure 10.24 : La visiosalle reconfigurable Rave de Fakespace Systems, copyright photo Fakespace Systems

qualité de résolution élevée. Les types d'interfaces visuelles employés dans les simulateurs de transport ont été à l'origine bien spécifiques. Maintenant, avec l'essor de la réalité virtuelle, les simulateurs de transport emploient les mêmes types d'interfaces visuelles que l'on vient de décrire précédemment, principalement les salles immersives. Les simulations de mouvement, en général couplés à la simulation visuelle, sont aussi de qualité (voir le chapitre 11). Une solution particulière est d'afficher les images sur un dôme, technique utilisée pour la simulation aérienne. Le dôme est un écran semi-sphérique d'une dizaine de mètres, l'observateur étant au centre et pouvant regarder sur 360° horizontalement et à 90° vers le haut. Deux projecteurs grand angle affichent sur le dôme les images à partir de faisceaux laser. L'affichage de l'intégralité de la scène n'est pas utile à chaque instant : en détectant le mouvement de la tête, on peut se contenter d'afficher les images à haute résolution dans sa direction et des images à faible résolution à la périphérie.

10.2.6 DIFFÉRENTS TYPES DE VIDÉOPROJECTEURS

Pour les différentes interfaces visuelles que nous venons de présenter, il y a plusieurs types de vidéoprojecteurs en concurrence :

- les vidéoprojecteurs tritubes (CRT) sont les plus anciens et les plus classiquement employés jusqu'à ce jour. Ils ont comme principaux avantages leur qualité d'image et leur haute résolution unitaire : jusqu'à 3200 par 2500 pixels en monoscopique et 1600 par 1200 pixels à 96 Hz en stéréoscopique, pour les très hauts de gamme. Leurs principaux inconvénients sont leur prix, leur puissance lumineuse modérée (250 à 450 lumens) et les dérives de leurs caractéristiques dans le temps qui impose une maintenance encore contraignante. Mais la stabilité de leurs caractéristiques progressent. Trois visites de maintenance par an sont suffisantes, surtout si les vidéoprojecteurs restent constamment sous tension. Leur haute qualité d'image est reconnue surtout pour leur absence de perception des pixels (ils font du balayage), leur noir est vraiment noir, idem pour le blanc, le contraste est excellent et la forme géométrique de leur image est adaptable à tout type d'écran, plat ou courbe ;
- les vidéoprojecteurs LCD sont d'un prix moins élevé en général, mais les prix sont plus importants quand on souhaite une puissance lumineuse élevée (fonction du type de lampe : de 2000 à 12000 lumens) ou une résolution élevée : 1280 par 1024. Cette résolution est la limite maximum qui le restera pendant les prochaines années à venir. Notons que la puissance lumineuse doit être élevée lorsqu'il faut afficher les images pour un groupe d'observateurs en salle immersive. Mais pour une utilisation mono observateur en RV, la puissance souhaitée est en général faible. Les vidéoprojecteurs LCD ont pour principal défaut de fournir des images avec des pixels rectangulaires qui laissent apparaître entre eux des interstices dégradant la qualité d'image (bien sûr, si l'application impose l'observation de prêt des écrans, ce qui est assez souvent le cas en RV). Le contraste de l'image est de 300 à 350. Les couleurs sont de moins bonnes qualités qu'avec les CRT, surtout pour le noir et plus difficilement ajustables (très utile pour certaines applications RV, comme la conception esthétique). Mais certains constructeurs, comme Barco, sont capables de mieux gérer les couleurs sur les LCD. Ils sont aussi capables, comme avec les CRT, de faire du « edge blending » (fusion des bords d'images) pour faire une grande image composée de plusieurs images unitaires de vidéoprojecteurs LCD ;
- les vidéoprojecteurs D-ILA utilisent des LCD, mais la création des images se fait par réflexion et non par transmission. Par rapport aux LCD, les interstices entre pixels

sont moins visibles, ce qui est appréciable ;

- les vidéoprojecteurs DLP, basés sur la modulation temporelle de micro miroirs, ont des caractéristiques plus comparables aux LCD. Ils ont la capacité de fournir des images à haute puissance lumineuse. Le contraste de l'image est de 400 à 450. Les couleurs sont de moins bonnes qualités qu'avec les CRT, surtout pour le noir, mais un peu mieux que les LCD.

Pour la vision stéréoscopique, faite avec les LCD ou les DLP en technologie passive, il faut avoir deux vidéoprojecteurs au lieu d'un seul CRT à double fréquence en technologie active (les prix sont à comparer). Un problème complémentaire se pose alors : la station graphique doit avoir deux canaux graphiques à fréquence normale pour les brancher sur les deux vidéoprojecteurs. Une autre solution, proposée par Barco avec le dispositif DUET est d'exploiter un seul canal à double fréquence pour le séparer en deux et le connecter à deux vidéoprojecteurs passifs. Le signal vidéo, n'étant pas digitalisé, garde sa qualité intrinsèque et les sources logiciels pour la stéréo active peuvent être exploitées dans ce cas, sans modification.

Pour la projection, le choix de l'écran est aussi important. Plusieurs types d'écran sont proposés commercialement. Dans le cas de la vision en relief avec polarisation, l'écran doit garder les polarisations après réflexion, ce qui est plus difficile à obtenir par rétro-projection. Si on observe l'écran de près, les couleurs peuvent être perçues différentes si l'angle de vision est grand par rapport à la normale de l'écran, sauf pour les écrans à lentilles de Fresnel qui permettent de voir en rétroprojection même à 90° de la normale au plan de l'écran. Pour les écrans courbes, la régularité de la surface doit être soignée pour garder une bonne qualité visuelle.

Les problèmes de déformation des images à projeter sur des écrans courbes ne sont pas simples à résoudre, surtout pour les dômes et pour afficher en rétroprojection sur un écran semi-cylindrique. L'entreprise Perception 3D a développé un système de déformation des images en temps réel, qui se fait matériellement (en «hard»), donc sans décalage temporel. Le problème dans ce cas est de garder la même qualité d'image en chaque zone après déformation.

10.3 INTERFACES VISUELLES PORTABLES

Grâce à l'essor de la réalité virtuelle, il est apparu un nouveau type de matériel pour la vision : les visiocasques (traduction de «Head Mounted Display», on peut aussi employer l'expression «casque immersif») dont l'objectif est de fournir une **vision stéréoscopique** à l'aide de deux petits écrans, un **grand champ visuel** correspondant à celui des yeux de l'utilisateur et **une immersion du regard**. Le visiocasque tend théoriquement (mais pas pratiquement !) vers «l'interface visuelle idéale» pour un seul observateur. En plus de la vision stéréoscopique, partielle ou totale, ces appareils disposent d'un capteur de localisation de la tête de l'utilisateur permettant d'afficher les images correspondant à l'orientation des yeux. La direction du regard est par définition la projection des axes des yeux sur le paysage. C'est donc le résultat de l'orientation du corps, de la tête et des yeux. Donc, dans un visiocasque, les images affichées correspondent à l'orientation de la tête mais pas à la direction des yeux, car aucun visiocasque commercialisé à ce jour ne détecte le mouvement des yeux. Cependant, dans la plupart des cas, il s'agit bien de la direction du regard. En effet, le mode de prise d'information visuelle se fait de la manière suivante : dans un premier temps, il y a rotation des yeux

dans leur orbite puis ensuite légère rotation de la tête pour mettre les yeux au milieu de leur orbite. Le champ de vision est théoriquement de 360 degrés en considérant la mobilité de la tête. En position fixe de la tête, le champ de vision d'un visiocasque n'est pas aussi grand que naturellement (ordre de grandeur : 120 degrés maximum au lieu de 180 degrés dans le plan horizontal), mais donne néanmoins une impression d'immersion visuelle. En vision naturelle, la perception n'est nette que dans le champ central. Si on observe un objet sur le contour du champ visuel, on tourne la tête instinctivement pour mieux le regarder. Les visiocasques sont donc moins efficaces pour attirer le regard sur le côté, leur champ de vision étant plus étroit. Dans certains visiocasques à champ de vision trop faible, on a la désagréable impression de voir à travers un tube. D'où le dilemme des constructeurs : avec un même type d'écran, si on augmente le champ de vision pour percevoir les objets de côté en écartant les deux écrans, on diminue la zone de superposition des images droite et gauche et donc la perception du relief. En complément, le son stéréophonique est en général proposé grâce à deux écouteurs.

10.3.1 ARCHITECTURE D'UN VISIOCASQUE

Le visiocasque doit présenter quatre fonctionnalités principales :

- un grand champ visuel ;
- une immersion du regard ;
- une vision stéréoscopique ;
- une résolution élevée.

Pour avoir **un grand champ visuel**, deux solutions sont possibles : soit un très grand écran, soit deux écrans proches des yeux. Le premier cas est exploité dans les visiosalles. Le deuxième cas est celui exploité dans les visiocasques, mais il implique deux obligations :

- ajouter un dispositif optique à grand champ visuel pour permettre l'accommodation des yeux sur des écrans très proches de ceux-ci ;
- disposer de deux mini-écrans à très bonne résolution ou, à défaut, si la résolution est faible, mettre au moins devant les mini-écrans LCD deux diffuseurs pour ne pas voir les pixels sous forme de gros pavés, séparés par des interstices noirs.

Dans ces deux obligations, les solutions techniques sont souvent opposées : par exemple, il n'est pas facile de fabriquer des écrans de très petite taille, constitués de nombreux pixels. Il est difficile d'obtenir un grand champ de vision avec des mini-écrans.

Pour **une immersion du regard**, il faut avoir un capteur de localisation et fixer les deux mini-écrans avec les optiques sur la tête. L'utilisateur en tire profit si l'ergonomie de l'interface est suffisante pour qu'elle devienne transparente. Ceci impose que les dispositifs soient très légers, confortables (bonne ventilation) en respectant des règles de sécurité. Le problème du choix des optiques est primordial. Pour projeter les images des mini-écrans, il faut des optiques de fort grossissement, mais ceci implique des distorsions géométriques, surtout si on désire de grands champs de vision. La vision

en relief exige aussi que les optiques soient bien positionnées pour faciliter la fusion des images. Il y a plusieurs solutions pour les optiques : soit avec plusieurs lentilles classiques, soit avec des lentilles de Fresnel, soit avec le dispositif appelé «Pancake window», les deux derniers cas permettant de diminuer le poids. Ce dernier système est basé sur un miroir sphérique semi-transparent et sur des polariseurs supprimant les rayons lumineux parasites. Un dispositif compact et léger a été conçu par la société Saab Avionics pour le visiocasque AddVisor, (Figure 10.25).



Figure 10.25 : Visiocasque AddVisor de Saab Avionics, copyright photo SimTeam

Pour obtenir **la vision stéréoscopique**, il suffit théoriquement d'avoir deux écrans et d'y afficher les bonnes images pour chaque œil. Vu les difficultés de fabrication des mini-écrans à très grand champ de vision, les constructeurs préfèrent augmenter le champ de vision total (des deux yeux) en ne créant qu'une zone réduite de recouvrement stéréoscopique (overlap). Celle-ci est souvent réglable sur les visiocasques (valeurs acceptables : de 100% à 70%). Un constructeur a proposé un recouvrement de 25%, ce qui ne laisse pas beaucoup de possibilité de voir en relief !

Nous pouvons résumer ces différentes contraintes techniques pour la conception d'une interface de vision individuelle (visiocasque) en quatre étapes dans l'ordre de difficultés croissantes (Figure 10.26).

Si on sait fabriquer des visiocasques à immersion visuelle grâce à l'incorporation d'un traqueur, la vision stéréoscopique est plus difficile à réaliser à 100%, surtout si on veut

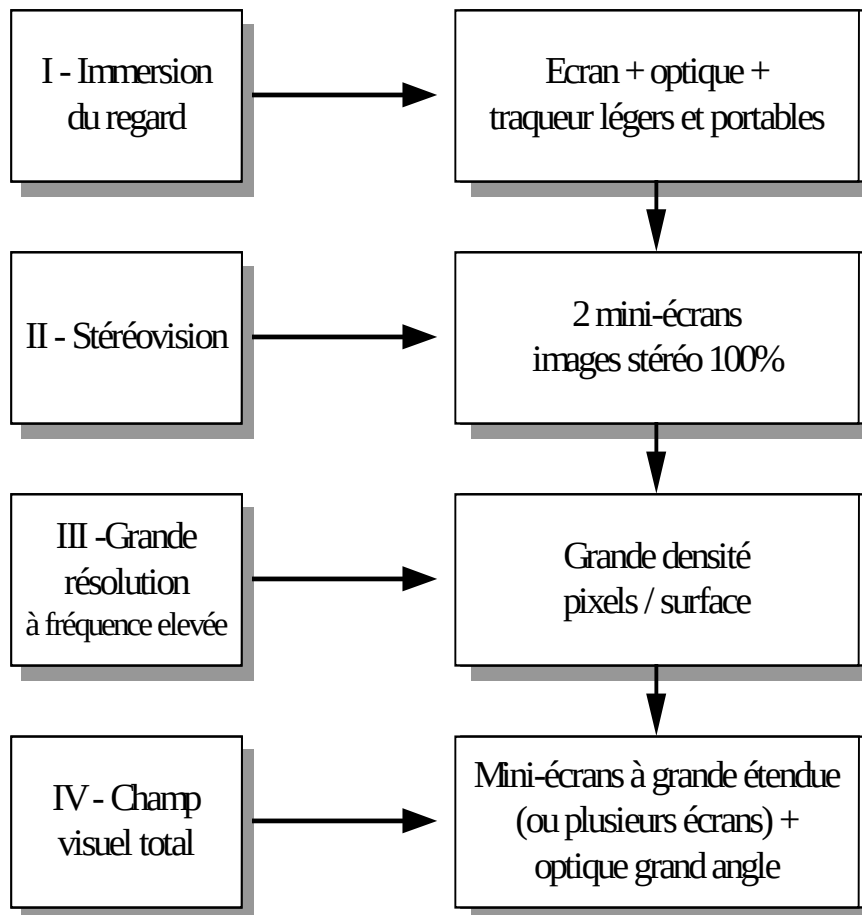


Figure 10.26 : Contraintes technologiques pour la conception d'une interface de vision individuelle

en même temps une très grande résolution : on n'est pas techniquement en mesure de répondre à ces exigences actuellement. Les difficultés augmentent encore si on souhaite se rapprocher du champ visuel humain. Les visiocasques sont proposés souvent avec un traqueur à six degrés de liberté et quelques fois pour des dispositifs moins chers avec seulement 3ddl (les orientations). Il ne faut pas forcément exploiter les six degrés de liberté, car cela peut dépendre des besoins qui sont analysés au niveau des I² sensori-motrices.

L'évolution des visiocasques est très lente et plus lente que ce que nous espérons. Les difficultés techniques, surtout optiques, sont plus importantes et surtout le marché de ce type d'interface est limité. La demande n'est donc pas moteur dans l'évolution des visiocasques. Les entreprises qui fabriquent les visiocasques sont souvent de petites sociétés.

10.3.2 VISIOCASQUES AVEC ÉCRANS À TUBE CATHODIQUE

Il existe actuellement deux grandes familles de visiocasques en se référant aux types d'écrans qu'ils utilisent : soit les écrans cathodiques miniatures, (dimension de la diagonale de l'écran : 1" environ), soit les écrans à cristaux liquides ou LCD (Liquid Crystal Display). Il existe aussi la possibilité d'amener les images par des fibres optiques à partir d'écran de haute résolution. Ces derniers types de visiocasques sont relativement chers et ont été développés principalement pour des applications aéronautiques [Shenker, 1987].

L'avantage principal de la technique des écrans cathodiques est la résolution élevée de l'affichage (jusqu'à 1280x1024 pixels par œil, mais attention, la fréquence des images peut être réduite de 60 à 30 Hertz dans ce cas). Les inconvénients en sont le poids, la taille (du tube), la consommation de ces écrans et surtout le prix, comparativement aux écrans à cristaux liquides (Figure 10.27). Un problème important actuel est l'impossibilité de fabriquer des écrans cathodiques miniatures *couleurs*. Devant les deux écrans sont donc placés deux obturateurs électroniques à dynamique très rapide (Liquid Crystal Shutters), qui permettent de filtrer alternativement une des trois couleurs rouge, vert et bleu. Les écrans cathodiques, ayant une fréquence de balayage de 180 hertz, sont synchronisés avec les obturateurs de telle façon que les trois couleurs RVB soient affichées à la fréquence de 60 hertz. On obtient ainsi à l'intérieur du casque des images stables en couleur grâce à la persistance rétinienne (Figure 10.28). Un constructeur a conçu un visiocasque (Eyegen3) à écrans cathodiques avec deux obturateurs mécaniques pour diminuer le coût du produit. Comme nous l'avons déjà signalé, un système optique performant doit être placé entre les écrans et les yeux, permettant une vision correcte.

La Distance Intra Oculaire (DIO) est en général réglable pour s'adapter à celle de la personne, de même que le réglage de la netteté des images. Le capteur de localisation, les écouteurs, un microphone éventuellement, et une partie de l'électronique sont incorporés au casque, tout en essayant de ne pas trop l'alourdir.

Cas particuliers :

Face au poids important de ce type de visiocasque, certains constructeurs pensent qu'il est préférable d'accrocher le visiocasque à un portique. C'est l'utilisateur qui manuellement oriente le dispositif de vision. Celui-ci est suspendu à un ensemble de leviers, qui

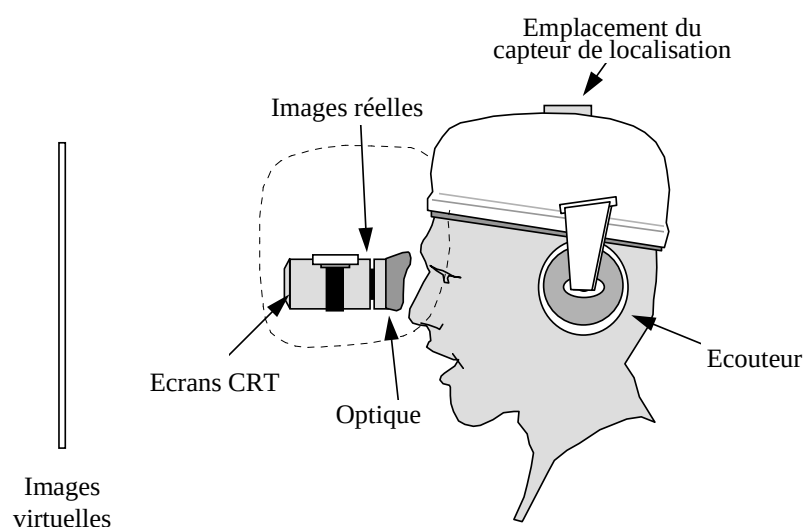


Figure 10.27 : Schéma de principe d'un visiocasque à écrans CRT

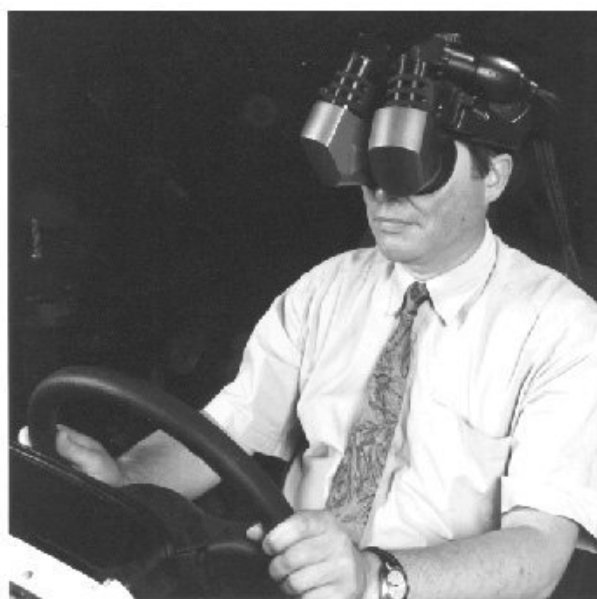


Figure 10.28 : Visiocasque à écrans CRT Datavisor 80 de n-vision, copyright photo PSA Peugeot Citroën

lui permet une orientation quelconque dans l'espace, tout en mesurant celle-ci grâce à des capteurs angulaires aux articulations. Un contrepoids permet un équilibrage mécanique du système pour que la personne puisse manipuler sans peine le visiocasque (Figure 10.29). Cette solution a un autre avantage : les mesures mécaniques de localisation de la tête sont bien plus performantes qu'avec les autres traqueurs : le temps de latence est de 2 ou 3 ms et la précision est très élevée.



Figure 10.29 : Système BOOM HF de Fakespace Labs, copyright photo Fakespace Labs

Un autre support mécanique permet de poser l'interface visuelle sur un bureau, tout en permettant l'orientation manuelle du regard (Figure 10.30). Cette interface a pour principal avantage d'avoir une très bonne résolution graphique (écrans cathodiques miniatures : 1280 x 960 pixels par œil), avec un champ de vision de 30° à 140° suivant les optiques. Le principal inconvénient est la liaison mécanique entre l'utilisateur et l'environnement et accessoirement l'occupation des mains, sauf si le système est fixé à la tête.

10.3.3 VISIOCASQUES AVEC ÉCRANS À CRISTAUX LIQUIDES

Les principales différences entre cette catégorie de visiocasques et ceux à écrans cathodiques sont leur plus faible prix et leur plus faible résolution des écrans LCD, même si certains modèles atteignent néanmoins une résolution de 1024 par 768 pixels pour le XL50 de Kaiser Electro Optics et 1280 par 1024 pixels pour le AddVisor de Avionics Saab. Ce dernier visiocasque, avec sa résolution SXGA, est une offre commerciale très intéressante, si elle est confirmée (Figure 10.25). La résolution des visiocasques à écrans LCD étaient si faible au début des années 90 que les constructeurs donnaient parfois la résolution en nombre de pixels élémentaires R, V ou B et non pas en pixels trichromiques, comme on l'indique sur les écrans cathodiques des visiocasques ou des ordinateurs. Il fallait donc bien lire les notices et parfois diviser par 3 le nombre total de pixels. Les visiocasques à cristaux liquides sont moins lourds et moins encombrants



Figure 10.30: Système PUSH 640 de Fakespace Labs, copyright photo Fakespace Labs

(la dimension de la diagonale de l'écran est inférieure à 3"). A noter qu'une entreprise, Kaiser Electro Optics, a proposé un visiocasque avec deux écrans à cristaux liquides pour chaque œil pour pallier leur principal défaut : leur faible résolution. La difficulté résidait dans la jonction entre les deux écrans qui n'était pas parfaite graphiquement. Ce type de visiocasque ne semble plus être commercialisé. Des recherches importantes sont entreprises actuellement pour augmenter le nombre de pixels sur chaque écran de type LCD, mais les résultats tardent à venir.

La structure de ces visiocasques est identique à celle des visiocasques à écrans cathodiques, à part les deux écrans et les deux diffuseurs servant à dépixelliser (Figure 10.31). La vision par des écrans à cristaux liquides à faible résolution donne l'impression d'observer la scène à travers un maillage dû aux interstices entre les pixels. Pour résoudre ce défaut, William Johnson a inventé un film polymère composé de millions de microlentilles, le Microsharp, qui dépixellise l'image sur les écrans miniatures à cristaux liquides. Il y a d'autres principes pour résoudre ce problème : on peut défocaliser en partie les images, mais cela apporte une certaine fatigue. Une autre solution est de mettre devant l'écran un «wavefront randomizer» qui est une collection de petits prismes réfléchissant suivant un angle variable la lumière (petits prismes : plus petits que la dimension d'un pixel) [Teitel, 1990].

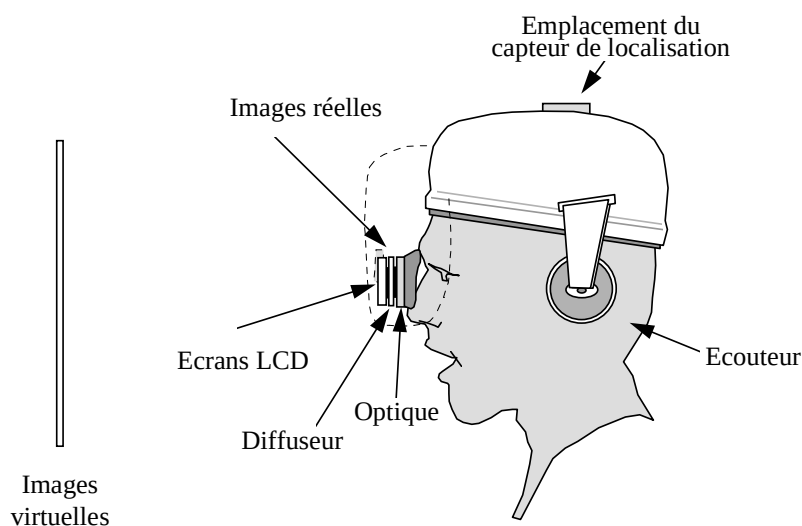


Figure 10.31 : Schéma de principe d'un visiocasque à écrans LCD

10.3.4 MODÈLE OPTIQUE D'UN VISIOCASQUE ET PROBLÈMES CONNEXES

10.3.4.1 Problèmes de la qualité visuelle d'un visiocasque

Une des grandes difficultés des visiocasques est de permettre d'observer des images sur des écrans très proches des yeux et pour cela on doit insérer des optiques entre les yeux et les écrans. Il faut parfaitement déterminer la géométrie et l'optique du visiocasque pour pouvoir calculer et afficher correctement les images stéréoscopiques sur les deux

écrans à cristaux liquides ou cathodiques. Si nous bougeons la tête, un objet fixe est vu sous différents angles et les dimensions de ses images sur les deux rétines varient. Mais notre cerveau perçoit sans effort cet objet comme étant fixe. Le visiocasque ne doit pas perturber cette perception visuelle des objets immobiles. Les images affichées sur les écrans du visiocasque doivent donc donner les mêmes sensations visuelles que lorsque l'on regarde un monde réel, ou du moins s'en rapprocher. Ceci ne dépend pas que des caractéristiques du visiocasque, mais aussi de la qualité du capteur de localisation de la tête et des algorithmes calculant les images à afficher. L'utilisation d'un visiocasque nécessite donc une localisation précise des deux écrans dans l'espace, ainsi que la position de ces écrans par rapport aux yeux. Plusieurs problèmes sont alors à distinguer :

- positionnement du capteur sur le visiocasque : il faut d'abord être capable de fixer le capteur sur le casque en un emplacement déterminé de celui-ci ;
- positionnement du visiocasque sur la tête : le casque doit être correctement positionné sur la tête de l'utilisateur, c'est-à-dire qu'on doit être certain que l'on affiche les images du bon point de vue ;
- localisation de la tête de l'utilisateur : enfin, la position et l'orientation du visiocasque dans l'espace doivent être connues avec précision afin de générer les bonnes images. La précision des capteurs électromagnétiques est-elle suffisante ? Il est toujours possible d'envisager l'utilisation de capteurs mécaniques donc la précision est bien meilleure et le temps de latence inférieur. Il faut cependant se demander si la gêne occasionnée par l'utilisation de capteurs mécaniques ne vient pas s'ajouter à celle provoquée par le casque, sauf à utiliser un système monté sur un bras articulé avec contrepoids.

Les images doivent suivre les règles de la perspective, des dimensions relatives et de la vision stéréoscopique. Ceci ne peut être déterminé qu'en connaissant les caractéristiques des systèmes optiques, de leur position par rapport aux deux écrans et aux yeux. Un modèle optique est présenté, comme exemple, en annexe à la fin du chapitre.

Plusieurs sources d'erreurs existent dans les modèles optiques. Certains utilisateurs de visiocasques considèrent leurs systèmes optiques comme de simples agrandisseurs sans faire intervenir les défauts des optiques. Ceux-ci créent des aberrations (sphériques, astigmatisme, distorsion,...). Il est préférable d'en tenir compte dans le modèle optique pour pouvoir corriger les images par logiciel. En stéréoscopie, les deux images virtuelles doivent être affichées en donnant une convergence correcte des yeux en fonction de la profondeur de l'objet observé. Mais l'accommodation et la convergence ne peuvent être en relation comme dans la réalité, les yeux devant accommoder au niveau des images virtuelles (voir les explications détaillées au paragraphe 10.6). La distance interpupillaire est variable d'une personne à l'autre. Pour éviter ces erreurs, il faut prévoir soit un réglage mécanique, soit des optiques avec sorties assez larges pour être employées par tout le monde. Les deux écrans peuvent avoir des axes optiques faiblement divergents et être décentrés, pour augmenter le champ de vision. Dans ces conditions, il est difficile de déterminer les caractéristiques exactes du champ de vision [Robinett, 1991].

Une fois réglés les problèmes optiques que génère tout visiocasque en réalité virtuelle, se pose les problèmes de l'inconfort psychologique dû à l'isolement dans le visiocasque et des perturbations dues au poids du visiocasque. L'utilisation d'un visiocasque est intrusive, en ce sens que le visiocasque est lourd, que ses mouvements peuvent être

plus ou moins contraints par les câbles vidéo et que le centre de gravité de la tête est déplacé. Plusieurs questions se posent donc :

- le poids du visiocasque a-t-il une influence sur la position de la tête ? Le déplacement du centre de gravité du couple tête-casque peut conduire l'utilisateur à chercher une nouvelle position moins fatigante ;
- les mouvements effectués par l'utilisateur sont-ils les mêmes avec le visiocasque ? Les paramètres entrant ici en compte sont la fatigue, la stabilité du couple tête-casque, les liaisons physiques entre le visiocasque et le générateur d'images.

Il est donc nécessaire de chercher à évaluer le biais induit par le simple port d'un casque, avant même se préoccuper de la vision elle-même, si bien sûr la posture de la tête de l'utilisateur est importante pour l'application (comme, par exemple, l'étude de la visibilité d'une scène routière depuis l'intérieur d'un véhicule). Pour ne pas être considérés comme une gêne réelle, il ressort que les visiocasques doivent être symétriques, être les plus légers possible et avoir un centre de gravité vers l'avant de la tête afin de profiter des muscles dorsaux du cou plus puissants et moins sujets à la fatigue. Des données chiffrées existent mais elles portent essentiellement sur les paramètres dynamiques puisqu'il s'agit d'étudier la résistance d'un pilote casqué lors de situation de vol courant : manœuvres courantes, combat aérien, éjection. Les études au sol portent sur les études d'impact.

Pour étudier tous ces problèmes et caractériser la qualité d'un visiocasque comme interface visuelle, il ne faut pas se contenter de demander à l'utilisateur : voyez-vous les images ? Sauf si on se contente de démonstrations «tape à l'œil» (c'est le cas de le dire !). Nous avons proposé un certain nombre de tests associés à l'utilisation professionnelle d'un visiocasque. Avant d'étudier l'immersion et l'interaction (I^2) mentales et fonctionnelles, il faut valider les I^2 sensori-motrices.

- localisation du visiocasque, deux problèmes majeurs :
 - la précision de la localisation de la tête est-elle suffisante pour l'application ?
 - le temps de latence induit par le capteur et par la création des images est-il acceptable ?
- vision, deux problèmes majeurs :
 - la compatibilité des champs de vision (horizontaux et verticaux des deux yeux, zone de recouvrement) dans le visiocasque en environnement virtuel avec les champs de vision exploités en réel ;
 - utilité, validité et contraintes de la vision stéréoscopique pour l'application.

La vision stéréoscopique est importante à étudier dans la mesure où se pose le problème de la convergence et l'accommodation des yeux vis-à-vis des écrans du visiocasque ainsi que le problème du maximum de décalage toléré (parallaxe angulaire limite de $1,2^\circ$ environ, voir au paragraphe 10.6) entre les deux images droite et gauche. Le Centre de Robotique de l'Ecole des Mines de Paris a élaboré un certain nombre de tests pour visiocasque, présentés au paragraphe de l'évaluation des interfaces visuelles.

10.3.5 VIDÉOLUNETTES

10.3.5.1 Vidéolunettes à écran LCD

Les vidéolunettes sont des interfaces visuelles légèrement différentes des visiocasques car celles-ci permettent de voir les images virtuelles tout en n'étant pas totalement coupé du monde extérieur. En corollaire, l'utilisateur perçoit alors des images sur un écran virtuel, flottant dans l'air, avec une immersion visuelle moins probante. Deux types de vidéolunettes sont proposées. Dans le premier cas, le champ de vision n'est pas entièrement obstrué par le dispositif optique et l'observateur peut voir sur le côté son environnement extérieur. Dans l'autre cas, les lunettes ne disposent que d'un seul écran.

Dans cette dernière solution, la personne peut focaliser son attention sur l'écran avec l'œil correspondant ou sur le monde extérieur avec l'autre œil. Cette technique, mise au point par la société américaine VIDI, permet d'afficher une image de résolution 640x480 pixels qui est vue à trois mètres avec une taille apparente de deux mètres de diagonale. Cet affichage est obtenu à l'aide d'un système d'imagerie monoculaire réfléchissant l'image à moyenne résolution d'un LCD (d'un pouce environ) vers l'œil dominant de l'observateur. Avec un tel dispositif, un opérateur peut visualiser toute information provenant d'un ordinateur portable accroché à sa ceinture (Figure 10.32). De nombreux domaines d'application sont visés comme la maintenance d'appareils ou la consultation de documents dans un atelier. Un dispositif similaire de la même société permet de voir la télévision, au lieu de consulter l'écran d'un ordinateur.

Les vidéolunettes à deux écrans se rapprochent plus des visiocasques si elles peuvent être équipées d'un traqueur. Dans ce cas, l'avantage principal est de ne pas entièrement couper la personne du monde extérieur pour des raisons fonctionnelles liées à l'application ou pour des raisons ergonomiques : moins de risque de vertiges, dispositif plus léger et plus supportable. L'inconvénient est une immersion visuelle moins forte.

Les vidéolunettes sans traqueur sont employées normalement pour l'affichage d'images vidéo monoscopiques, bien qu'elles puissent théoriquement afficher des images stéréoscopiques. La télévision en relief n'existant pas actuellement, elles n'exploitent pas cette fonctionnalité pour la «télévision individuelle». Elles sont normalement équipées d'écouteurs. Les vidéolunettes sont peu onéreuses (quelques milliers de francs maximums) et ont pour principales utilisations les jeux vidéo, les applications ludiques en réalité virtuelle et la télévision individuelle (Figure 10.33).

Des récentes vidéolunettes (ou visiocasque ?) arrivent sur le marché du jeu. C'est ce que propose Siemens avec le «Global Pl@yer» : il s'agit d'un casque immersif, a priori sans traqueur de tête, accompagné d'un volant de commande. L'originalité de cette interface visuelle et du volant est de fonctionner sans câble, mais avec des liaisons radio.

10.3.5.2 Vidéolunettes à technique RSD

Un cas particulier de vidéolunettes est en développement. Au lieu d'être affichées sur des écrans, les images sont créées par deux faisceaux laser balayant les verres des

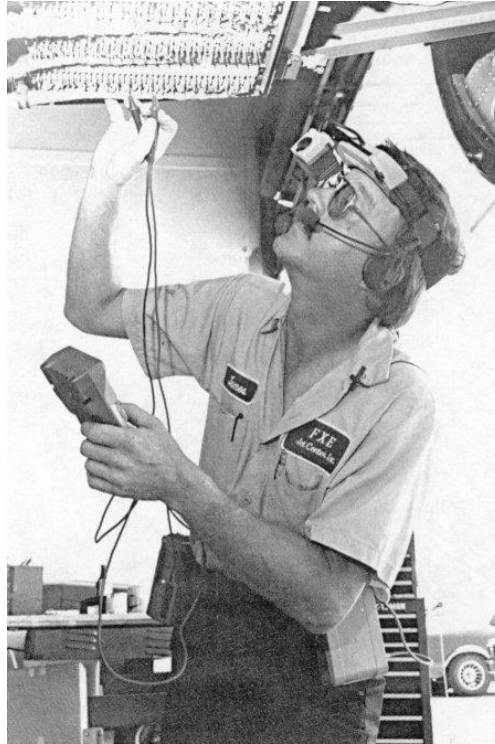


Figure 10.32 : Vidéolunettes à un seul écran pour ordinateur à la ceinture, copyright photo Xybernaut Corporation



Figure 10.33 : Les célèbres Vidéolunettes I-Glasses de Virtual I O, copyright photo ThetaScan

lunettes. Le principe est original : deux diodes laser émettent deux faisceaux qui aboutissent aux photorécepteurs des deux rétines, après réflexion sur les verres. Grâce aux balayages des faisceaux, les images se forment sur les rétines. Des dispositifs optiques contrôlent les balayages horizontaux et verticaux et les rayons sont modulés en intensité. Les avantages du dispositif sont de supprimer les volumineuses optiques comme dans les autres visiocasques et d'avoir la possibilité théorique de haute résolution. Il reste plusieurs difficultés à surmonter dont la couleur qui ne sera possible qu'avec le développement de diodes laser pour le bleu et le vert. Une analyse détaillée peut être lue dans la bibliographie de Holgrem [Holgrem, 1994]. Ces vidéolunettes, qui sont encore à l'état de prototype pour les modèles les plus innovateurs, sont développées par la société MicroVision en partenariat avec le laboratoire de technologie de l'interface humaine de Seattle (HITL) depuis plusieurs années. Elles reposent sur un nouveau principe : la technique RSD élimine l'écran pour projeter l'image directement sur la rétine, permettant ainsi d'obtenir des images plus lumineuses, plus contrastées et avec une meilleure résolution. Cette technique apporte de plus de nombreux avantages :

- très légère, le système peut donc être inclus sur des lunettes ;
- faible puissance consommée (les diodes lasers ne consommant que très peu d'énergie) ;
- beaucoup plus de couleurs en théorie ;
- champ de vision très étendu, plus de 120° ;
- haute résolution espérée.

Fonctionnement :

La technique RSD ou VRD (Virtual Retina Display) utilise un faisceau de lumière modulée qui est appliquée à la rétine, alors utilisée comme un écran de projection. Pour créer ce faisceau et donc l'image on utilise une source de lumière cohérente (trois faisceaux laser modulés dans le cas d'une image couleur) dont on fait correspondre l'intensité à l'intensité du pixel projeté. Le faisceau ainsi créé balaye alors la rétine pour placer chaque pixel de l'image à la position appropriée sur la rétine. De façon à autoriser une source vidéo comme source des images, on utilise un balayage «classique» pour le faisceau de lumière en balayant chaque ligne consécutivement (balayage vidéo normal).

De façon à placer correctement les pixels sur la rétine, il faut de plus que la pupille de sortie du système VRD soit coplanaire avec la pupille d'entrée de l'œil, celui-ci s'adapte alors pour que le faisceau forme un point sur la rétine. Le point de la rétine éclairé à tout instant dépend de l'angle d'entrée du faisceau qui varie constamment suivant le balayage, l'intensité reçue en ce point dépendant quant à elle de l'intensité du faisceau. Une électronique de commande permet de stabiliser l'image en synchronisant le balayage vidéo avec le balayage de la rétine.

Ces vidéolunettes sont apparues comme une solution viable pour la réalité virtuelle seulement très récemment avec l'apparition de diodes laser bleues et vertes de faible puissance, qui permettent avec une diode rouge de générer quasiment toutes les couleurs existantes. En effet, la nature monochromatique des sources permet d'avoir des couleurs de bases extrêmement pures. Par modulation des trois faisceaux lasers, on reconstitue alors n'importe quelle couleur.

Sécurité : S'il semble montré que cette technique est la plus prometteuse de toutes celles utilisées, il faut néanmoins s'intéresser au problème que pose l'utilisation d'un

laser directement sur le fond de l'œil. Néanmoins, l'utilisation de laser à basse puissance, tel qu'utilisé avec cette technique, ne pose pas de problèmes vis-à-vis de la norme de sécurité américaine. Actuellement, la société MicroVision commercialise un dispositif de réalité augmentée utilisant la technique RSD en monochrome (rouge), résolution 640x480 (VGA, une version SVGA sera disponible courant 2000), et en 64 niveaux de couleurs. Ce dispositif ne pèse que 657 grammes avec une alimentation sous 12 Volts en continu.

10.3.6 VISIOCASQUE ET DISPOSITIF SEMI TRANSPARENTS

Il existe des visiocasques à écrans cathodiques qui permettent de superposer à la vision du monde réel des images de synthèse (See-through HMD ou Visuel de casque). Ceux-ci, relativement onéreux, ont été conçus pour certaines applications, comme l'affichage simultané des tableaux de bord et de la vue extérieure pour les pilotes d'avion [Chung, 1989]. Quelle que soit l'interface, elle repose soit sur un écran semi-réfléchissant sur lequel est envoyé une ou plusieurs images qui se superposent à la vision humaine classique. Le principe de l'écran est utilisé dans les avions de chasse «anciens» sous la forme d'un petit écran inclinée sur lequel sont projetées les informations nécessaires au pilotage. Pour la voiture du futur, le pare-brise servira d'écran réflecteur sur lequel seront projetées les informations de vitesse, compte-tours, ...mais aussi les informations de guidage du véhicule. On peut aussi imaginer utiliser ce système pour remplacer les rétroviseurs en projetant l'image filmée par des caméras externes. Un cas particulier de dispositif immersif est le système semi-transparent de ReachIn Technologies qui associe un miroir semi-transparent avec un moniteur (Figure 10.34).



Figure 10.34 : Bureau immersif avec écran semi-transparent de ReachIn Technologies, Copyright photo SimTeam

10.4 CRITÈRES DE CHOIX D'UNE INTERFACE VISUELLE

Le choix d'un type d'interface pour une application donnée dépend de critères **fonctionnels, techniques et économiques**. Comme nous l'avons indiqué dans notre approche théorique de la réalité virtuelle, suivant les I^2 fonctionnelles souhaitées pour l'application, il faut proposer à l'utilisateur des PCV spécifiques (Primitives Comportementales Virtuelles). Celles-ci sont accomplies par les utilisateurs selon des processus mentaux basés sur des schèmes ou des métaphores. Pour les réalisations effectives des PCV, le concepteur du dispositif RV doit donc choisir les interfaces sensorielles et motrices adéquates. La solution n'est pas simple à déterminer dans bien des cas. Le choix de tel type d'interface sensorielle peut être lié au choix de l'interface motrice ou réciproquement. Par exemple, le choix d'un visiocasque, pour la vision, associé à un tapis roulant, pour la locomotion, n'est guère envisageable car l'utilisateur ne peut pas marcher normalement sur le tapis s'il ne voit pas ses propres pieds. Au-delà des incompatibilités entre interfaces comportementales, nous pouvons donner des critères de choix qui ne sont que des indications pour le concepteur, celui-ci devant analyser plus précisément suivant l'application, suivant les I^2 mentales et les I^2 sensori-motrices. Par exemple pour la vision, après analyse des I^2 sensori-motrices, nous devons pouvoir déterminer l'acuité visuelle minimum exigée, les champs de vision horizontal et vertical minimums et la nécessité ou non des images en relief et de l'immersion du regard. Tout cela est possible si l'analyse suivant notre démarche (ou une autre similaire) a été faite précisément.

La nécessité ou non des images en relief est une question qui doit toujours être posée (et que l'on nous a souvent posée). Mais beaucoup trop de concepteurs de dispositifs RV utilisent les images en relief sans raison, parce que l'on leur propose techniquement et souvent parce qu'ils pensent que cela fait mieux. Au delà de l'effet un peu spectaculaire des images en relief, il faut faire le choix de celles-ci que sur la base de raisons sérieuses. N'oubliez pas que ce type d'images procurent au minimum des fatigues oculaires et des rejets pour quelques pour cents de la population qui ne peuvent fusionner les images en relief. Les images en relief ont leur utilité quand nous devons observer des environnements proches (quelques mètres maximum) et que l'application exige :

- soit les positions exactes des deux yeux, par exemple, étude de la visibilité depuis l'intérieur d'une voiture ;
- soit une très bonne perception de la profondeur, par exemple, l'observation des volumes et des formes exactes d'une planche de bord, vue de la position du conducteur.

La création et les contraintes des images en relief sont exposées dans le paragraphe 10.6. Ceci étant introduit, que pouvons nous donner comme critères généraux de sélection d'une interface visuelle ? L'ordre proposé des critères de sélection peut paraître discutable (et il l'est !) et dépend bien sûr des possibilités techniques actuelles.

Première étape : élimination de certains types d'interfaces

Nous avons des critères fonctionnels (principalement d'élimination de solution), basés principalement sur les I^2 mentales et les caractéristiques de la PCV retenue.

(Notation : PES = Pas d'élimination de Solution ; c'est à dire que le critère n'intervient pas pour éliminer certains types d'interfaces visuelles).

Critère N° 1 : faut-il l'immersion totale du regard, pour observer autour de soi ?

- si oui : solution de type visiocasque, visiocube (CAVE) ou fenêtre immersive ;
- si non : PES ;

Critère N° 2 : faut-il l'immersion concentrique du regard, pour observer un objet ou une scène sous différents angles, pour mieux percevoir la profondeur ?

- si oui : pas de fenêtre immersive ;
- si non : PES ;

Critère N° 3 : faut-il l'échelle 1 ?

- si oui : pas de moniteur ou de bureau immersif si la scène réelle est «trop grande» par rapport à la taille des écrans des interfaces précitées ;
- si non : PES ;

Critère N° 4 : l'utilisateur peut-il être isolé visuellement, psychologiquement ?

- si oui : PES ;
- si non : pas de visiocasque ;

Critère N° 5 : l'utilisation de l'interface doit-elle se faire sans ou avec peu d'apprentissage ?

- si oui : pas de visiocasque ;
- si non : PES ;

Critère N° 6 : y a-t-il plusieurs utilisateurs qui doivent être immergés en même temps avec l'interface visuelle ?

- si oui : pas de visiocasque, pas de moniteur. et pas de bureau immersif, si plus de 2 utilisateurs ;
- si non : PES ;

Critère N° 7 : y a-t-il incompatibilité de certains types d'interface visuelle avec les autres interfaces comportementales (motrices ou sensorielles) ?

- si oui : élimination des interfaces correspondantes ;
- si non : PES ;

Deuxième étape : sélection du type d'interface et caractérisation

Cas 1 : Tous les types d'interfaces commercialisées n'ont pas été éliminés

Avec les interfaces visuelles restantes, on fait la sélection de la meilleure solution et des caractéristiques de l'interface suivant des critères techniques et économiques basés principalement sur les I² sensori-motrices.

- vision stéréoscopique nécessaire³, suivant la qualité de perception de la profondeur exigée ou non des images en relief et de l'immersion du regard ;
- tête de l'utilisateur traquée⁴, suivant la qualité de perception de la profondeur exigée ou suivant l'obligation de points de vue différents commandés par la tête, exigée par l'application ;
- éventuellement, nombre de faces d'écran du visiocube, déterminé à partir de l'immersion du regard minimum exigé ;

³Ce critère n'est pas un critère d'élimination de certains types d'interfaces, car pour toutes les interfaces visuelles la vision en relief est possible techniquement à ce jour.

⁴Idem que la note précédente.

- dimensionnement de(s) écran(s)⁵, calculé à partir des champs de vision horizontal et vertical minimums ;
- résolution minimum des écrans, calculée à partir de l'acuité visuelle minimum exigée, monoscopique et stéréoscopique éventuellement ;
- fréquence des images⁶, calculée à partir de la fluidité des images nécessaires ;
- temps de latence⁷ maximum entre l'action virtuelle et la réponse visuelle, déterminé expérimentalement (dépend de la PCV, des I² mentales et sensorielles) ;
- qualités visuelles des images⁸ : luminosité, contraste, rendu des couleurs, absence de perception des pixels ;
- coût de l'interface acceptable : coût d'achat («prêt à voir») ou coût de sa propre réalisation, s'il n'y en pas de solution proposée commercialement répondant aux caractéristiques technico-économiques («sur mesure»).

Cas 2 : Tous les types d'interfaces commercialisées ont été éliminés en étape 1.

- soit vous vous lancez dans votre propre conception d'interface visuelle «sur mesure», bon courage !
- soit vous faites appel à un centre de recherche⁹ pour vous aider dans une conception «sur mesure» ;
- soit vous informez votre hiérarchie que tout n'est pas possible dans ce monde, en particulier en réalité virtuelle, et qu'il faut abandonner le projet d'un dispositif RV.

10.5 TESTS VISUELS SUR LES INTERFACES

Il est souhaitable de faire une série de tests visuels avant d'acheter un produit pour apprécier la qualité d'un appareil. Mais attention, les résultats ne dépendent pas que des caractéristiques intrinsèques de l'interface visuelle mais aussi de la capacité de l'ordinateur en images de synthèse et de votre propre système visuel. Ce dernier est particulièrement prépondérant pour les images en relief.

10.5.1 TESTS VISUELS SUR LES ÉCRANS STÉRÉOSCOPIQUES

Tests techniques :

- détermination du chevauchement temporel des images en relief (crosstalk) : pour ce test, il faut afficher **qu'une** image droite (ou gauche) et fermer l'œil correspondant. Vous ne devez rien percevoir en théorie. La séparation des images n'étant pas parfaite avec les systèmes à lunettes, une image «fantôme» est visible faiblement. Celle-ci

⁵Il peut avoir plusieurs écrans pour composer une même image continue (par exemple, le cas des salles immersives ou de notre magasin virtuel) pour respecter les critères d'acuité visuelle minimum et de champ de vision minimum.

⁶La fréquence des images dépend aussi de la puissance de la station graphique employée et de la complexité du modèle de l'environnement virtuel.

⁷Idem que la note précédente.

⁸Caractéristiques surtout très importantes pour le choix entre types de vidéoprojecteurs.

⁹Le Centre de Robotique de l'Ecole des Mines... par exemple.

doit être la plus atténuée pour ne pas gêner la fusion des images stéréoscopiques. Il faut tester avec les trois couleurs rouge, bleue et verte, en se rappelant que la séparation est plus délicate pour le vert ;

- observation de la qualité de la résolution de l'écran en affichant les lignes verticales d'un pixel d'épaisseur ;
- appréciation des contrastes et de la luminance des images ;
- observation de la fréquence d'affichage des images. Il faut obtenir 25 Hz minimum pour chaque œil. La fluidité des mouvements rapides d'un objet doit être testée.

Tests psychophysiques :

- détermination de la limite de fusion des images stéréoscopiques (angle maximum de parallaxe horizontale). Elle est fonction de la qualité du dispositif et de la capacité de fusion de votre système visuel ;
- appréciation de la qualité de perception de la profondeur avec des images de synthèse (et avec des images réelles, si besoin).

Tests ergonomiques :

- encombrement et poids des lunettes ;
- possibilité de regarder d'autres écrans monoscopiques ;
- fatigue visuelle due à la stéréoscopie.

10.5.2 TESTS VISUELS SUR LES ÉCRANS VISIOPASQUES

Les tests techniques et psychophysiques précédents (sauf le test du «crosstalk») peuvent être faits pour les visiocasques stéréoscopiques. Pour étudier les problèmes et quantifier la validité de l'interfaçage pour chaque utilisateur, le Centre de Robotique de l'Ecole des Mines de Paris a élaboré un certain nombre de tests qui peuvent se décliner en quatre catégories :

- les tests dits préliminaires, qui servent à valider le couple capteur-visiocasque ;
- les tests statiques qui doivent être effectués avant chaque nouvelle utilisation du casque ;
- les tests dynamiques qui sont effectués pour comparer la visibilité en environnement réel et en environnement virtuel, sur des exemples simples ;
- les tests ergonomiques pour étudier le confort et la fatigue du visiocasque.

Tests préliminaires

Il existe un certain nombre de tests à mener avant toute utilisation généralisée du visio-casque.

- répétitivité des réglages IPD¹⁰ ;

¹⁰Distance inter-pupillaire

- qualité et résolution des images affichées (Pour cela, on peut demander à l'utilisateur de lire un texte à l'écran);
- perception des interstices entre pixels, en focalisant ou non votre regard sur les écrans LCD;
- positionnement vertical des écrans (à vérifier en traçant des lignes horizontales sur les écrans à même hauteur et voir si ces lignes coïncident);
- Positionnement des écrans par rapport au champ de vision souhaité et mesure de celui-ci en fonctionnement en affichant des lignes en limites de vision. Il ne faut pas oublier que les optiques sont circulaires et que les écrans sont rectangulaires. Comment sont-ils positionnés entre eux (il y a deux cas principaux) et quelle est l'incidence sur les champs de vision horizontal et vertical ? Comment se fait précisément la zone de recouvrement binoculaire (overlap) ? Peut-on avoir en même temps champ de vision horizontal maximum et zone de recouvrement maximum ? (la réponse est non !). Sur la Figure 10.35, une façon de positionner écran et optique. Et avec les deux écrans et les deux optiques, comment sont-ils positionner visuellement ensembles ? Nous laissons au lecteur rechercher la solution ! Et ne comptez pas toujours sur le constructeur pour vous donner la réponse !

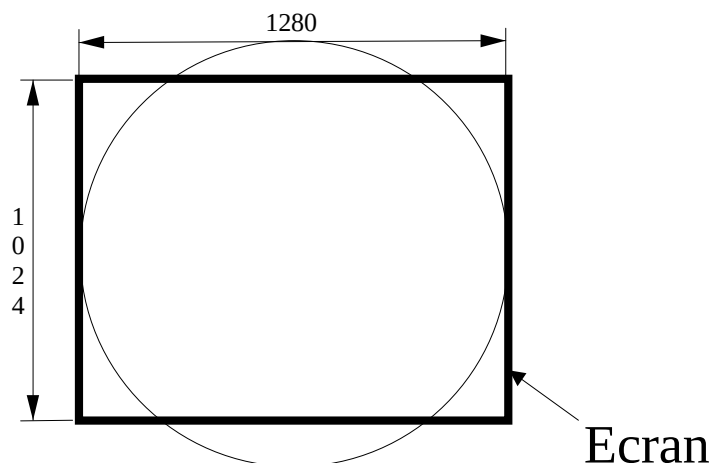


Figure 10.35 : Positionnement des écrans dans le champ de vision

Tests pour l'estimation du recouvrement : faire avec de simples lignes verticales sur fond noir pour chaque observateur, qui doit voir la même ligne un œil après l'autre dans la zone de recouvrement

Tests statiques

- positionnement du casque : on fait le réglage de l'IPD et on demande à l'utilisateur de voir des rectangles blancs affichés sur les bords des écrans : sont-ils entièrement visibles ?
- acuité visuelle monoscopique : par exemple, séparation d'une ligne noire horizontale d'épaisseur 1 pixel entre deux lignes blanches horizontales d'épaisseur 1 pixel;
- acuité visuelle stéréoscopique : détermination de la possibilité de perception de profondeur en vision stéréoscopique sans indices monoculaires.

Tests dynamiques

- immersion visuelle et immersion du regard. Test possible pour l'utilisateur : suivre avec sa tête un objet oscillant sur l'écran. D'autres tests doivent dépendre de l'application envisagée en fonction des I² mentales.

Tests ergonomiques

- encombrement et poids du visiocasque ;
- réglages aisés de la distance intra oculaire, de la netteté et de la fixation de l'appareil sur la tête ;
- fatigue visuelle due à l'immersion et à la stéréoscopie ;
- confort thermique.

10.6 CRÉATION D'IMAGES STÉRÉOSCOPIQUES

Vision stéréoscopique :

Rappelons d'abord que celle-ci n'est pas indispensable pour la perception de la profondeur. En vision monoculaire, en employant les sept règles ou certaines d'entre elles, il est possible de rendre compte de la troisième dimension sur un écran 2D d'ordinateur (de même sur un téléviseur). Certains constructeurs de visiocasques ont préféré proposer la vision sur un seul écran monoscopique pour avoir une résolution double. Cette solution n'est pas à dédaigner tant que la résolution des écrans sur visiocasques n'atteindra pas celle des très bons écrans d'ordinateurs. Dans une application de réalité virtuelle, tout concepteur doit se poser la question : la vision en relief est-elle indispensable ? Est-elle préférable à une vision monoscopique de meilleure résolution ?

Si le choix du dispositif technique permettant l'affichage d'images stéréoscopiques est important, ce n'est pas le seul problème à résoudre pour créer le relief. Bien des personnes pensent qu'une fois le matériel choisi, la création d'images en relief est une question relativement simple à traiter. Nous voudrions montrer dans les paragraphes suivants que les algorithmes pour les images de synthèse ou les paramètres des caméras pour les images réelles sont à choisir judicieusement. Pour les images en relief, des compromis doivent être pris. Ils se basent sur des considérations géométriques, psychophysiques et cognitives. La méconnaissance de ces critères a fait échouer dans le passé bien des projets exploitants les images stéréoscopiques.

Nous rappelons que la vision en relief dépend du moyen technique mis en œuvre mais aussi du système visuel de l'observateur. Et avec très peu de technique, on peut percevoir le relief par une gymnastique oculaire : avec des stéréogrammes, avec de simples dessins ou avec deux photos droite et gauche sur une simple feuille 2D, on perçoit le relief en louchant légèrement et en accommodant sur la feuille. Dans la Figure 10.36, les cercles et les rectangles sont à différentes profondeurs. Le principe est expliqué en annexe à la fin du chapitre.

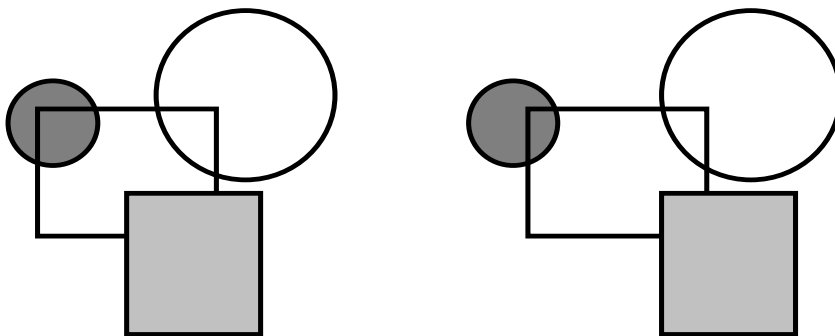


Figure 10.36 : Vision en relief à partir d'une simple image

10.6.1 PRINCIPLE

Il s'agit de créer sur un plan deux images, chacune correspondant à chaque œil, ceux-ci ne voyant qu'une image grâce à un dispositif technique. A un point dans l'espace correspond un point sur le plan pour chaque image. La distance entre ces deux points est appelée parallaxe. Celle-ci produit la disparité rétinienne qui fournit l'effet stéréoscopique de profondeur, chaque œil percevant une image similaire, **mais pas identique**, à celle d'une vision spatiale réelle. Nous insistons sur le fait que les images créées ne sont pas identiques, les images observées par les yeux ne sont que des images planes obtenues par projection géométrique de solides dans l'espace. Ces projections planes induisent des contraintes géométriques et psychophysiques aux images planes que n'a pas la vision spatiale. Nous explicitons ces contraintes dans les paragraphes suivants. La valeur de la parallaxe est fonction de la distance entre l'objet et l'écran, et on a la classification suivante :

Parallaxe positive (non croisée) :

Quand l'objet observé est situé virtuellement derrière l'écran, la parallaxe est positive. Les points homologues gauche et droite sont respectivement affichés à gauche et à droite (Figure 10.37).

Parallaxe nulle :

Quand la partie d'un objet virtuel se situe au niveau de l'écran, les points homologues ont une parallaxe nulle. Ils sont affichés aux mêmes endroits (Figure 10.38).

Parallaxe négative (croisée) :

Quand l'objet observé est situé virtuellement devant l'écran, la parallaxe est négative. Les points homologues gauche et droite sont respectivement affichés à droite et à gauche (Figure 10.39).

Comme nous l'avons écrit précédemment, la parallaxe doit être petite pour ne pas créer de grandes difficultés pour la vision stéréoscopique. Dans le cas contraire, le cerveau est incapable de fusionner les deux images ou il le fait avec une certaine fatigue. L'objectif est donc de créer des images stéréoscopiques avec faible parallaxe tout en gardant un effet de profondeur. Suite aux études expérimentales précédemment décrites, on a

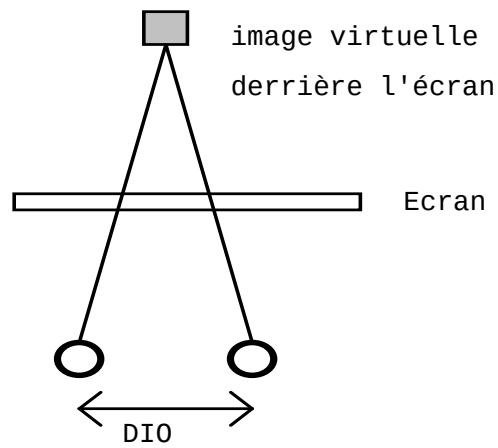


Figure 10.37 : Parallaxe positive permettant de visualiser des objets derrière l'écran ;
DIO : Distance Intra Oculaire (65 mm en moyenne)

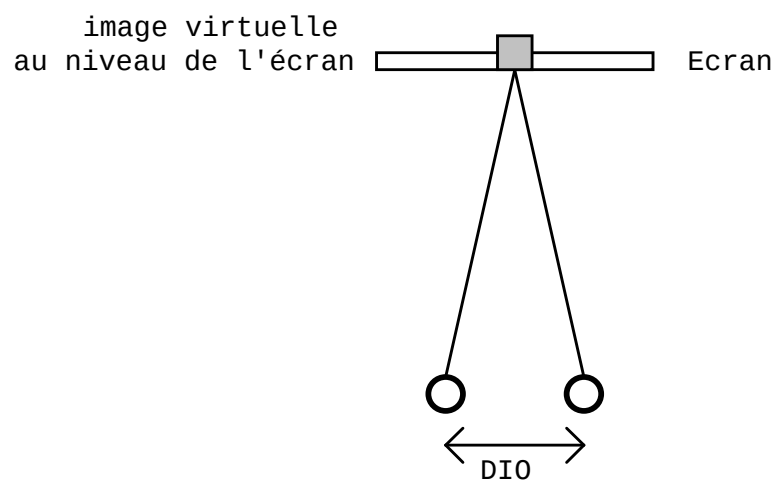


Figure 10.38 : Parallaxe nulle permettant de visualiser des objets au niveau de l'écran

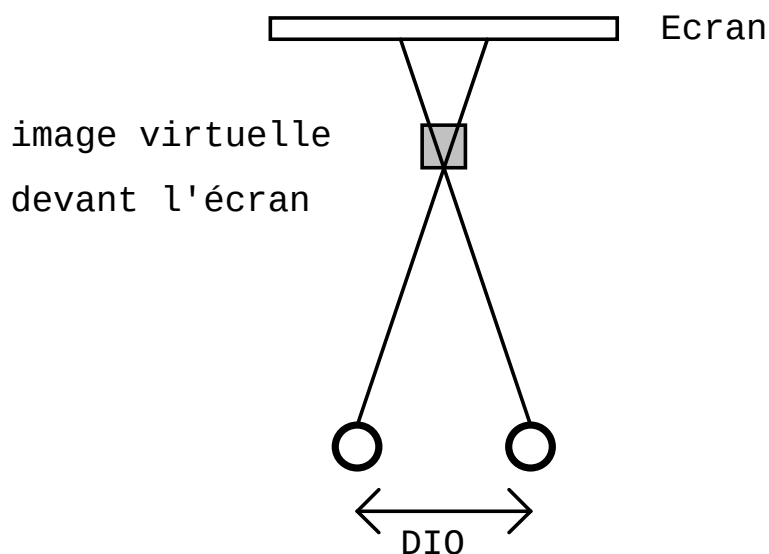


Figure 10.39 : Parallaxe négative permettant de visualiser des objets devant l'écran

montré la difficulté de fusionner deux images planes ayant un angle de parallaxe supérieur à $1,2^\circ$ ou à la limite $1,5^\circ$ pour un stimulus de deux secondes en parallaxe positive. Les conditions sont plus restrictives en cas de mouvements rapides des objets et si les fréquences spatiales des images sont élevées, comme nous l'avons explicité sur le chapitre sur la vision humaine (voir le chapitre sur les sens humains, paragraphe 4.2.5, pour les valeurs numériques). Le problème est plus crucial pour les écrans à faible largeur, l'opérateur doit s'en éloigner suffisamment pour avoir un angle de parallaxe suffisamment petit, mais son champ de vision diminue relativement. L'augmentation d'immersion visuelle créée par la vision en relief peut être annihilée par la diminution du champ de vision.

Nous avons signalé que la relation convergence-accommodation peut être mise en défaut artificiellement lorsque nous regardons des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait converger les axes optiques sur lui tandis que les yeux s'accommodent sur l'écran (Figure 10.40). Cette modification de la relation convergence-accommodation est susceptible de créer des difficultés visuelles et de la fatigue pour l'observateur, sauf quand l'objet se situe au niveau de l'écran. En conclusion, il faut se contraindre à positionner l'objet principal à proximité du plan de l'écran.

Des considérations cognitives impliquent un autre problème. Un écran d'ordinateur ou de télévision est perçu par le cerveau comme une fenêtre limitée par les bords de l'écran. Ceux-ci peuvent être source de confusion. Des objets affichés en parallaxe négative, donc situés devant l'écran par la stéréopsie, et partiellement occultés par les bords de l'écran (mentalement derrière la fenêtre) créent un conflit dans leur localisation. Ce phénomène peut empêcher la fusion des images et perturbe l'observateur. Il faut en déduire qu'il est rarement souhaitable de créer des objets «sortant de l'écran» malgré le côté attrayant que cela procure. Ce procédé n'est à employer que pour les

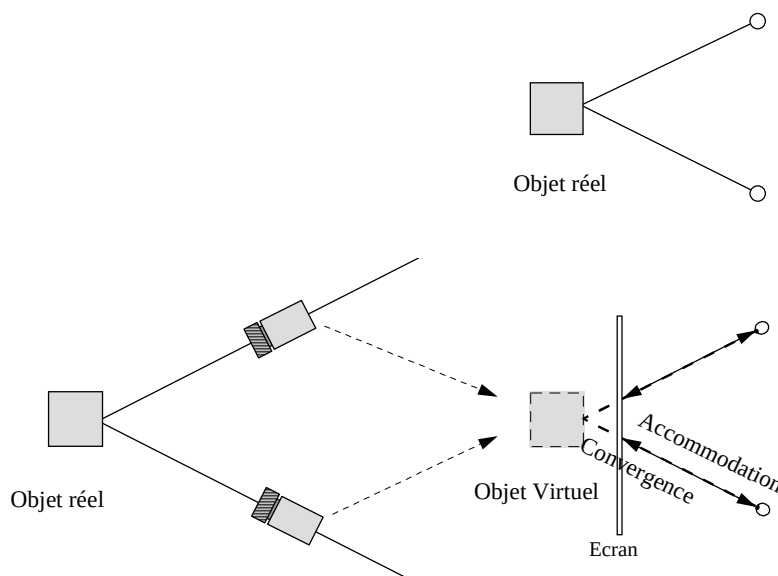


Figure 10.40 : Différence entre vision naturelle et restitution stéréoscopique

très larges écrans des démonstrations spectaculaires des films en relief. Vous pourrez remarquer que les publicités des systèmes stéréoscopiques vantent souvent par l'image ce phénomène qui est impossible à reproduire sur ces dispositifs !

Si l'observateur se translate parallèlement à l'écran, il voit les objets se déplacer anormalement : il perçoit des mouvements appelés pseudoscopiques (Figure 10.41).

Les yeux de l'observateur se translatant vers la droite, il voit le point P se déplacer vers la gauche au point P', car les points Id et Ig sont fixes. Ce phénomène est rarement perceptible, car l'observateur est souvent assis, ce qui implique une translation latérale de la tête trop faible pour entrevoir ces faux mouvements. Un schéma similaire permet de comprendre que si l'observateur recule ou avance face à l'écran, il y a étirement ou raccourcissement de la profondeur perçue (Figure 10.42).

Ce phénomène implique qu'il n'y a théoriquement qu'une seule distance à l'écran (la distance orthostéréoscopique) susceptible de ne pas déformer les objets selon l'axe de la profondeur. Le cerveau tolère en pratique une marge de déformation entre la dimension en profondeur et les dimensions parallèles à l'écran. Si l'observateur est susceptible de bouger sa tête et si on désire parfaire la vision, il est souhaitable de déterminer la position de la tête grâce à un capteur de localisation spatiale : des fournisseurs proposent ce dispositif, principalement sur les bureaux immersifs et dans les visiocubes. En conclusion, l'affichage est théoriquement prévu pour un seul observateur à une position bien déterminée. En pratique, ces mouvements pseudoscopiques sont gênants dans des cas particuliers. Par exemple en téléopération, l'opérateur peut avoir besoin de se déplacer en commandant son manipulateur et il risque d'être perturbé visuellement par ces faux mouvements.

Sur quels critères doit-on se baser pour positionner les deux caméras fictives ou réelles ?

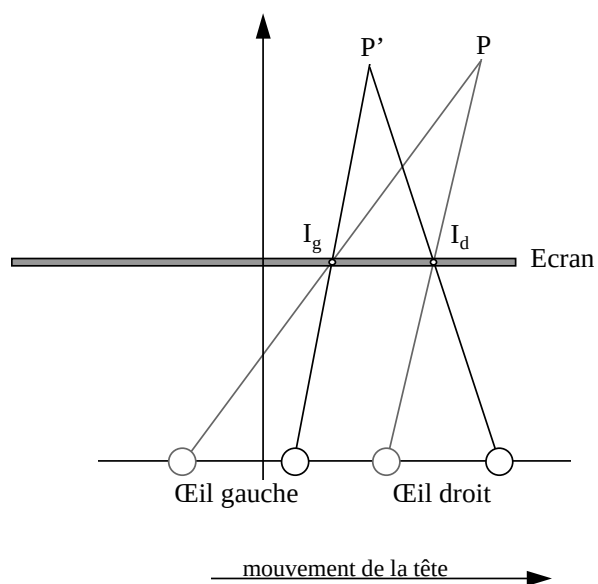


Figure 10.41 : Mouvements pseudoscopiques quand l'observateur bouge la tête non traquée

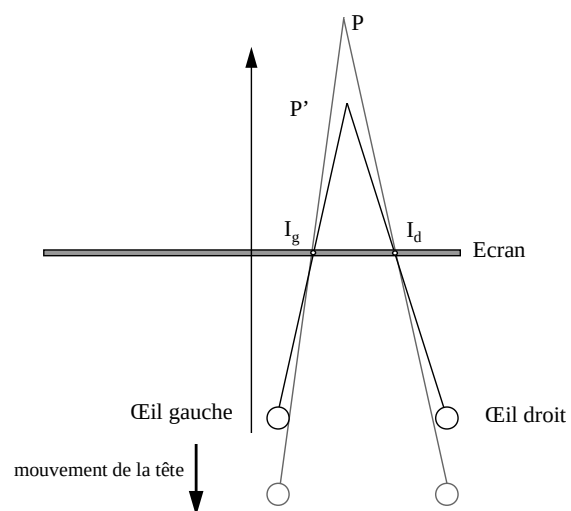


Figure 10.42 : Mouvements pseudoscopiques quand l'observateur bouge la tête non traquée

A priori, on peut penser que l'on doit faire converger les deux axes optiques vers la zone principale à observer. Dans ce cas, les projections sur les deux plans donnent des images avec des parallaxes horizontales et verticales (ces images correspondent, à un facteur multiplicatif près, aux deux images affichées sur l'écran (Figure 10.43).

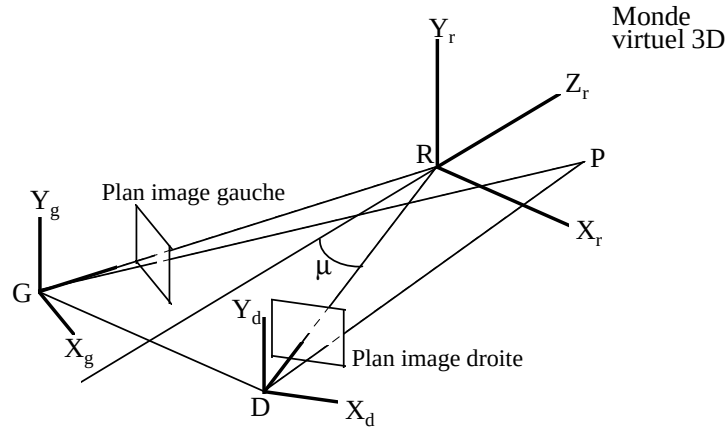


Figure 10.43 : Positionnement des deux caméras fictives ou réelles dont les axes optiques convergent

Mais ce cas n'est guère souhaitable car il impose des parallaxes verticales qui rendent difficile la fusion des images. Par exemple, les projections droite et gauche de la face avant d'un cube centré sur l'axe Z_r donnent les images suivantes sur la Figure 10.44.

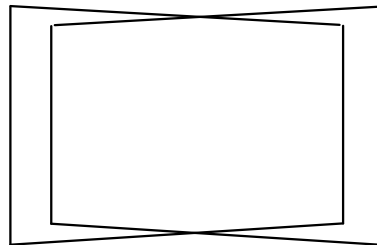


Figure 10.44 : Les parallaxes verticales sont exagérées sur la figure.

Des études ont montré qu'il est difficile de fusionner pour des parallaxes verticales d'angle supérieur à 20' d'arc [Jules, 1971]. Il est donc souhaitable, mais pas indispensable, que les deux axes optiques soient parallèles pour éviter la distorsion due à la rotation des images et les parallaxes verticales, qui procurent des gênes à l'affichage [McAllister, 1990], [Baker, 1987]. Les algorithmes permettant la création d'images stéréoscopiques de synthèse dans ce cas sont relativement simples et rapides (calculs de translations et de projections).

Une autre règle sur les parallaxes horizontales des images est à considérer pour permettre une vision aisée. Deux objets voisins au niveau de l'affichage, observés en même temps, ne doivent pas avoir une trop grande différence de parallaxes horizontales, donc

de profondeur, pour être facilement fusionnables. Cette contrainte peut parfois être plus exigeante que la condition de base sur les parallaxes horizontales.

On peut penser que l'écartement des axes optiques des caméras doit être égal à la DIO de l'observateur. Si cela est possible et souhaitable pour des scènes de dimensions «normales», ceci ne peut être envisagé dans les cas extrêmes :

- pour des vues aériennes stéréoscopiques, l'écartement est de plusieurs centaines de mètres pour rendre un effet de profondeur, qui n'est pas naturel mais utile ;
- pour une vision d'objets minuscules (insectes par exemple), l'écartement est inférieur à la DIO ;
- le choix de cette grandeur est important pour la vision stéréoscopique. Retenons les deux règles primordiales suivantes :
- plus l'écartement est grand, plus l'effet de profondeur est important ;
- plus la parallaxe, fonction de l'écartement, est grande, plus la fusion des images par l'observateur est difficile à réaliser.

10.6.2 CHOIX DES PARAMÈTRES STÉRÉOSCOPIQUES

Les conditions théoriques de création et de restitution d'images stéréoscopiques permettent de choisir les paramètres stéréoscopiques. Pour la création des images, deux classes de solutions sont étudiées : les axes optiques des caméras sont parallèles ou sont convergents. Dans les deux cas, on doit respecter les contraintes psychophysiques résumées ci-dessous :

A La relation accommodation-convergence doit être peu modifiée pour éviter la fatigue visuelle : l'objet principal observé doit avoir une parallaxe nulle (ou presque nulle).

B La fusion des deux images doit être possible et facile, impliquant peu de fatigue oculaire :

B1 La parallaxe horizontale doit être limitée, mais assez grande pour percevoir la profondeur. Les limitations des parallaxes horizontales sont en général les mêmes en valeur positive et négative ($1,2^\circ$ à $1,5^\circ$). Mais dans le cas d'un petit écran, le jaillissement des objets devant celui-ci est déconseillé.

B2 La parallaxe verticale doit être nulle ou très limitée ($20'$ d'arc).

B3 Deux plans voisins au niveau de l'affichage, observés en même temps, ne doivent pas avoir une trop grande différence de profondeur.

C Pour mieux percevoir en relief, il faut que la vision restituée soit isomorphe si possible à la vision du monde réel. On parle de vision orthostéréoscopique. Celle-ci doit respecter les lois de perspective et de perception de la profondeur par stéréopsie. Ceci implique trois conditions :

C1 L'écartement entre les deux caméras est égal à la distance interoculaire.

C2 Les images stéréoscopiques ne doivent pas être étirées ou aplaties en profondeur, donc les angles des champs de vision à l'acquisition et à la restitution doivent être identiques ou peu différents. Ceci implique pour un écran vu à la distance d_o , que celle-ci doit être égale au produit de la focale f par l'agrandissement M de l'objet : $d_o = f.M$

C3 La distance de l'objet principal observé par rapport aux caméras doit être égale à la distance de son image par rapport à l'observateur.

Ces conditions sont employées dans un cas précis de téléprésence visuelle au paragraphe suivant et sont exprimées analytiquement en annexe à la fin du chapitre.

10.6.3 CRÉATION D'IMAGES EN RELIEF POUR LA TÉLÉOPERATION

10.6.3.1 Téléprésence visuelle stéréoscopique

Nous présentons une application de téléopération en vision stéréoscopique développée à la Direction des Etudes et Recherches d'EDF, en partenariat avec notre centre de recherche de Robotique à l'Ecole des Mines de Paris. L'objectif global est d'étudier les apports des techniques de la Réalité Virtuelle pour la Téléopération.

Les travaux de téléopération dans les centrales nucléaires peuvent être améliorés par une téléprésence visuelle stéréoscopique. Les tâches de téléopération consistent à mettre en place des outillages, à pratiquer des contrôles, à réaliser des rechargements de soudure, etc., dans un environnement distant et souvent hostile. Il faut rendre les opérateurs téléprésents dans l'environnement de travail en fournissant des informations pertinentes et réalistes. Ces informations proviennent d'images issues de caméras vidéo et d'images de synthèse obtenues à partir de modélisations. Aujourd'hui nous sommes dans une phase où les deux sources sont exploitées le plus souvent séparément, leur fusion n'étant que du domaine expérimental. Le cas pratique étudié consiste en une série de tâches de téléopération dans une boîte à eau d'un générateur de vapeur (espace confiné d'un quart de sphère de 3,5 m de rayon). L'objectif recherché est une meilleure **immersion visuelle** de l'opérateur dans l'espace de travail. Dans tous les cas, le téléopérateur doit toujours être capable de se situer aisément par rapport à la tâche qu'il accomplit.

Nous exploitons deux techniques pour la vision en relief. Leur principe repose sur une séparation temporelle des deux images pour les deux yeux, l'écran stéréoscopique étant à double fréquence (100 Hz) afin de conserver 25 images par seconde pour chaque œil. La séparation se fait soit par polarisation circulaire de la lumière et filtrage de celle-ci par des lunettes passives, soit par obturation électronique des images au niveau de lunettes actives à cristaux liquides. Un capteur de localisation électromagnétique permet de déterminer les 6 degrés de liberté de la tête de l'opérateur à la fréquence de 120 mesures par seconde. Ce capteur permet d'asservir les mouvements des caméras aux mouvements de la tête (Figure 10.45)

10.6.3.2 Etude de la vision en relief

La fusion des deux images affichées sur un écran ne peut se faire sans fatigue oculaire que sous certaines contraintes, principalement :

1. La relation accommodation-convergence doit être peu modifiée pour éviter la fatigue visuelle.
2. La parallaxe horizontale doit être limitée, mais assez grande pour percevoir la profondeur.
3. La parallaxe verticale doit être nulle ou très limitée.

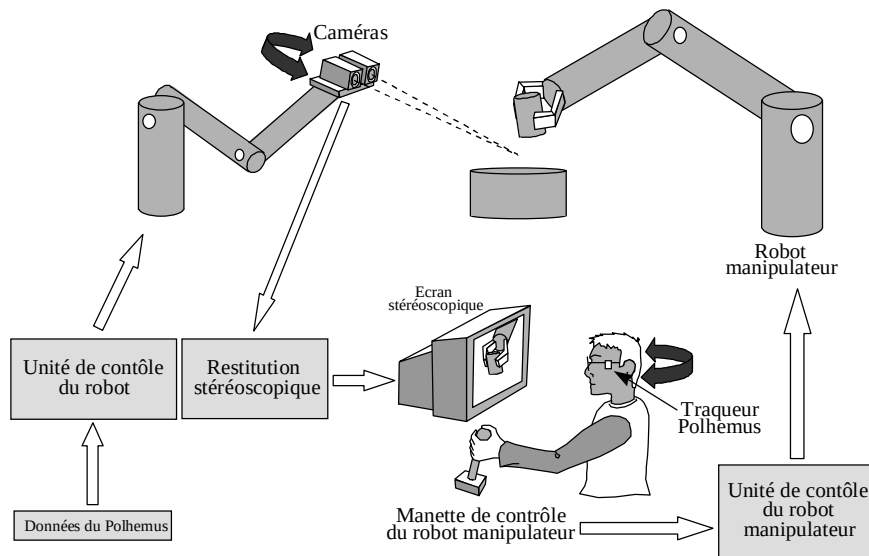


Figure 10.45 : Dispositif de téléopération et de téléprésence stéréoscopique

4. Deux plans voisins au niveau de l’affichage, observés en même temps, ne doivent pas avoir une trop grande différence de profondeur.
5. Les images stéréoscopiques doivent être peu étirées ou aplaties en profondeur.

Pour déterminer théoriquement les solutions envisageables pour l’acquisition d’images en relief avec variation de focales f (zoom de 16 à 48 mm) dans un espace limité de 0,5 m à 2,5 m, nous prenons en compte les contraintes psychophysiques précédentes. Les hypothèses pour cette étude sont les suivantes. L’observation se fait par un opérateur qui est placé à la distance optimum $d_o = 80$ cm de l’écran. Il doit pouvoir travailler longtemps en vision stéréoscopique sans fatigue exagérée induite par ce type de vision. Comme nous l’avons écrit, les prises de vue peuvent se faire par des caméras à axes optiques parallèles avec décalage logiciel entre les deux images ou à axes optiques convergent vers l’objet principal.

Formulation des contraintes dans le cas des axes optiques parallèles
(Voir en annexe à la fin du chapitre pour l’établissement des formules)

Contrainte C2 : pour ne pas avoir d’aplatissement ou étirement en profondeur de l’image, il faut que la déformation en profondeur.

$$\rho = d_0 / f M \quad (10.1)$$

où M est l’agrandissement des images de l’écran par rapport à celles des CCD ($M = 36$ dans notre cas). Donc pour une distance précise de l’observateur à l’écran, il n’y a

qu'une valeur de focale des caméras où la vision est sans déformation (vision orthostéréoscopique). Mais on estime que ρ peut varier entre 0,7 et 1,4 sans trop perturber et fatiguer la vision. Dans ces conditions, on choisit $d_o = 0,8 \text{ m}$ (pour $f = 16 \text{ mm}$: $\rho = 1,4$ - pour $f = 22 \text{ mm}$: $\rho = 1$ - pour $f = 32 \text{ mm}$: $\rho = 0,7$). Mais pour les forts grossissements des zooms ($f > 32 \text{ mm}$), les images sont perçues aplaties en profondeur. La distance $d_o = 0,8 \text{ m}$ implique un champ de vision correct face à l'écran sans percevoir ses lignes.

Contrainte A : la relation naturelle entre accommodation et convergence des yeux est peu perturbée si l'objet principal observé a une parallaxe nulle $z_{po} = 0$. Le décalage D_h des images doit être réglé en conséquence :

$$z_{po} = \frac{f.M.e_c}{2.D_h} \quad (10.2)$$

Contrainte B1 : la parallaxe horizontale doit être limitée expérimentalement : $P_h < 1,5^\circ = 0,026 \text{ rd}$

$$P_h = \frac{f.M.e_c \cdot \left(\frac{1}{z_{po}} - \frac{1}{z} \right)}{d_o} < 0,026 \quad (10.3)$$

Il faut prendre la distance e_c (écart entre les 2 caméras) la plus petite possible mécaniquement, dans notre cas égale à 65 mm, pour que la variation possible de z autour de z_{po} soit la plus grande.

Si $z > z_{po}$ il faut $z < 1 / (1/z_{po} - 0,026 \cdot d_o / f.M.e_c)$

Si $z < z_{po}$ il faut $z > 1 / (1/z_{po} + 0,026 \cdot d_o / f.M.e_c)$ (à éviter sur petit écran)

Contrainte B2 : la parallaxe verticale est nulle dans le cas des caméras à axes parallèles.

Contrainte B3 : deux plans voisins au niveau de l'affichage, observés simultanément, ne doivent pas avoir une trop grande différence de profondeur. Le seuil S à ne pas dépasser est défini par les proximités, inverses des distances z_1 et z_2 des deux plans :

$$\frac{1}{z_1} - \frac{1}{z_2} < S \quad (10.4)$$

avec z exprimé en mètre, S a pour limite 1 m^{-1} et il est préférable que S soit égal à $0,5 \text{ m}^{-1}$.

10.6.3.3 Dédutions des contraintes

Les contraintes qui limitent la zone de vision en profondeur sont les N° 2 et 4. On trace les deux premières inéquations en fonction de différentes valeurs de la focale f (16, 22, 32 et 48 mm) et pour la troisième inéquation on prend z_1 ou $z_2 = z_{po}$ (Figure 10.46).

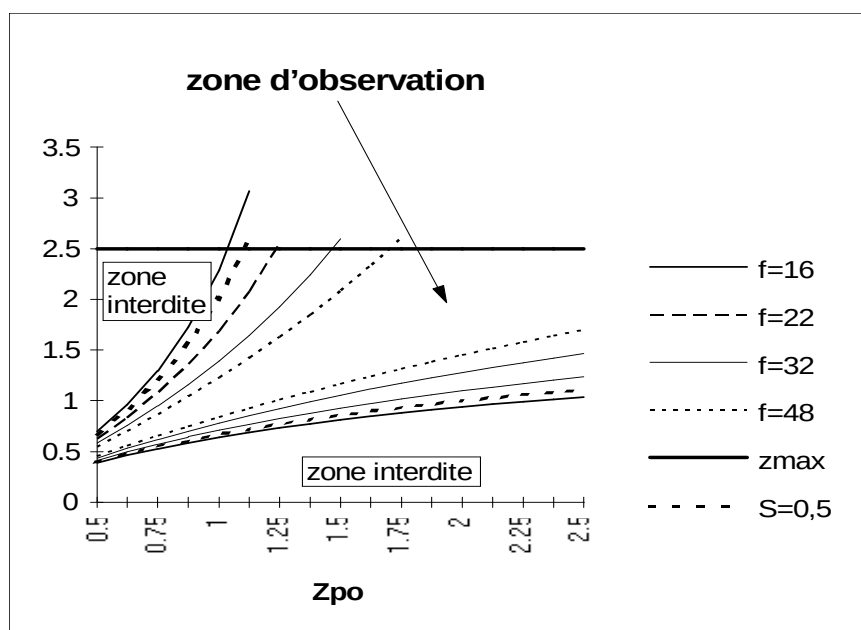


Figure 10.46: Zone d'observation possible sans fatigue oculaire en fonction des focales

10.6.3.4 Limitation de la vision en relief

Pour $f \geq 16$ mm la contrainte B3 est toujours vérifiée quand la contrainte B1 l'est aussi. Ce qui revient à dire que lorsque les parallaxes horizontales sont limitées, les différences de profondeur entre les objets sont aussi suffisamment limitées pour permettre une vision sans fatigue. Pour des focales < 16 mm, qu'il vaut mieux ne pas utiliser car la déformation en profondeur est trop importante, $\rho > 1,4$, c'est la contrainte B3 qui limite la zone d'observation. La vision est sans problème pour des focales entre 16 et 32 mm et pour une zone d'observation supérieure à 1,5 m. (z_{po} peut prendre dans ce cas toutes les valeurs $> 1,5$ m). Par contre si la zone d'observation est comprise entre 1 et 1,5 m, celle-ci est limitée (fonction de f) autour de la distance z_{po} qui doit être la distance de l'objet principal observé. La vision avec grande focale $f > 48$ mm se fait avec aplatissement en profondeur. La vision rapprochée (inférieure à 1 m) est très difficile. Elle ne peut se faire qu'avec faible focale et que si les objets observés sont à profondeur voisine.

Ces calculs théoriques sur les zones de vision acceptables ont été vérifiés expérimentalement. Les résultats des tests de vision en relief dans les mêmes conditions confirment que des zones sont inobservables pour certaines valeurs de focales des zooms. En pratique, cela nous incite à rendre impossible à l'opérateur les choix de f et de z_{po} ne respectant pas les contraintes. Pour cela, nous interdisons par logiciel les conditions inadéquates de vision stéréoscopique. L'étude dans le cas des caméras à axes convergents donne des conditions et des courbes similaires, que le lecteur peut trouver dans notre communication [Fuchs, 1995].

10.6.4 LIMITATION DE LA FATIGUE OCULAIRE EN VISION STÉRÉOSCOPIQUE

10.6.4.1 Problématique de la fatigue oculaire

Nous avons signalé dans le chapitre sur les sens et les réponses motrices que la fusion aisée de deux images stéréoscopiques dépend de leurs parallaxes en tous points homologues et aussi de leurs fréquences spatiales en ces points homologues. Nous souhaitons limiter la fatigue visuelle, due aux écrans stéréoscopiques, par un contrôle de la disparité entre les images gauche et droite *mais aussi par un contrôle du contenu en fréquences spatiales des images*. Cette méthode est donc plus efficace que celle décrit précédemment car maintenant nous prenons en compte les fréquences spatiales locales en plus de la disparité. La diminution de la fatigue oculaire est basée sur les fréquences spatiales et les disparités, en exploitant une fonction de confort visuel dans l'image. Cette fonction permet de déterminer si un point de l'image fatigue l'observateur et si le point image doit faire l'objet d'un traitement. La fonction de confort a été déterminée expérimentalement par plusieurs chercheurs. Nous utilisons la fonction de confort stéréoscopique obtenue à partir des données de Wöpking, [Wöpking, 1995]. Ce dernier a réalisé des tests de confort (relativement à la fatigue oculaire) en fonction de la disparité et du contenu fréquentiel en chaque point image : quand les fréquences spatiales sont élevées, la fusion des images stéréoscopiques est plus difficile et plus fatigante (par exemple, pour les objets représentés seulement par leurs contours, donc à fréquence élevée). Quand la disparité est de $1,5^\circ$ ($90''$), la fatigue oculaire est incommode pour une fréquence spatiale élevée de 23 cpd, légèrement incommode à 5,6 cpd et imperceptible à 1 cpd (Figure 10.47) :

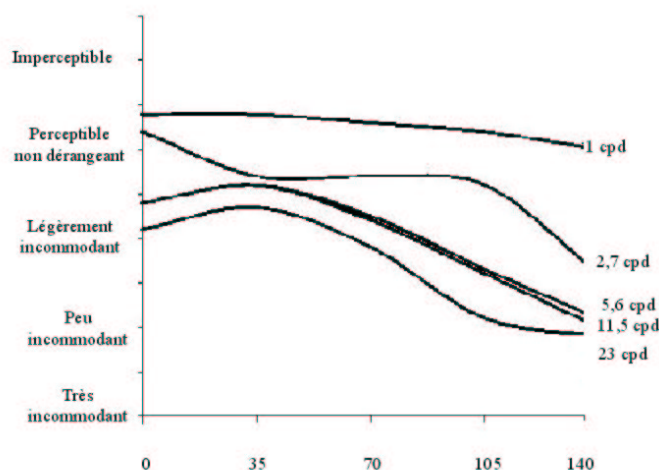


Figure 10.47: Fonction de confort oculaire en vision en relief par rapport à la disparité donnée en minute d'angle, Wöpking, 1995

10.6.4.2 Méthode par filtrages fréquentiels

Deux solutions peuvent être envisagées pour limiter la fatigue visuelle :

- soit masquer les points où les fréquences spatiales et les disparités ne sont pas compatibles (par rapport aux résultats de Wöpking). Solution : le rideau virtuel et le calque virtuel ;
- soit tronquer le contenu en fréquences spatiales dans les hautes fréquences pour faciliter la fusion des images (pour se retrouver dans une zone de confort sur la courbe de Wöpking). Solution : le flou adaptatif.

Des résultats de ces filtrages sont présentés à partir du couple d'images stéréoscopiques de la Figure 10.48.

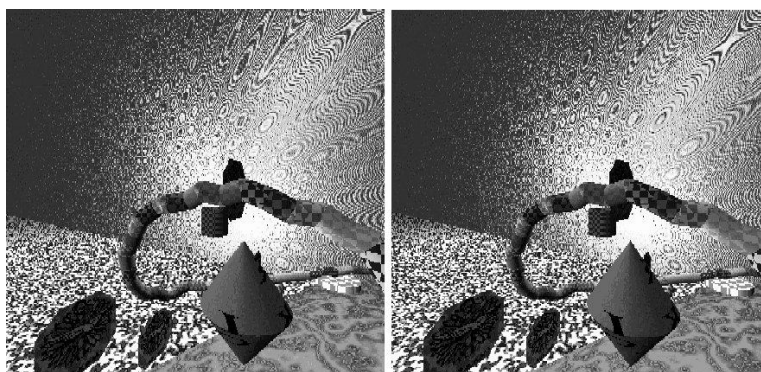


Figure 10.48 : Images de base pour les différents filtrages (ces images peuvent être vues en relief comme la figure 10.36)

La technique du calque virtuel (Figure 10.49) consiste à étendre une zone de flou sur les points des images ne satisfaisant pas la condition de confort stéréoscopique : les régions pour lesquelles la fusion est difficile semblent alors être vues à travers une vitre dépolie.

Avec le masquage du flou adaptatif (Figure 10.50), les traitements des images se font en premier par la décomposition fréquentielle de l'image par ondelettes. Au lieu de tester la compatibilité de la disparité et de la fréquence locale, nous choisissons d'adapter la fréquence locale à la disparité. C'est une technique de flou adaptatif car, en chaque point de l'image, son contenu fréquentiel est rendu adapté à la disparité locale. Les hautes fréquences, « difficiles à fusionner », sont supprimées : l'image est décomposée par traitement en ondelettes et elle est ensuite recomposée à partir des coefficients des ondelettes, sauf ceux correspondant aux hautes fréquences.

10.6.4.3 Résultats expérimentaux

Nous avons testé ces filtrages sur une tâche d'interaction dans un monde virtuel. Les testeurs déplaçaient un anneau le long d'un fil sans toucher ce dernier, l'anneau en-

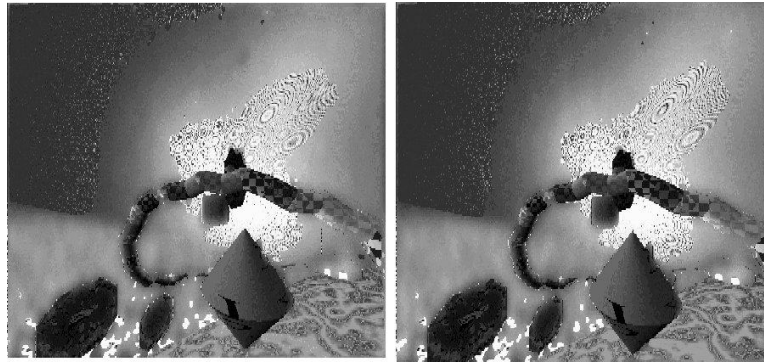


Figure 10.49 : Le calque virtuel : les zones dont la disparité et la fréquence locale sont incompatibles du point de vue de la fonction de confort sont flouées

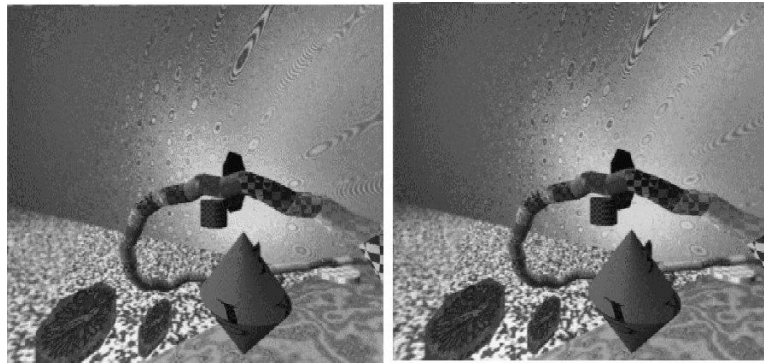


Figure 10.50 : Le flou adaptatif : là où la disparité et la fréquence locale sont incompatibles, les hautes fréquences ont été supprimées

tourant ce fil. Les testeurs manipulaient l'anneau virtuel avec un traqueur à six degrés de liberté et observaient la scène grâce à un écran stéréoscopique. Suivant les tests, les images étaient filtrées ou non. Chaque expérience durait un certain temps pour que la fatigue oculaire puisse être ressentie. Le résultat principal est que le filtrage par le flou adaptatif est nettement préféré aux autres filtrages et à l'absence de filtrage, en terme de diminution de fatigue. Par contre, il peut surprendre, surtout pour les novices des images en relief, car il dégrade légèrement la netteté des images. Tous les résultats qualitatifs et quantitatifs sont présentés dans la thèse de Jérôme Perrin [Perrin, 1998b].

10.6.4.4 Conclusion

Mais ces traitements sont à l'heure actuelle difficile à réaliser en temps réel. Pour les environnements virtuels, la carte de disparité est facilement déterminée à partir de la carte de profondeur (Z buffer), mais les fréquences spatiales locales sont plus longues à calculer par décomposition en ondelettes. Quand les calculs seront réalisables en temps réel, ces filtrages seront souhaitables, voire indispensables. Pour toute application de réalité virtuelle exploitant les images en relief dans un cadre professionnel, le filtrage par flou adaptatif permet une perception de la profondeur sans fatigue oculaire exagérée.

10.6.5 CRÉATION D'IMAGES EN RELIEF POUR UNE REVUE DE PROJET

Dans notre projet IMAVE pour l'entreprise Sommer Allibert (Immersion Adaptée à un Véhicule), que nous avons déjà présenté, l'objectif principal est la visualisation d'une planche de bord à partir de la place du conducteur. L'application doit permettre de faire de la revue de projet entre personnels de différents corps de métier. Après étude des I² sensori-motrices, il fut convenu que la perception de la profondeur devait être la meilleure possible (surtout pour la perception des courbures et des formes). La vision de la planche de bord de la place du conducteur est-elle possible en vision stéréoscopique, sans fatigue exagérée ? La planche de bord devant être aussi observée quand le conducteur se penche, il fut décidé que la tête du conducteur serait traquée.

Sans faire tous les calculs dans ce paragraphe, une solution est possible si on place l'écran au niveau de la planche de bord : à la distance optimum par rapport aux yeux de telle façon que les deux limites avant et arrière de la planche de bord impliquent les mêmes parallaxes horizontales, en valeur absolue. Avec les dimensions réelles, nous obtenons une solution, mais qui est limitée car les parallaxes horizontales sont voisines de 1,5° (Figure 10.51). Le problème est plus délicat si l'observateur veut regarder la planche de bord de plus près : la parallaxe horizontale négative dépasse les 1, 5° en valeur absolue. Il a fallu trouver un compromis : quand le conducteur s'approche de la planche de bord, donc de l'écran, nous diminuons l'effet stéréoscopique de profondeur (en pratique, les deux points de vue des caméras virtuelles se rapprochent et ne sont donc plus aux positions nominales des yeux). Les résultats sur l'observation de la planche de bord ont été satisfaisants pour une centaine de personnes, qui ont pu voir la planche de bord en relief, *sans de fatigue oculaire*.

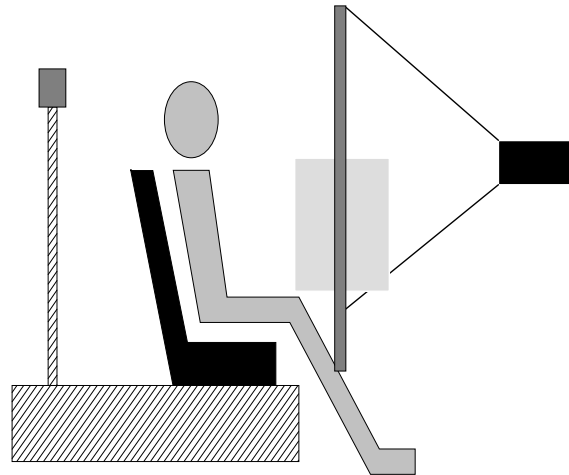


Figure 10.51 : Schéma montrant la position virtuelle de la planche de bord, vue en relief

10.7 ÉVALUATION DES TECHNIQUES STÉRÉOSCOPIQUES

10.7.1 INTÉRÊTS DE LA VISION EN RELIEF

Des recherches ont été entreprises pour quantifier l'apport de la vision stéréoscopique pour un utilisateur. Les études ont surtout porté sur les systèmes de vision à un écran, car les visiocasques sont encore trop récents. Les principales recherches portent sur l'amélioration de la perception de la profondeur due à la vision stéréoscopique, sur la fatigue visuelle supplémentaire et sur les problèmes inhérents à la vision en relief (crosstalk, fusion, etc.). La recherche de Tetsuya Mityashita [Tetsuya, 1990] porte sur les causes de la fatigue et sur les améliorations possibles pour la diminuer. La fatigue visuelle est plus faible pour une grande profondeur de champ. Elle n'est que légèrement supérieure pour une vision avec disparité par rapport à une vision sans disparité (l'expérience a été réalisée avec deux écrans cathodiques miniatures).

10.7.2 CHOIX DES PARAMÈTRES DE LA VISION EN RELIEF

Comme nous l'avons écrit précédemment, la parallaxe doit être petite pour ne pas créer de grandes difficultés pour la vision stéréoscopique. Dans le cas contraire, le cerveau est incapable de fusionner les deux images ou il le fait avec une certaine fatigue. L'objectif est donc de créer des images stéréoscopiques avec faible parallaxe tout en gardant un effet de profondeur, et de peu perturber la relation entre convergence et accommodation. Il faut, si possible, positionner l'objet principal observé proche du plan de l'écran.

Suite à des études expérimentales [Valyus, 1962], on a démontré la difficulté de fusionner deux images planes ayant un angle de parallaxe supérieur à 1,5 degrés : le paramètre limitatif est l'angle de parallaxe (angle de vision pour deux points homologues) et non la distance parallaxe. Nos études entreprises dans le cadre de la téléopération avec des

images réelles indiquent une limite d'angle de parallaxe de $1,2^\circ$ [Fuchs, 1995].

L'étude de Yei-Yu Yeh [Yeh, 1990] examine la limite de la fusion et l'appréciation de la profondeur. La limite de fusion est analysée en fonction de différents paramètres des stimuli affichés sur un écran cathodique avec lunettes actives : la couleur, le temps d'affichage, la position verticale sur l'écran et la parallaxe positive ou négative. Les résultats impliquent que la limite de fusion (donnée en angle de parallaxe) est très faible pour un stimulus bref de 0,2 seconde : soit 27 minutes d'arc (parallaxe négative), soit 24 minutes d'arc (parallaxe positive) comparativement à un stimulus de 2 secondes : soit 5 degrés (parallaxe < 0), soit 1,5 degrés (parallaxe > 0). Il est donc plus difficile d'observer en relief des images en mouvement rapide. La couleur rouge est plus facile à fusionner que la couleur blanche (pour des stimuli de 2 secondes : 6,2 degrés pour le rouge et 3,7 degrés pour le blanc en parallaxe < 0). Les erreurs d'appréciation de la profondeur sont déterminées expérimentalement : l'erreur (d'angle de parallaxe) est de 2,2 minutes pour des images ayant des angles de parallaxe entre 0 et 30 minutes.

Les études entreprises au Cerma (Centre d'Etudes et de Recherche de Médecine Aérospatiale) du Ministère de la Défense donnent des résultats corroborant les précédents, en étant plus précis. La limite de fusion est aussi fonction de la fréquence spatiale des images, ce qui n'est pas étudié dans les travaux cités précédemment. Leurs études permettent de conclure que de plus grandes disparités horizontales sont fusionnables lorsque les stimulations visuelles sont constituées de basses fréquences spatiales. Ils ont mis en évidence deux mécanismes impliqués dans la fusion selon le temps du stimulus : une fusion immédiate si la disparité n'est pas trop grande et une fusion non immédiate pour des disparités plus grandes, mettant en jeu des mouvements de vergence réflexes des yeux. Les résultats donnent entre autres : à la fréquence spatiale de 4,8 cycles/degré, fusion immédiate jusqu'à 20' d'arc et fusion maximale à 52' d'arc environ. A la fréquence spatiale de 0,22 cycles/degré, fusion immédiate jusqu'à 80' d'arc et fusion maximale à 176' d'arc environ [Perrin, 1998].

Les images doivent être affichées pour chaque œil avec une résolution et une fréquence suffisantes, que ce soit sur un ou sur deux écrans : il est souhaitable d'avoir au moins 25 images par seconde pour chaque œil, donc une fréquence de trame de 50 Hz (60 Hz norme américaine). Suivant les techniques employées, il faut dans certains cas avoir des écrans à 120 Hz pour les écrans stéréoscopiques et à 180 Hz en monochrome pour les visiocasques à tube cathodique. Il faut bien noter que certains constructeurs proposent des écrans à de telles fréquences, mais au détriment de la résolution, qui est diminuée par rapport à des fréquences plus basses. Plus la résolution est élevée, plus les images sont de bonne qualité. Le seuil de satisfaction dépend des opérations que l'on désire effectuer avec l'affichage. La résolution doit être meilleure si on souhaite lire ou écrire un texte sur l'écran (le cas des ordinateurs) ou afficher des images esthétiques (le cas des films).

10.7.3 COMPARAISON DES SOLUTIONS À UN ÉCRAN

Le principal avantage des écrans autostéréoscopiques réside dans l'absence de lunettes pour l'observateur (grand intérêt en aéronautique et dans les lieux publics). Comme première indication, la qualité des images des deux procédés par séparation avec lunettes est comparable, les temps de fermeture et d'ouverture étant proches. La différence se situe plus au niveau de la commodité d'utilisation des lunettes polarisantes par rapport à celles qui possèdent un obturateur électronique, ainsi qu'à la limite plus

faible des dimensions des écrans avec lunettes passives.

10.7.4 COMPARAISON ENTRE LES ÉCRANS DES VISIOCASQUES

Les écrans cathodiques miniatures (CRT) sont plus lourds, plus fragiles et plus chers mais ont surtout une meilleure résolution. Les écrans à cristaux liquides pourront les concurrencer si des progrès sont réalisés pour leur résolution. C'est le principal challenge technologique en réalité virtuelle qui est engagé actuellement par les constructeurs d'écrans : réaliser de très petits écrans (un pouce environ) à très haute définition et maîtriser leur production avec un faible prix de revient. Un consortium d'entreprises américaines (Kaiser Electro-Optics, Kopin) est en train de développer de nouveaux mini-écrans à technologie AMLCD ou AMEL, qui pourraient atteindre une résolution comparable ou meilleure que les écrans cathodiques : jusqu'à 2 000x2 000 pixels. La compagnie britannique GEC Marconi travaille sur des écrans à cristaux liquides de petite taille.

10.8 CONCLUSION

Les systèmes à un écran sont des produits qui existent commercialement depuis plusieurs années, même si le marché est petit comparativement à celui des écrans monoscopiques. Ce sont donc des produits stables, fiables et vendus par des entreprises renommées en vision. Les visiocasques, produits prometteurs, vont-ils être développés sur grande échelle ? Pour l'instant, c'est souvent de petites entreprises qui commercialisent les visiocasques. Les grandes entreprises internationales des jeux vidéo se sont associées à des entreprises travaillant dans la réalité virtuelle pour le développement de visiocasques et de traqueurs à très bon marché. Mais malgré certaines annonces un peu hâtives, il n'y a pas eu l'essor attendu par certains des visiocasques dans ce secteur.

Il y a 3 à 5 ans, de nouveaux visiocasques étaient annoncés presque tous les mois. Mais depuis trois ans, il n'y a pas de grandes avancées technologiques pour les visiocasques. Ceux-ci ont gagné en qualité ergonomique mais restent de faible résolution pour les visiocasques à écrans à cristaux liquides. On ne pourra franchir une étape importante pour les interfaces visuelles que lorsque les mini écrans à haute résolution seront disponibles. Ceci ne peut être envisager sérieusement que s'il y a un marché important pour ce type d'interface, ce qui n'est pas le cas actuellement.

L'utilisateur d'un système de vision stéréoscopique doit retenir que la restitution visuelle est similaire à la vision naturelle, mais non identique. Cette restitution est fonction de critères *géométriques et psychophysiques*. N'oublions pas que chaque homme réagit différemment devant la vision en relief. L'immersion visuelle semble théoriquement un atout complémentaire en réalité virtuelle, mais les problèmes techniques et ergonomiques qu'elle pose sont à prendre en compte pour un emploi fréquent et de longue durée.

10.9 ANNEXES

10.9.1 *RESTITUTION PAR IMAGES VOLUMIQUES*

Il s'agit de créer une image virtuelle totalement en trois dimensions. La société Texas Instruments propose un tel système créant une image spatiale : l'Omni View.

Un miroir de surface hélicoïdale tourne rapidement (30 tours par seconde environ) dans une enceinte semi-sphérique. Des rayons laser sont envoyés sur lui pour créer l'image correspondant à la position du miroir. En commandant les lasers, l'ordinateur reconstitue la luminosité et la couleur de chaque point à visualiser. Par persistance rétinienne, l'observateur voit une véritable image spatiale. Le volume de la demi sphère est de 24 pouces de diamètre. L'affichage est de 7000 voxels, de diamètre 1,2 mm, et la fréquence est de 30 Hz.

Ce système est conçu pour des applications particulières (militaire, contrôle aérien, médecine, etc.). Il n'a pas les mêmes objectifs d'emploi que les dispositifs classiques d'affichage, puisque son utilisation n'est intéressante que si l'observateur se déplace autour des images volumiques. Ceci impose en pratique que la personne soit debout pour profiter des fonctionnalités du système.

10.9.2 *DISPOSITIFS POUR LA POLARISATION DE LUMIÈRE*

10.9.2.1 *Polarisation de la lumière*

La lumière peut être considérée comme une onde électromagnétique caractérisée dans un milieu isotrope par deux vecteurs orthogonaux : un champ électrique et un champ magnétique situés dans un plan perpendiculaire au rayon lumineux. Les récepteurs usuels (œil, caméra, etc.) ne sont sensibles qu'au champ électrique. Celui-ci a un vecteur qui change d'orientation rapidement et aléatoirement dans les trains d'onde émis par les sources lumineuses. La lumière a dans ce cas une symétrie de révolution autour de sa direction de déplacement. Certains phénomènes permettent de filtrer la lumière naturelle pour obtenir une lumière polarisée rectiligne ou circulaire. Un des phénomènes produisant la polarisation est la double réfraction produite par certains cristaux : les mini cristaux dichroïques sont orientés et inclus dans un film plastique (polaroid).

10.9.2.2 *Obturbateur à cristaux liquides pour l'affichage stéréo à multiplexage temporel*

Les cristaux liquides sont des matériaux qui sont dans une phase particulière de la matière, en générale dans une phase nématique : les molécules sont ordonnées en orientation. Elles sont approximativement parallèles entre elles selon leur axe longitudinal (comme dans un cristal) et ne sont pas liées en translation entre voisines (comme dans un fluide). Des facteurs externes peuvent modifier leur orientation comme le contact avec une surface «brossée» d'un matériau. Un champ électrique peut aussi modifier l'orientation des molécules (Figure 10.52).

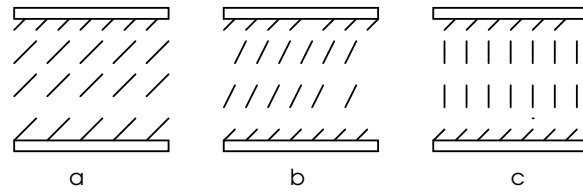


Figure 10.52: En a : champ électrique faible, en b : champ moyen, en c : champ fort

Ce type de cellule (ECB : Electrically Controllable Birefringence) peut être employée pour changer l'état de polarisation de la lumière, le passage de la lumière dépendant de l'orientation des molécules. Pour un rayon lumineux de direction quelconque et de longueur d'onde λ_v passant à travers une cellule, les deux composantes parallèle et perpendiculaire à l'orientation des molécules traversent le cristal liquide à vitesse différente. On obtient un déphasage entre les deux ondes et une direction différente en sortie pour le rayon lumineux :

Elles ont des longueurs d'onde différentes $\lambda_e = \lambda_v / n_e$ et $\lambda_o = \lambda_v / n_o$ (λ_v longueur d'onde dans le vide, n_e et n_o indices de réfraction). Elles entrent en phase et ressortent déphasées. Soit d la longueur de la cellule et soit un rayon lumineux entrant perpendiculairement à la surface de la cellule, la composante de direction extraordinaire (indice e) a un angle de phase $2\pi d/\lambda_e$ en sortie et la composante de direction ordinaire (indice o) a un angle de phase $2\pi d/\lambda_o$, (Figure 10.53), d'où le déphasage $= 2\pi d(n_e - n_o)/\lambda_v$.

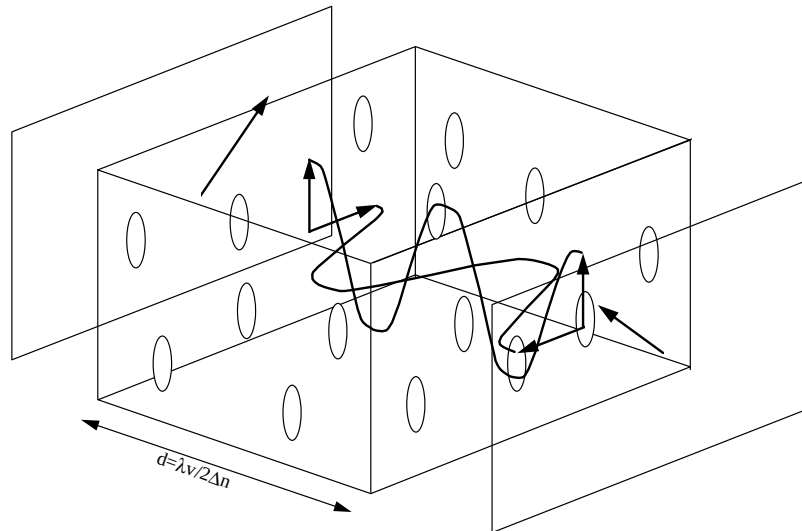


Figure 10.53: Cheminement de la lumière dans la cellule

Exemple : on obtient un retardateur d'une demi-longueur d'onde en prenant $d = \lambda_v / 2 \cdot (n_e - n_o)$ d'où le déphasage $= \pi$ (d est fonction de la longueur d'onde).

Si le retardateur d'une demi-longueur d'onde est placée entre deux polariseurs croisés à 45° , l'ensemble donne une grande transmission de la lumière. Le dispositif ECB implique : si le voltage est élevé (30V, le cas c de la figure précédente), n_e est égal à n_o et donc la transmission est faible. Réciproquement, si le voltage est faible (le cas a de la figure précédente), n_e est différent de n_o et la transmission est élevée.

Un autre type commun de cellule est le Twisted Nematic (TN). Elle est similaire à une cellule ECB, sauf que les deux surfaces striées de contact sont à 90° l'une de l'autre. On dispose parallèlement deux plaques de verres sur lesquelles on dépose faces intérieures deux films de polymère dont leur direction de «brossage» sont perpendiculaires. Ceux-ci obligent les molécules de cristaux liquides à se vriller de 90° . En envoyant une lumière polarisée, celle-ci tourne aussi de 90° . En mettant deux polariseurs croisés sur les faces externes des verres, on laisse passer la lumière. Par contre en appliquant un champ électrique, on oblige les molécules du cristal à rester parallèles au champ et la lumière ne passe plus.

Ce type de cellule (TN) peut être utilisée pour des lunettes actives, mais leur vitesse de basculement d'un état à un autre pose un problème : quand un voltage est appliqué, elle passe à faible transmission en 1 ms, mais quand le voltage est supprimé, il faut plus de 10 ms pour obtenir la haute transmission. La cellule ECB peut être aussi employée pour des lunettes actives, mais leur vitesse de basculement d'un état pose le même problème que la cellule TN. Des recherches ont été effectuées pour améliorer les vitesses de changement d'état. Un constructeur (Tektronix) a conçu un type de cellule (π -cell) à faible épaisseur ($5\mu\text{m}$) dans ce but. Il reste un problème à résoudre pour les lunettes passives, car on a vu que dans les cellules ECB l'épaisseur d dépend de la longueur d'onde. Pour résoudre ce problème, il faut que pour l'ensemble du chemin lumineux de chaque œil *le retard en phase soit nul*. D'où le dispositif présenté sur la Figure 10.54.

Durant le temps où l'image droite est affichée, la combinaison de la ECB (de déphasage nul) avec les polariseurs retardateurs croisés donne un retard en phase nul pour tout le chemin de l'œil gauche et donc empêche la lumière de passer pour cet œil quelque soit la longueur d'onde, tandis qu'elle passe pour l'œil droit.

Avec ces dispositifs et d'autres similaires [McAllister, 1993], on obtient une bonne transmission des images désirées avec peu de chevauchement entre elles. Comme l'affichage se fait par balayage de l'écran, le chevauchement peut être plus visible en haut de l'écran (images fantômes) et dépend de la couleur affichée. Il est important pour la couleur verte. Il est recommandé de tester la qualité de l'affichage stéréoscopique par des images simples.

10.9.3 MODÉLISATION D'UN VISIOCASQUE

Un modèle optique d'un visiocasque, l'EyePhone de la société VPL, est décrit dans les bibliographies [Robinett,91]. Ce visiocasque fut un des premiers visiocasques à être fabriqué et commercialisé. Malgré son ancienneté, sa modélisation optique est néanmoins intéressante à étudier (Figure 10.55).

Nous avons décrit les principales difficultés rencontrées pour la conception des optiques des visiocasques. Les optiques devant être très proches des yeux avec des grands

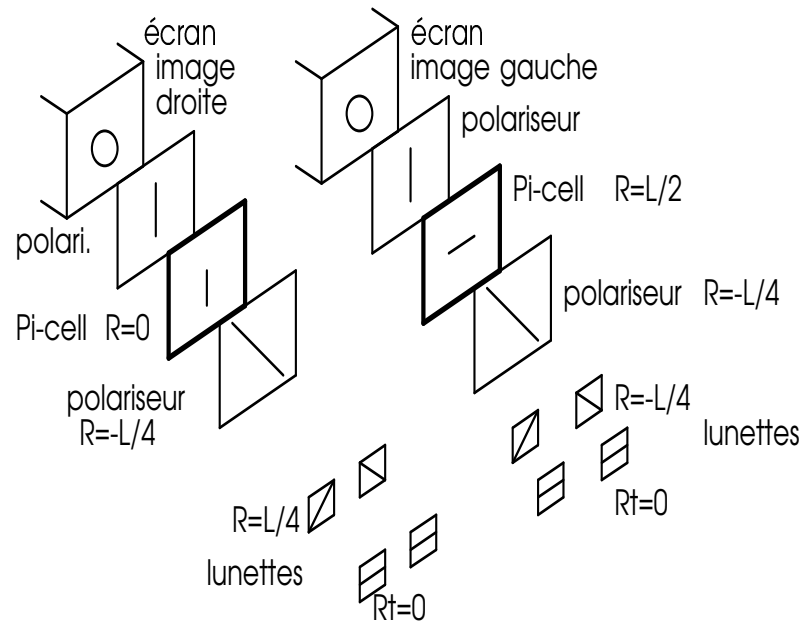


Figure 10.54: R = Retard, R_t = Retard sur tout le long du parcours et L = longueur d'onde

champs de vision, des distorsions sont inévitables. Elles peuvent être compensées en affichant les images sur les écrans avec des distorsions inverses de telle façon que les déformations soient pratiquement négligeables en sortie des optiques. Une approximation polynomiale du troisième degré pour ces déformations est suffisante et est exploitée sur certains visiocasques.

Soient L_e et L_v les largeurs de l'écran et de l'image virtuelle, on normalise les distances R_e et R_v :

$R_{en} = R_e / L_e$ $R_{vn} = R_v / L_v$ (L_v / L_e est le rapport d'agrandissement)

on prend pour approximation des distorsions :

$R_{vn} = R_{en} + K \cdot R_{en}^3 = R_{en} \cdot (1 + K \cdot R_{en}^2)$ avec K coefficient à déterminer.

comme on a :

$R_{en}^2 = X_{en}^2 + Y_{en}^2$

$R_{vn}^2 = X_{vn}^2 + Y_{vn}^2$

on obtient pour les coordonnées :

$X_{vn} = X_{en} \cdot (1 + K \cdot (X_{en}^2 + Y_{en}^2))$

$Y_{vn} = Y_{en} \cdot (1 + K \cdot (X_{en}^2 + Y_{en}^2))$

On peut prendre pour approximation inverse une fonction du troisième degré avec un coefficient K' , même si les deux approximations ne sont pas théoriquement inverses l'une de l'autre.

$X_{en} = X_{vn} \cdot (1 + K' \cdot (X_{vn}^2 + Y_{vn}^2))$

$Y_{en} = Y_{vn} \cdot (1 + K' \cdot (X_{vn}^2 + Y_{vn}^2))$

En connaissant K' , caractéristique des optiques, on peut calculer les images à dessiner sur les écrans en fonction des images virtuelles que l'on veut voir sans distorsions.

Modèle stéréoscopique :

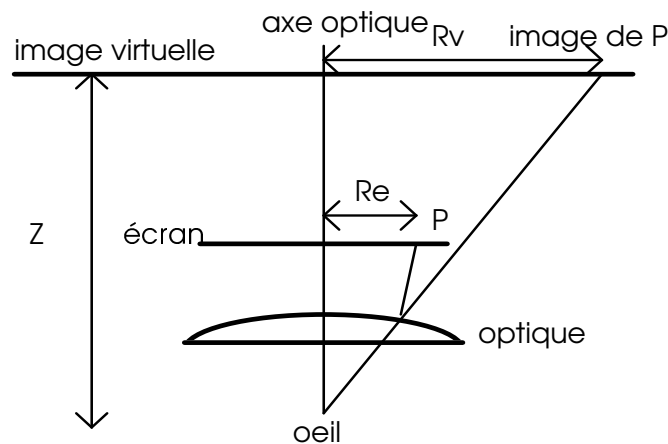


Figure 10.55 : Modèle optique pour l'affichage d'une image d'un visiocasque

Il est possible de calculer les paramètres employés dans le modèle stéréoscopique si on connaît les caractéristiques des systèmes optiques, de leur position par rapport aux deux écrans et par rapport aux yeux. Si ceux-ci changent de position (variation de la DIO ou de la distance œil-optique), les paramètres du modèle sont modifiés.

A partir de la position spatiale du point virtuel 3D (Figure 10.56), on peut déterminer les positions des deux points images virtuels P_d et P_g d'une façon similaire aux calculs effectués pour la vision stéréoscopique sur un seul écran. Et à partir de ces deux points, on peut calculer les coordonnées des points à afficher I_d et I_g à l'aide des caractéristiques des optiques. Dans la bibliographie de Robinett [Robinett,91], les calculs de la variation du champ de vision montrent que le champ est fonction de la distance de l'œil à la lentille (eye relief distance). Si les yeux se rapprochent des optiques, le champ de vision augmente. Mais cette distance doit être assez grande pour que les observateurs, portant des lunettes, puissent mettre le visiocasque (distance d'environ 30 mm pour le Eyephone). De même la profondeur de l'image virtuelle est fonction de la distance séparant les écrans des optiques. Si cette distance augmente, la profondeur de l'image, et donc la distance d'accommodation, peut croître jusqu'à l'infini (pour l'Eyephone, la distance d'écartement est de 400 mm environ).

10.9.4 PERCEPTION DU RELIEF SUR UNE FEUILLE

A partir d'une feuille, on peut percevoir le relief en vision binoculaire grâce à la stéréopsie. Ce phénomène est perceptible en contrôlant son système visuel pour l'utiliser d'une façon inhabituelle : il faut découpler la convergence et l'accommodation des yeux. Cette contrainte est la même que lors de l'observation d'écran stéréoscopique, mais dans ce dernier cas elle est modérée et seulement subie par le spectateur. Par contre, pour la perception du relief sur une feuille, l'observateur doit lui-même imposer à ses yeux cette contrainte. Il faut donc loucher devant la feuille (au point C) et accommoder au niveau de la feuille (Figure 10.57). C'est une habitude à prendre. Cet exercice demande un certain temps au début et fatigue les yeux. Il n'est pas conseillé de le faire trop longtemps. Une autre solution, sans croiser les images droite et gauche,

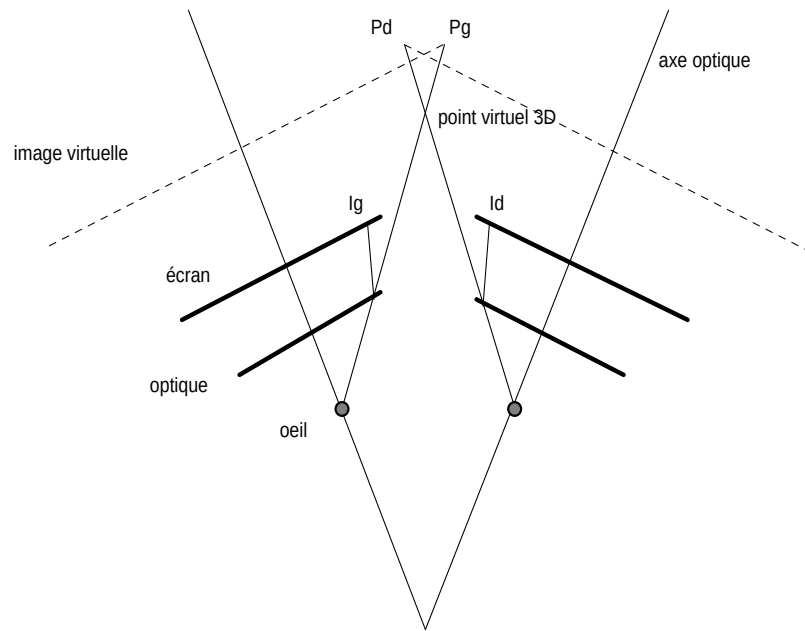


Figure 10.56 : Modèle optique d'un visiocasque

est possible, mais dans cette disposition il faut converger derrière la feuille.

10.9.5 FORMULES DE LA VISION STÉRÉOSCOPIQUE

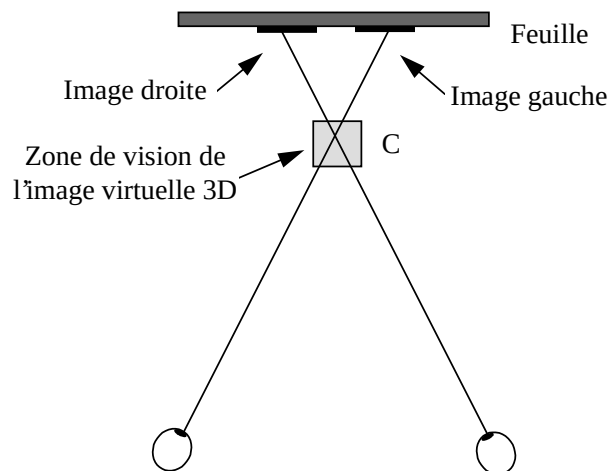
10.9.5.1 Liste des paramètres utilisés

Les paramètres fixés sont les suivants :

- e_o = distance inter-oculaire
- h_c = hauteur du CCD
- h_e = hauteur de l'écran
- l_c = largeur du capteur CCD
- l_e = largeur de l'écran
- $M = l_e/l_c = h_e/h_c$ agrandissement des images.

Les paramètres variables sont :

- f = focale
- e_c = écart entre les caméras
- d_o = distance entre l'écran et l'observateur
- d_c = distance entre les caméras et l'objet principal observé
- d_v = distance entre les caméras et le point de convergence
- P_h et P_v = parallaxes angulaires horizontales et verticales de maximum P_{h_M} et P_{v_M}



Les yeux accommodent sur la feuille et convergent en C

Figure 10.57: Schéma du principe de perception du relief sur une feuille

- $a = \frac{1}{2}$ angle de convergence des caméras : $\tan a = e_c/2d_v$
- D_h = décalage horizontal sur l'écran des deux images gauche et droite
- z_{po} = la profondeur du plan dans la scène observée ayant une parallaxe nulle
- x, y, z coordonnées d'un point de l'espace observé par rapport au repère des caméras
- X, Y, Z coordonnées d'un point de l'espace visuel par rapport à la tête de l'observateur
- x_M, y_M, z_M coordonnées maximums et x_m, y_m, z_m coordonnées minimums pour les points visibles par les caméras
- μ grandissement axial de l'image dans la direction de l'axe de prise de vue
- λ grandissement frontal de l'image dans la direction perpendiculaire de l'axe de prise de vue
- $\rho = \mu/\lambda$ déformation en profondeur (étirement ou aplatissement)
- $\gamma = 1/\rho$ grossissement angulaire

La détermination des images d'objets virtuels est identique à la création d'images d'objets réels filmés par deux caméras. On va donc considérer que deux caméras virtuelles, placées à l'endroit du point de vue de l'observateur, prennent les images du monde virtuel. Les formules sont les mêmes, que la prise de vue soit réelle ou soit virtuelle (Figure 10.58).

10.9.5.2 Les axes optiques des caméras sont parallèles

Dans la scène observée :

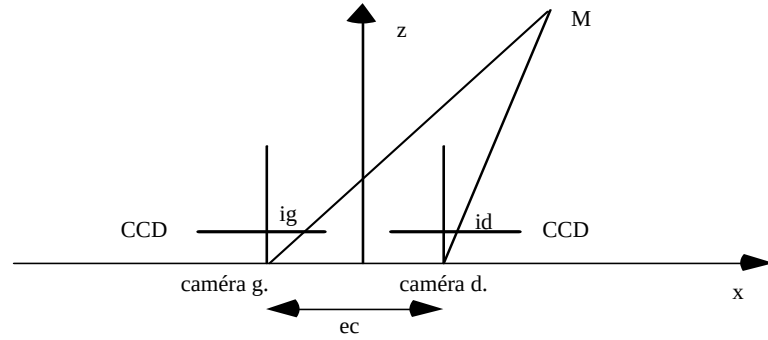


Figure 10.58 : Coordonnées dans la scène observée

Dans les repères centrés sur les capteurs CCD des caméras, on a les coordonnées des points images gauche et droite i_g et i_d :

$$i_g = \left(f \frac{(x + e_c/2)}{z}, f \frac{y}{z} \right), i_d = \left(f \frac{(x - e_c/2)}{z}, f \frac{y}{z} \right) \quad (10.5)$$

Dans l'espace visuel de l'observateur :

Ces deux points images donnent sur l'écran les deux points affichés I'_g et I'_d . Leurs coordonnées par rapport au repère centré sur l'écran sont égales aux précédentes multipliées par le facteur M . Si on ne procède pas à une translation des deux images ainsi créées, tous les objets seraient vus à parallaxes négatives, donc devant l'écran.

La parallaxe horizontale $P_h = -\frac{f M e_c}{z d_o}$ est négative.

On fait donc deux décalages pour les deux images de même valeur D_h , on obtient les points affichés I_g et I_d de coordonnées, (Figure 10.59).

$$I_g = \left(M f \frac{x + e_c/2}{z} - D_h, M f \frac{y}{z} \right) \quad (10.6)$$

$$I_d = \left(M f \frac{x - e_c/2}{z} + D_h, M f \frac{y}{z} \right) \quad (10.7)$$

d'où

$$P_h = -\frac{f M e_c}{z d_o} + 2 \frac{D_h}{d_o} \quad (10.8)$$

On peut alors choisir la profondeur $z = z_{po}$ pour que $P_h = 0$ à cette distance (celle de l'objet principal) :

$$z_{po} = \frac{f M e_c}{2 D_h} \text{ et } P_h = f M e_c \frac{1/z_{po} - 1/z}{d_o} \quad (10.9)$$

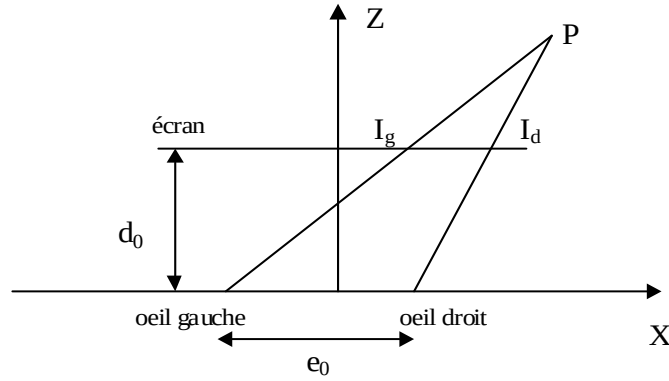


Figure 10.59 : Coordonnées dans l'espace visuel de l'observateur

Remarque : Le système visuel humain permet de percevoir précisément la profondeur relative, la perception absolue étant peu précise. Ceci implique que le cerveau situe les objets les uns par rapport aux autres, un de ceux-ci servant de référence. Dans notre cas, la référence est le bord de l'écran [Cahen, 1990].

On a les relations suivantes :

$$\begin{aligned} (X + e_o/2)/Z &= (X_{I_g} + e_o/2)/d_o \\ (X - e_o/2)/Z &= (X_{I_d} - e_o/2)/d_o \\ Y/Z &= Y_I/d_o \end{aligned}$$

d'où :

$$\begin{aligned} e_o &= Z(X_{I_g} - X_{I_d} + e_o)/d_o \\ Z &= e_o d_o / (X_{I_g} - X_{I_d} + e_o) \\ Z &= e_o d_o / (f M e_c (1/z - 1/z_{po}) + e_o) \end{aligned}$$

Remarque : Z est indépendant de x et de y . Si $z = z_{po}$, alors $Z = d_o$.

Inversement, X et Y dépendent de z :

$$Y = \frac{e_o f M y}{z(f M e_c(1/z - 1/z_{po}) + e_o)} \quad (10.10)$$

$$X = \frac{x e_o f M}{z(f M e_c(1/z - 1/z_{po}) + e_o)} \quad (10.11)$$

On peut alors déterminer les paramètres de l'image reconstituée :

Le grandissement axial μ est égal à Z/z :

$$\mu = \frac{e_o d_o}{f M e_c(1 - z/z_{po}) + e_o z} \quad (10.12)$$

Le grandissement frontal λ est égal à X/x :

$$\lambda = \frac{e_o f M}{f M e_c(1 - z/z_{po}) + e_o z} \quad (10.13)$$

La déformation en profondeur ρ est égale à μ/λ :

$$\rho = \frac{d_o}{f M} \quad (10.14)$$

Le grossissement angulaire γ des champs de vision entre la prise de vue et la restitution visuelle est égal au rapport des tangentes quand les angles sont petits, d'où :

$$\gamma = \frac{X/Z}{x/z} = \frac{1}{\rho} = \frac{f M}{d_o} \quad (10.15)$$

Remarque : ρ et γ sont indépendants de e_o , e_c et de z_{po} .

Formulation des contraintes dans le cas des axes optiques parallèles :

A l'objet principal observé doit avoir une parallaxe nulle (ou presque nulle), donc il faut qu'approximativement $z_{po} = d_c$, ce qui ne pose pas de problème.

B1 la parallaxe horizontale doit être limitée ($<1,5^\circ=0,026\text{rad}$) et pas négative pour les objets vus sur le bord des écrans.

$$P_h = f.M.e_c.(1/z_{po} - 1/z)/d_o$$

Si $z > z_{po}$ il faut $z < 1/(1/z_{po} - 0,026.d_o/f.M.e_c)$

Si $z < z_{po}$ il faut $z > 1/(1/z_{po} + 0,026.d_o/f.M.e_c)$ (à éviter)

B2 la parallaxe verticale doit être nulle, ce qui est le cas pour des caméras à axes parallèles.

B3 Deux plans voisins au niveau de l’affichage, observés simultanément, ne doivent pas avoir une trop grande différence de profondeur. Le seuil S à ne pas dépasser est défini par les proximités, inverses des distances z_1 et z_2 des deux plans :

$$(1/z_1 - 1/z_2) < S$$

Avec z exprimé en mètres, S a pour limite 1 m^{-1} et il est préférable que $S = 0,5 \text{ m}^{-1}$.

10.9.5.3 Les axes des caméras convergent

Dans la scène observée :

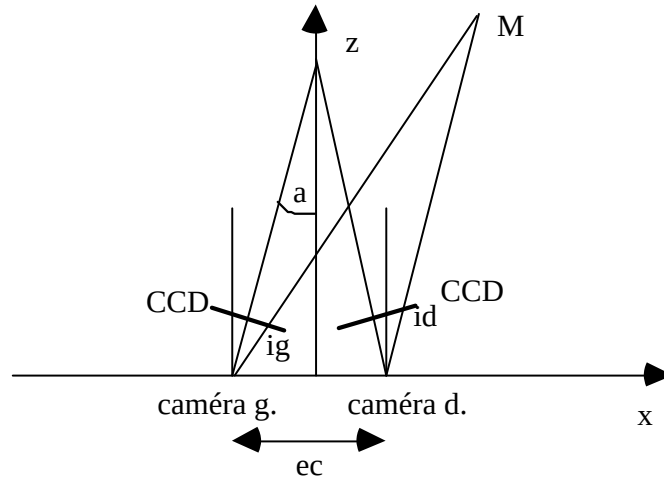


Figure 10.60 : Coordonnées dans la scène observée

Dans le repère centré sur le capteur CCD de la caméra de gauche, on a les coordonnées du point image i_g , (Figure 10.60) :

$$i_g = \left(f \frac{(x + e_c/2) \cos a - z \sin a}{(x + e_c/2) \sin a + z \cos a}, f \frac{y}{(x + e_c/2) \sin a + z \cos a} \right) \quad (10.16)$$

Dans le repère centré sur le capteur CCD de la caméra de droite, on a les coordonnées du point image i_d :

$$i_d = \left(f \frac{(x - e_c/2) \cos a + z \sin a}{-(x - e_c/2) \sin a + z \cos a}, f \frac{y}{-(x - e_c/2) \sin a + z \cos a} \right) \quad (10.17)$$

Dans l'espace visuel de l'observateur :

Ces deux points images donnent sur l'écran les deux points affichés I'_d et I'_g . Leurs coordonnées par rapport au repère centré sur l'écran sont égales aux précédentes multipliées par le facteur M , (Figure 10.61) :

$$I'_g = \left(Mf \frac{(x + e_c/2) \cdot \cos a - z \sin a}{(x + e_c/2) \sin a + z \cos a}, Mf \frac{y}{(x + e_c/2) \sin a + z \cos a} \right) \quad (10.18)$$

$$I'_d = \left(Mf \frac{(x - e_c/2) \cos a + z \sin a}{-(x - e_c/2) \sin a + z \cos a}, Mf \frac{y}{-(x - e_c/2) \sin a + z \cos a} \right) \quad (10.19)$$

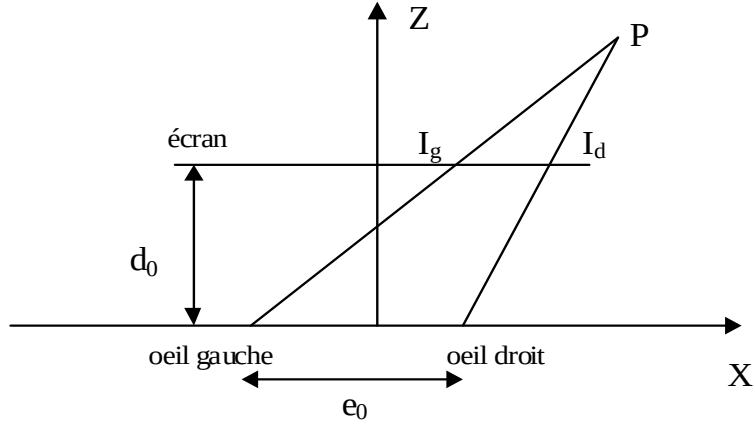


Figure 10.61 : Coordonnées dans l'espace visuel de l'observateur

Avec l'image géométrique identique à l'image psychophysique, on obtient :

$$\begin{aligned} (X + e_o/2)/Z &= (X_{I_g} + e_o/2)/d_o \\ (X - e_o/2)/Z &= (X_{I_d} - e_o/2)/d_o \end{aligned}$$

d'où :

$$e_o = \frac{Z(X_{I_g} - X_{I_d} + e_o)}{d_o}, Z = \frac{e_o d_o}{X_{I_g} - X_{I_d} + e_o} \quad (10.20)$$

on pose alors

$$\begin{aligned} A &= (x + e_c/2) \cos a - z \sin a & B &= (x + e_c/2) \sin a + z \cos a \\ C &= (x - e_c/2) \cos a + z \sin a & D &= -(x - e_c/2) \sin a + z \cos a \end{aligned}$$

On obtient

$$Z = \frac{e_o d_o}{fM \left(\left(\frac{A}{B} - \frac{C}{D} \right) + \frac{e_o}{fM} \right)} \quad (10.21)$$

ce qui donne

$$X = e_o/2 + Z(X_{I_d} - e_o)/d_o \quad (10.22)$$

$$= \frac{e_o(X_{I_g} + X_{I_d})}{2(X_{I_g} - X_{I_d} + e_o)} \quad (10.23)$$

soit

$$X = \frac{e_o d_o \left(\frac{A}{B} + \frac{C}{D} \right)}{2 \left(\left(\frac{A}{B} - \frac{C}{D} \right) + \frac{e_o}{fM} \right)} \quad (10.24)$$

et

$$\frac{X}{Z} = \frac{Mf}{2d_o \left(\frac{A}{B} + \frac{C}{D} \right)} \quad (10.25)$$

Quand a est petit, $\sin a$ est proche de zéro et $\cos a$ proche de 1 :

$$\frac{X}{Z} = \frac{Mf}{2d_o} ((x + e_c/2)/z + (x - e_c/2)/z) \quad (10.26)$$

$$= \frac{Mfx}{zd_o} \quad (10.27)$$

D'où approximativement $\gamma = (X/Z)/(x/z) = 1/\rho = fM/d_o$

Le cerveau ayant un fort pouvoir d'adaptation, le rapport ρ doit être seulement proche de 1. On estime qu'il peut varier entre 0,7 et 1,4 sans trop perturber et fatiguer la vision stéréoscopique.

Formulation des contraintes dans le cas des axes optiques convergent :

Approximativement, la parallaxe horizontale $P_h = (X_{I_d} - X_{I_g})/d_o$ (formule exacte pour un point 3D affiché au centre de l'écran) vaut alors

$$P_h = \frac{Mf}{d_o \left(\frac{C}{D} - \frac{A}{B} \right)} \quad (10.28)$$

En mettant au même dénominateur, en négligeant $\sin a^2$, car a est proche de zéro, et $e_c^2/4$ devant z^2 , on obtient, avec $\tan a = e_c/2d_v$:

$$P_h = \frac{2Mf(x^2 + z^2 - zd_v) \tan a}{z^2 d_o} \quad (10.29)$$

avec l'approximation $\tan a = a$:

$$P_h = \frac{2Mfa(x^2 + z^2 - zd_v)}{z^2 d_o} \quad (10.30)$$

L'équation (10.31) donne la parallaxe horizontale

$$P_h = \frac{2Mfe_c(x^2 + z^2 - zd_v)}{z^2 d_o d_v} \quad (10.31)$$

La parallaxe verticale n'est pas nulle :

$$P_v = \frac{(Y_{I_d} - Y_{I_g})}{d_o} \quad (10.32)$$

Après simplification, on obtient :

$$P_v = \frac{Mfxye_c}{z^2 d_o d_v} \quad (10.33)$$

D'où les conditions B :

$$\mathbf{B1} : |P_h| = \left| \frac{Mfe_c(x^2 + z^2 - zd_v)}{z^2 d_o d_v} \right| < 1.5^\circ = 0,026\text{rad}$$

$$\mathbf{B2} : |P_v| = \left| \frac{Mfxye_c}{z^2 d_o d_v} \right| < 20' = 0,0058\text{rad}$$

$$\mathbf{B3} : (1/z_1 - 1/z_2) < S \quad \text{avec } S = 1 \text{ ou mieux } S = 0,5m^{-1}$$

En conclusion, comparativement par rapport au cas des axes parallèles, cette deuxième solution impose des contraintes complémentaires à cause de la parallaxe verticale. Le faible avantage est une vision plus similaire à la vision naturelle pour les perspectives.

10.10 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Baker, 1987] J. Baker, «*Generating images for a time-multiplexed stereoscopic computer graphics system*», SPIE Vol. 761, p. 44-52, (1987).
- [Cahen, 1990] O. Cahen, livre : «*L'image en relief*», édition Masson, (1990).
- [Chung, 1989] Chung, Harris, «*Brooks, exploring virtual worlds with HMD*», SPIE vol. 1083 3D visualization and display tech. 1989,42-52.
- [Eichenlaub, 1990] J. Eichenlaub, 1990, «*Autostereoscopic display for use with a personal computer*». proc. SPIE : stereoscopic displays and applications. 1256 :156-63.
- [Fuchs, 1995] P. Fuchs, D. Ernadote, D. Maman, and C. Laugeau. «*Téléprésence visuelle stéréoscopique*». In Congrès Informatique 95 : Interface des Mondes Réels et Virtuels, Montpellier, (1995).
- [Holgrem, 1994] D.E. Holgrem, «*Scanner laser displays for virtual reality*». Presence 2(3) :171, (1994).
- [Isono, 1990] I. Haruo, M. Yasuda, NHK science and technical research laboratories, 1-10-11, Kinuta, Setagaya, Tokyo, 157 Japan (1990).
- [Jules, 1971] B. Julesz, «*Foundation of Cyclopean Perception*», University of Chicago press, Chicago, (1971).
- [Kaneko, 1987] E. Kaneko, «*Liquid Crystal TV Display : Principles and applications of liquid Crystal Display*», D. Reidel Publishing Company, (1987).
- [Lipton, 1991] L. Lipton, «*The CrystalEyes handbook*», editor S. Boris, Stereographics Corporation, San Raphael, (1991).
- [McAllister, 1993] D. F. McAllister, «*Stereo Computer Graphics*», éditeur Hanson and Tarjan, Princeton University Press, serie in computer science, (1990).
- [McAllister, 1990] D. McAllister, L. Hodges, «*Rotation algorithm artifacts in stereoscopic images*», Optical Engineering, (august 1990), vol 29 N°8, p 973-976.
- [Moreau, 2001] G. Moreau, P. Fuchs, «*Virtual Reality in the Design Process : from Design Review to Ergonomics Studies*», ESS'01, pp. 123–130 octobre 2001.
- [Perrin, 1998] J. Perrin, P. Fuchs, ENSMP (France) ; C. Roumes, CERMA (France) ; F. Perret, ENSTA (France), «*Improvement of stereoscopic comfort through control of disparity and spatial frequency content*», 3387-14, AeroSense, 13–14 April 1998, Visual Information Processing VII, Proceedings of SPIE Vol. 3387.
- [Perrin, 1998b] J. Perrin, «*Profondeur et binocularité : algorithmie, étude psychologique et intérêt pour l'ergonomie des interfaces stéréoscopiques* », Thèse de l'Ecole des Mines de Paris en collaboration avec le CERMA (Centre d'Etudes et de Recherches de Médecine Aéronautique), (1998).
- [Robinett, 1991] W. Robinett, J. Rolland, «*a computational model for the stereoscopic optics of a HMD*», SPIE vol 1475, Stereo. display and applications II, 1991, 140-160.
- [Teitel, 1990] M. Teitel, «*the eyephone, a HMD*», SPIE vol. 1256 stereoscopic display and applications 1990, 168-171.
- [Tetsuya, 1990] M. Tetsuya, U. Tatsuo, «*Cause of fatigue and its improvement in stereoscopic displays*», Proceedings of SID, Vol. 31/3, p.249-254, (1990).
- [Tilton, 1987] H.Tilton, 1987,«*The 3D oscilloscope : a practical manual and guide*», Englewood Cliffs,NJ : Prentice-Hall.
- [Valyus, 1962] N.A. Valyus, «*Stereoscopy*», (London, New York) Focal Press (1962).
- [Wöpking, 1995] M. Wöpking, «*Viewing comfort with stereoscopic* », Journal SID, 3 :101-103, (1995).

- [Yeh, 1990] Y. Yeh and L. Silverstein. «*Limits of fusion and depth judgment in stereoscopic color display*». In The Human Factors Society, volume 32, pages 45–60, (1990).
- [http1] <http://www.alioscopy.com>

Adresses web des constructeurs :

Interfaces visuelles à support fixe

Ecrans d'ordinateur triple ou stéréoscopique et lunettes associées

Dimension Technologies Inc.	http://www.dti3d.com
I-O Display Systems	http://www.i-glasses.com
Nuvision	http://www.nuvision3d.com
Panoram Technologies	http://www.panoramtech.com
Stereographics	http://www.stereographics.com
Vrex	http://www.vrex.com

Fenêtre immersive

Virtual Research Systems	http://www.virtualresearch.com

Visiobureaux

Barco	http://www.barco.com
Elumens Corporation	http://www.virtual-reality.com
Fakespace Systems	http://www.fakespacesystems.com
TAN	http://www.tan.de/english/intro.html

Visiosalles et visiocubes

EVL of University of Illinois	http://www.evl.uic.edu/pape/CAVE
Fakespace Systems	http://www.fakespacesystems.com
SEOS Displays LTD	http://www.seos.co.uk
Silicon Graphics Inc.	http://www.sgi.fr
TAN	http://www.tan.de
Vrex	http://www.vrex.com

Interfaces visuelles portables

Visiocasques CRT

BAE-Reflectone	http://www.unites.uqam.ca/grri/publications
Fakespace Systems	http://www.fakespacesystems.com
Fakespace Labs	http://www.fakespacelabs.com
Kaiser Electro-Optics	http://www.keo.com
n-Vision	http://www.nvis.com
Virtual Research Systems	http://www.virtualresearch.com

Visiocasques LCD

Cybermind	http://www.cybermind.co.uk
Interactive Imaging Systems	http://www.iisvr.com/NSindex.html
Kaiser Electro-Optics	http://www.keo.com
SEOS Displays LTD	http://www.seos.co.uk
Virtuality Group	http://vtd.fel.tno.nl/virtgrp
Virtual Research Systems	http://www.virtualresearch.com

Vidéolunettes

Canon	http://www.canon.jp
Cgsd	http://www.cgsd.com
Cy-Vis	http://www.personaldisplay.com
Hitachi	http://www.hitachi.com
InterSense	http://www.isense.com
I-O Display Systems	http://www.i-glasses.com
Liquid Image	http://www.liquidimage.ca
MicroOptical	http://www.microopticalcorp.com
MicroVision	http://www.retinaldisplays.com
Olympus	http://www.eye-trek.com
Optics 1	http://www.optics1.com
Personal Mointor	http://www.personalmonitor.com
Siemens	http://www.siemens.com
Sony	http://www.cybermind.nl
Virtual Vision	http://www.virtualvision.com
Xybernaut Corporation	http://www.xybernaut.com

Visiocasques semi-transparentes

ReachIn Technologies	http://www.reachin.se

Page de comparaison de visiocasques	http://www.stereo3d.com/hmd.htm

11 LES INTERFACES À SIMULATION DE MOUVEMENT ET LES INTERFACES À SIMULATION DE CLIMAT

Philippe Fuchs

11.1 GÉNÉRALITÉS

Les sensibilités à la position dans l'espace, au mouvement et aux forces exercées sur les muscles sont les trois domaines liés à la proprioception. Les interfaces proprioceptives se divisent en deux catégories : les interfaces à retour d'effort qui ont pour objectif d'appliquer des forces contrôlables sur le corps et les interfaces à simulation de mouvement qui ont pour objectif de modifier l'orientation du corps dans l'espace ou de lui faire subir des accélérations. Mais nous devons remarquer que même si ces deux catégories sont techniquement différentes, les interfaces de simulation de mouvement imposent des efforts, incontrôlables directement, sur le corps de l'utilisateur, ainsi que des sensations tactiles. Dans la suite de ce chapitre, nous présentons les interfaces à simulation de mouvement.

Techniquement, le déplacement du corps d'une personne, ou de plusieurs dans les cabines de simulation, est théoriquement sans problème technique insurmontable. Il suffit de se donner les moyens, mais en pratique cela impose des mécanismes très puissants et par conséquent à prix très élevés. C'est donc bien à la base un problème mécanique à résoudre pour tout concepteur de ce type d'interface. Le poids des personnes étant ce qu'il est, nous ne pouvons jamais espérer avoir des interfaces de simulation de mouvement à prix très modiques. Il faudra toujours fournir une énergie relativement importante, source imposant des dispositifs techniques conséquents. Vu les investissements qu'elles demandent, ces interfaces ne sont donc fabriquées et commercialisées que pour des applications professionnelles ou pour des salles de jeux et de spectacles. Si elles pouvaient être peu onéreuses, il est bien évident que le marché des jeux vidéo les aurait exploitées pour des jeux personnels.

Nous pouvons les classer en deux grands domaines, qui ne sont pas strictement disjoints :

- les sièges et autres dispositifs mono utilisateur de simulation, qui changent l'orientation du corps de la personne et(ou) qui parfois appliquent des accélérations alternatives et limitées en intensité ;
- les cabines de simulation, multi utilisateurs en général, modifient l'orientation du corps des personnes et leur imposent de fortes accélérations. Dans ces cabines, le dispositif d'affichage du monde virtuel est inclus dans la cabine fermée.

La vision du monde virtuel est donc liée au mouvement de la cabine, ce qui n'est pas le cas avec les sièges de simulation, sauf quelques cas particuliers. Dans les interfaces mono utilisateur, la personne n'est pas forcément assise, elle peut être couchée ou debout. Nous verrons aussi que les plates-formes à 6 ddl des cabines de simulation peuvent être aussi exploitées avec moins de puissance pour les interfaces mono utilisateur. Il n'y a donc pas de frontières bien nettes entre ces deux catégories et il y

a une grande diversité de dispositifs, en particulier en simulant tous les véhicules de transport. Signalons, comme nous l'avons déjà indiqué, que la cabine du véhicule est souvent réelle et non virtuelle, sauf si elle est sujet d'étude.

Les accélérations ne peuvent pas être aussi importantes que désirées ou imposées dans la direction voulue. Aussi sont-elles parfois simulées différemment et ne correspondent pas alors exactement au mouvement réel. Dans l'exemple classique de simulation de véhicule, on baisse la cabine vers l'avant, quand il y a un freinage du véhicule virtuel, pour suggérer au cerveau, grâce aux systèmes vestibulaires, une décélération linéaire, plutôt que de réaliser cette décélération. Car cette dernière imposerait un mécanisme plus cher et surtout plus encombrant. L'utilisateur est abusé sensoriellement car il n'a pas de repère visuel du monde extérieur et parce qu'il voit le véhicule virtuel en train de décélérer dans le monde virtuel. Il n'a donc pas la sensation d'être incliné. Il y a donc dans ce cas précis une limite de décélération ou d'accélération d'intensités maximales inférieure à l'attraction terrestre. Dans tous les cas, il y a des limites physiques et techniques aux simulations de mouvements à accélération importante. Il est impossible actuellement de simuler, par exemple, les accélérations d'une voiture tournant rapidement autour d'un rond-point. Et dans le cas extrême, le simulateur sans biais du départ en fusée exigerait d'être réalisé à partir d'une vraie fusée ou d'un ascenseur à poudre pour vaincre l'accélération terrestre d'une façon identique au cas réel ! En pratique dans ce cas extrême, les simulations de mouvements et d'accélérations ne sont pas parfaitement identiques. Mais cela n'empêchent pas les astronautes à s'entraîner à subir les accélérations les plus fortes imposées à des hommes. C'est le cas limite d'interface de simulation de mouvement : la centrifugeuse pour l'entraînement des astronautes :

A l'opposé, il existe des dispositifs très simples techniquement qui ne sont pas des interfaces de simulation de mouvement. Ceux-ci permettent à la personne de lui imposer des mouvements *non commandables* par l'ordinateur. La personne peut aussi s'imposer elle-même des mouvements. C'est possible si l'utilisateur est accroché à une plate-forme du genre gyroscope. Un traqueur doit être employé pour que la vision du monde virtuel soit en correspondance avec les mouvements de l'utilisateur. Ces dispositifs sont exploités pour les jeux et pour tester les capacités neurologiques et physiques d'un futur pilote, par exemple. Ils peuvent donc être utilisés en réalité virtuelle, comme élément perturbateur de l'environnement virtuel (Figure 11.1) :

Un domaine d'exploitation de ces simulateurs est le domaine des jeux et des loisirs. Dans ce cas, ils sont installés dans des salles de jeux ou sur des camions, permettant un amortissement plus rapide des investissements engagés. Ces cabines de simulation sont équipées d'un ensemble de sièges (seize, par exemple), mais dans ce cas l'interactivité est faible car les participants sont presque toujours passifs par rapport à la simulation proposée. La visualisation peut être collective ou individuelle par visiocasque.

Un autre domaine est celui des transports, qui soient terrestres, maritimes ou aériens (transport ferroviaire, voitures, engins militaires, avion, bateau, etc.). Certains simulateurs sont plus simples à réaliser que d'autres car il y a moins de degrés de liberté à commander ou des accélérations maximales plus limitées. Les simulateurs de train sont construits avec une simple interface visuelle (un seul écran) et ont ou non des actionneurs pour les mouvements. Dans le cas de leur simulation, la cabine est montée sur un système assurant trois déplacements pour subir des accélérations linéaires, latérales et de légères inclinaisons pour simuler les virages. Depuis de nombreuses années, les simulateurs de vol sont conçus pour donner aux pilotes la sensation de se déplacer dans l'espace. C'est un domaine important au niveau économique, dont les solutions

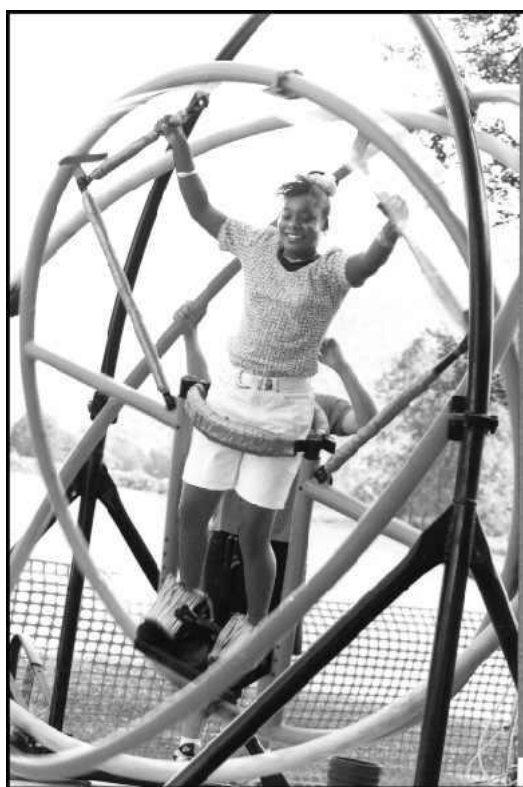


Figure 11.1 : Gyroscope de Aerotrim, copyright photo Aerotrim

proposées sont très satisfaisantes (les pilotes d'avion se forment partiellement et homologuent certaines de leurs compétences sur simulateurs d'avion !). Pour les véhicules terrestres, des plates-formes mobiles 3 ou 6 axes ont été réalisées, mais les simulations sont plus éloignées des phénomènes réels que pour les avions.

Nous ne parlerons pas dans ce chapitre, des effets sur l'homme des simulateurs de mouvement, en particulier du «mal du simulateur», explicité au chapitre sur les effets sur l'homme des environnements.

11.2 SIÈGES DE SIMULATION POUR MONO-UTILISATEUR

11.2.1 GÉNÉRALITÉS

Ces interfaces mono utilisateur ont la possibilité de modifier l'orientation du corps de la personne. Cette dernière est soit assise sur un siège mobile, supporté par une base fixe, soit debout sur une plaque mobile, la personne étant entourée d'une protection circulaire. La personne peut être aussi couchée, si l'application l'exige. Ces dispositifs sont souvent équipés de poignées de commande pour permettre la navigation dans une scène virtuelle. Si des accélérations doivent être subies par la personne, il s'agit en pratique de faibles accélérations dues aux vibrations appliquées au siège ou à son inclinaison. Cette technique est par exemple employée dans des salles de cinéma interactif, où les vibrations et les inclinaisons des sièges sont synchronisées avec les mouvements virtuels projetés à l'écran pour une meilleure immersion. Le siège de simulation est aussi employé dans les salles de jeux pour réaliser un simulateur de voiture dans lequel le joueur navigue dans une scène virtuelle. Pour une meilleure immersion, les inclinaisons du siège sont synchronisés avec le mouvement du volant.

Les sièges de simulation peuvent être classés selon le nombre de degrés de liberté qu'ils imposent à l'utilisateur. Nous pouvons avoir des sièges :

- à un degré de liberté, par exemple, des vibrations verticales pour simuler des secousses suivant l'état d'une chaussée et des suspensions du véhicule ;
- à deux degrés de liberté, par exemple, un siège de moto ,
- à trois degrés de liberté, par exemple, un delta plane ;
- à six degrés de liberté, par exemple, une planche de surf.

Les mouvements des sièges sont commandés par des moteurs électriques, par des vérins pneumatiques ou par des vérins hydrauliques, plus puissants. Ces interfaces ont en général au moins deux inclinaisons, tangage et roulis, 25° à 45° suivant les modèles, et certains peuvent se soulever jusqu'à vingt-cinq centimètres. Un système à six degrés de liberté est plus proche mécaniquement des cabines de simulation. Il est constitué d'un siège se trouvant à deux mètres du sol et de six vérins électro-pneumatiques linéaires très puissants. Trois vérins supportent le siège et les trois autres permettent de faire tourner l'ensemble sur lui-même. Un cas particulier mérite d'être souligné. Le simulateur équestre Persival, construit vers 1987 pour le Cadre noir de Saumur, permet d'appliquer au cavalier des mouvements et des efforts adéquats à son apprentissage, avec affichage de la scène virtuelle sur un écran.

11.2.2 LES DIFFÉRENTS PRODUITS DE SIMULATION MONO-UTILISATEUR

11.2.2.1 Dispositifs à un degré de liberté

Il s'agit souvent dans ce type d'interface des vibreurs, qui sont utilisés seuls ou en complément des plates-formes de mouvement pour rajouter du réalisme par des hautes fréquences dans le déplacement d'un véhicule. Ils peuvent être exploités pour l'étude du confort d'un véhicule ou pour les jeux. Le vérin du vibreur de Siège de Servos and Simulation crée des vibrations dans une direction.

Nous pouvons trouver aussi d'autres systèmes à un degré de liberté comme le simulateur de mouvement de la NAVY, basé simplement sur une rotation autour d'un axe.

11.2.2.2 Dispositifs à deux degrés de liberté

RPI Entertainment vend le Cyberchair, utilisé essentiellement dans les salles de jeu, il est constitué d'un fauteuil ressemblant à celui d'une voiture. Pour le modèle standard, la plate-forme du mouvement fournit 20° en tangage et 30° en roulis. On peut avoir comme option aussi le mouvement en lacet grâce à trois axes.

L'entreprise Flyit Simulation propose un simulateur personnel de mouvement avec au maximum plus ou moins 13° en tangage et en roulis (Figure 11.2).

11.2.2.3 Dispositifs à trois degrés de liberté

Le système DreamGlider de Technologies Dreamality comprend un système supportant un hamac accroché à un trapèze qui fournit à l'utilisateur la sensation de mouvement vertical et de balancement. L'utilisateur contrôle la direction du planeur par les forces qu'il exerce sur une barre de commande, comme pour un deltaplane. Le système est connecté à un PC, qui inclut un générateur des images de synthèse. Trois effets spéciaux ont été développés pour augmenter l'immersion de l'utilisateur. La première de ceux-ci est l'odorat, le système autorise jusqu'à quatre parfums différents selon la hauteur simulée. Le deuxième effet est l'addition d'un servomoteur hydraulique capable de relever et baisser l'appareil DreamGlider pour fournir des accélérations verticales, du type ascenseur. Des ventilateurs sont inclus pour produire la sensation de la brise (Figure 11.3).

Le CyberMotion de CineMotion incorpore deux sièges individuels avec un système de contrôle de mouvement pneumatique qui offre 3 ddl pour la version standard. Il fournit jusqu'à 18° en tangage et en roulis et 125 mm de soulèvement de haut en bas. Chaque siège CineMotion a une manche à balai pour intervenir dans le jeu.

Le «Micro Simulator» de Panoram technologies peut être employé en «petit» simulateur de vol, comme son nom l'indique. Le système triple écran permet d'avoir un champ de vision assez grand. Comme ce système n'est pas trop lourd, il est solidaire du fauteuil mobile. Un nouveau type d'actionneurs pneumatiques est employé pour un faible coût et une faible maintenance. Le tangage et le roulis sont possibles jusqu'à plus



Figure 11.2 : Le Personal Motion Simulator de Flyit Simulation, copyright photo Flyit Simulation



Figure 11.3 : Le simulateur de deltaplane DREAMGlider de Dreamality Technologies, copyright photo Dreamality Technologies

ou moins 30° et de plus ou moins 15° pour le lacet. La bande passante est proche de 200 Hz (Figure 11.4) :

11.2.2.4 Dispositifs à six degrés de liberté

Produit de ViRtogo, le «Cyber Air Base» est une chaise basée sur des vérins pneumatiques qui procurent des mouvements avec 6 ddl, avec un temps de latence inférieur à 40 ms. Le modèle haut de gamme a comme spécifications techniques : 25° en tangage et en roulis, 35° en mouvement en lacet et 23 cm pour le soulèvement de haut en bas (Figure 11.5).

IntellSeat est développé et vendu par Torus Systems. Il est conçu comme une nouvelle solution différente des actionneurs hydrauliques ou électromagnétiques. Ce système utilise six réservoirs pneumatiques individuels placés dans le siège et en arrière du coussin de la chaise pour fournir l'illusion du mouvement. La ceinture du siège est contrôlée pour donner des sensations de mouvement additionnelles en la serrant ou en la défaisant selon les cas pour des sensations d'accélération ou de décélération afin d'augmenter l'immersion de l'acteur dans la scène virtuelle. Un système additionnel utilise un atomiseur contenu dans le bras de la chaise pour vaporiser de petites quantités d'eau sur l'utilisateur.

Un autre produit est constitué d'une moto à 6ddl sur laquelle le joueur s'installe en position allongée et visualise des images à travers des optiques. Tous les mouvements de la moto sont simulés (Figure 11.6).

Le siège pour simulation «Explorer» de Motionbase est un siège fermé avec six degrés de liberté commandés par des vérins électriques et un affichage composé de trois écrans. La société Developpements Dennen a mis sur le marché une nouvelle interface à simulation de mouvement, le «PemRAM 3ddl», basée sur des actionneurs électromagnétiques précis (PemRAM). Dans ces actionneurs électromagnétiques, l'espace sous



Figure 11.4: Le micro simulator de Panoram technologies avec 3 écrans, copyright photo Panoram Technologies



Figure 11.5: Le siège de simulation Cyberchair, copyright photo ViRtogo



Figure 11.6 : Un simulateur de moto

les pistons est rempli avec de l'air à une pression suffisante pour supporter le poids mort de la charge utile. Ce système du contrepoids permet à la base du mouvement de rester stable, en équilibre neutre, et le mouvement dynamique est fourni par les forces impulsives de l'actionneur électromagnétique. Les forces générées par l'actionneur sont ressenties comme des accélérations par l'utilisateur. Le système inclue un algorithme permettant de compenser les changements du centre de gravité à chaque mouvement du corps de l'utilisateur, ce qui rend le système précis, lui permettant de calculer les forces exactes qu'il exercera sur l'utilisateur. Ces actionneurs ont une consommation faible, typiquement un dixième de celle utilisée par les systèmes hydrauliques. La bande passante de ces actionneurs est élevée ce qui leurs procurent précision et rapidité (vibrations dépassant les 30 Hz) Le «PemRAM 3ddl» a un roulis de 20° , un tangage de -24° à $+18^\circ$ et un soulèvement de haut en bas de 180 mm. Le «PemRAM 6ddl» a un roulis et un tangage de 20° , un soulèvement de haut en bas de 200mm, un mouvement en lacet de 25° , un mouvement d'avant en arrière de 250 mm et un mouvement de gauche à droite de 200 mm. Nous pourrions citer, sans être exhaustif, d'autres types de sièges et de dispositifs variés mono utilisateur, comme le simulateur de «planche de surf».

11.3 CABINES DE SIMULATION POUR MULTI UTILISATEURS

11.3.1 EVOLUTION DES SIMULATEURS

Les simulateurs avec mouvement de cabine, nommés souvent simplement «les simulateurs», sont employés depuis plusieurs dizaines d'années en aéronautique et dans le domaine militaire. Ces systèmes permettent des immersions multisensorielles de très hautes qualités, indispensables pour la formation et l'entraînement des pilotes. Les simulateurs sous dôme sont de deux types : le dôme est soit mobile avec la cabine, soit indépendante de cette dernière. Les actionneurs sont les pièces maîtresses du sys-

tème. Ils sont pneumatiques, électriques ou hydrauliques pour les charges de plusieurs tonnes. Les actionneurs sont souvent disposés deux par deux sur chaque coté d'un triangle, permettant les six degrés de liberté sur des débattements de 20 à 30 degrés. Ils peuvent atteindre des vitesses proche d'un mètre par seconde et les accélérations angulaires de la cabine sont supérieures à $100^\circ/\text{s}^2$. Ils doivent pouvoir déplacer des masses de plusieurs tonnes et vaincre l'inertie de la cabine. Leurs caractéristiques dynamiques sont les plus importantes. Suivant leurs capacités de mouvements et d'accélérations durant ceux-ci, les simulateurs pourront ou non répondre aux objectifs de simulation. Rappelons aussi qu'une accélération, d'intensité pas trop élevée, dans un mouvement de translation du véhicule virtuel peut être judicieusement remplacée par une inclinaison de la cabine pour profiter de l'attraction terrestre. Les systèmes vestibulaires des utilisateurs seront identiquement sollicités et si la vision est cohérente, la simulation sera interprétée comme un mouvement de translation accéléré.

Les simulateurs sont installés dans des locaux adaptés et sont fixés sur des socles en béton pour immobiliser leur base. Les investissements et les coûts de fonctionnement imposent un taux d'utilisation important (Figure 11.7).



Figure 11.7: Simulateur d'avion

Mais à la suite des simulateurs d'avion, se sont développés différents simulateurs, d'abord pour les transports : simulateurs de train, de voiture, d'engin militaire et de bateau. Des simulateurs particuliers ont été conçus, comme le simulateur d'engin d'abatage d'arbre «Arvester», réalisé par l'ENSAM pour permettre la formation de la conduite de cet engin forestier. Comme pour presque tous les simulateurs de transport, la cabine ou le poste de pilotage est réel, seul la visualisation est virtuelle (Figure 11.8).



Figure 11.8 : Simulateur d'engin d'abattage d'arbre Arvester conçu par l'ENSAM et l'Institut de l'Image, copyright photo ENSAM

D'autres simulateurs de mouvements ont été dernièrement développés à l'intention des parcs à thèmes pour faire des simulateurs d'engin ou pour faire du «cinéma interactif» (Figure 11.9).



Figure 11.9 : Simulateur pour les loisirs

11.3.2 PLATES-FORMES MOBILES

Avant de présenter les cabines complètes, nous explicitons les différentes plates-formes, qui peuvent être employées pour créer son simulateur spécifique. Servos and Simulation offre des systèmes de commande «tout électrique». Les systèmes de mouvement à deux, trois, cinq, six axes et le siège miniature sont disponibles dans une variété de poids et de taille. En système miniature, ils proposent une plate-forme de petites dimensions (un triangle équilatéral de 37.5 cm de côté) pouvant supporter une charge de

130 Kg. Le système de déplacement électrique à deux axes de Servos est conçu spécialement pour des applications commerciales et militaires. Son coût réduit lui permet cependant de trouver sa place dans l'industrie des sports et de loisirs (Figure 11.10).

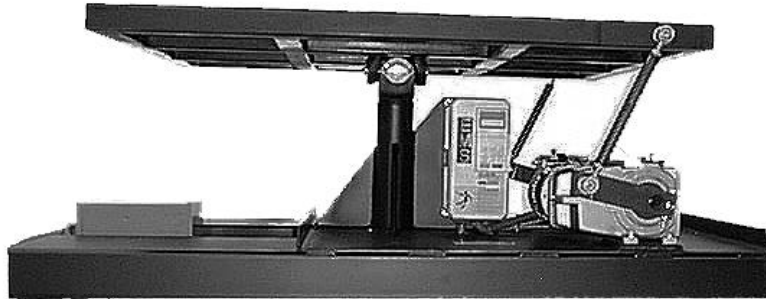


Figure 11.10 : Système de déplacement électrique à deux axes de Servos and Simulation, coyright photo Servos and Simulation

Les vérins sont électriques, ce qui facilite l'installation et la maintenance. Les angles maximaux de tangage et de roulis sont de 15° , avec des vitesses maximales de 45° par seconde. Commercialisé par la société MOOG, le système «4DOF9000 E» a 4ddl. Il peut supporter au-dessus de sa base une charge maximale de 6000 Kg. Ces deux inclinaisons tangage et roulis sont de 17° , avec des vitesses maximales de $24^\circ/\text{s}$. Il a un soulèvement de haut en bas de 0,41 m (vitesse maximale de 0,53/s et accélération de 0,8 G) et une inclinaison d'avant en arrière de 0,57 m (vitesse maximale de 0,36/s et accélération de 0,4 G), (Figure 11.11). La société MOOG utilise des actionneurs électriques, certes faciles à installer et à maintenir, mais qui ont une puissance faible et ne permettent pas une sensation d'accélération forte comme les systèmes hydrauliques.

Commercialisée par la société Motionbase, la plate-forme Maxcue à 6ddl est basée sur des servomoteurs électriques. Elle peut supporter une charge maximale de 1,000kg, 1,500kg ou 3,000kg suivant le modèle. Elle peut fonctionner jusqu'à une fréquence de 25 Hz. Pour le modèle le plus puissant, les caractéristiques sont les suivantes : tangage à $\pm 34^\circ$, le roulis à $\pm 28^\circ$ et le mouvement de lacet à $\pm 44^\circ$ avec une vitesse maximale de $50^\circ/\text{sec}$ environ et une accélération maximale de $15^\circ/\text{s}^2$; le soulèvement de haut en bas jusqu'à 500 mm, l'inclinaison d'avant en arrière et le balancement de gauche à droite jusqu'à 900 mm environ, les trois mouvements avec une vitesse maximale de 0,7 m/s et une accélération moyenne de plus ou moins 0,6 g (Figure 11.12).

Commercialisé par la société Servos and Simulation, le système électrique à six axes est conçu spécialement pour des utilisations commerciales et militaires. Son coût réduit lui permet cependant de trouver sa place dans l'industrie des sports et des loisirs. Les vérins sont électriques supportant une charge de 906 Kg, ayant le tangage, le roulis et le mouvement de lacet jusqu'à 15° avec une vitesse maximale de $15^\circ/\text{sec}$ et une accélération maximale de $15^\circ/\text{s}^2$. Le soulèvement de haut en bas, l'inclinaison d'avant en arrière et le balancement de gauche à droite vont jusqu'à 15 cm avec une vitesse maximale de 0,3 m/s et une accélération maximale de $0,3 \text{ m/s}^2$. La Figure 11.13 représente

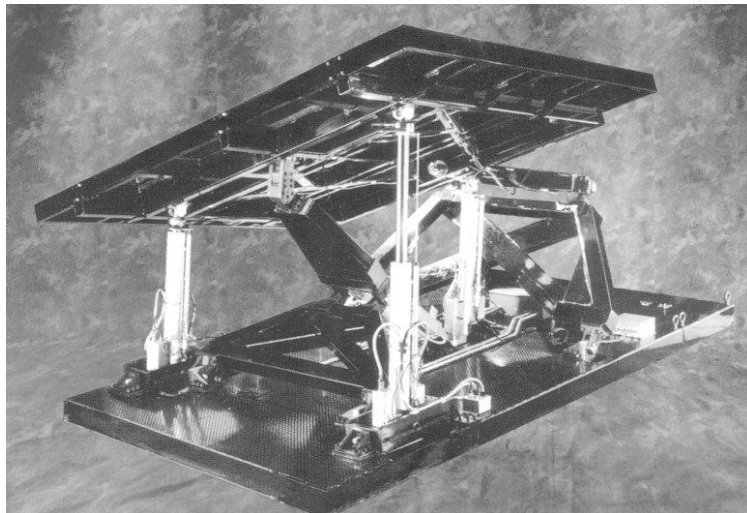


*Figure 11.11 : Plate-forme à 4 ddl avec des vérins électriques de Moog Inc.,
copyright photo Moog Inc.*



Figure 11.12 : La plate-forme Maxcue de Motionbase, copyright photo Motionbase

le système .

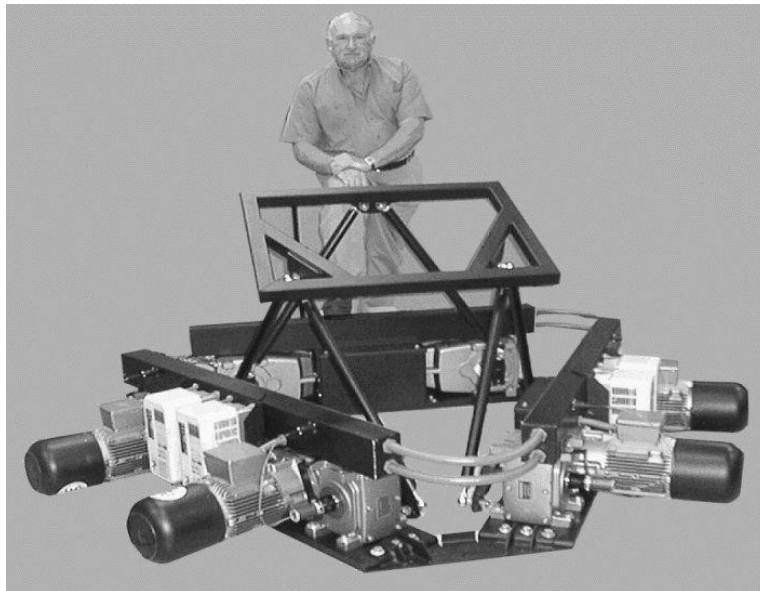


Figure 11.13 : Système de mouvement à six axes de Servos and Simulation, copyright photo Servos and Simulation

11.3.3 LES DIFFÉRENTES CABINES DE SIMULATION

Des cabines de simulation à un seul siège et des cabines à plusieurs sièges sont réalisées et commercialisées par différentes entreprises. La société AI GROUP (anciennement Thomson Entertainment) propose différentes cabines et dispositifs pour les parcs à thèmes. Le *Venturer S2* est une cabine deux places pour les activités ludiques. Il fonctionne avec 3 ddl à partir de 3 moteurs électriques «brushless» (Figure 11.14). Toujours dans la même entreprise, il y a des simulateurs accueillant 4, 10, 20, 25 ou 40 personnes pour le «cinéma interactif» (motion theatres). Des projecteurs d'images monoscopiques ou stéréoscopiques sont incorporés aux dispositifs.

La société Motionbase a conçu une série de cabines de simulation à partir de ses plates-formes, comme la cabine «PLC» à la forme originale pour une personne (Figure 11.15).

Nous ne présentons pas les classiques simulateurs d'avion et d'autres moyens de transport. Nous citons seulement pour finir cette présentation non exhaustive le système le plus complet et le plus complexe au monde pour la simulation d'aéronefs : le VMS (Vertical Motion Simulator) est une création du NASA Ames Research Center pour l'entraînement des pilotes d'avion ou d'hélicoptères et des astronautes. C'est un simulateur à 6ddl à grandes courses en translation : 18,3 m (60 feet) en vertical, 12,2 m (40 feet) et 2,4 m (8 feet) horizontalement, simulant les sensations de décollage des navettes spatiales. Le système VMS, installé dans un grand bâtiment de grande hauteur, peut déplacer dans les 3 axes une cabine qui est interchangeable : ICAB. Celle-ci est fixée sur une classique plate-forme, elle-même liée au VMS. La plate-forme a 3 ddl,



Figure 11.14: Cabine de simulation Venturer S2 de AI GROUP (Thomson Entertainment), copyright photo AI Group



Figure 11.15: La cabine PLC de Motionbase, copyright photo Motionbase

permettant les rotations sur 90° . L'ensemble fait un poids de 70 tonnes qui se déplace grâce à un contrepoids. L'accélération verticale peut atteindre $6,7 \text{ m/s}^2$ (environ $3/4 \text{ g}$).

11.4 CONCLUSION SUR LES INTERFACES À SIMULATION DE MOUVEMENT

Les interfaces à simulation de mouvement sont à un stade stable depuis quelques années. Il n'y a pas de changement important, seulement des variantes avec des accessoires divers. Les actionneurs sont de plus en plus électriques, justifiés par leur bas prix, leur facilité d'installation et de maintenance. Mais en terme de sensations, les systèmes hydrauliques sont largement plus puissants, mais ils sont onéreux. Actuellement, quelques centres de recherches essaient d'exploiter professionnellement des visiocasques à la place des vidéoprojecteurs, qui ont un poids très contraignant quand on les embarque sur les plates-formes. Selon l'usage qu'on veut faire de ces interfaces, les prix varient de plusieurs ordres de grandeurs. Tous ces simulateurs de mouvement sont relativement très chers. Leur achat ne peut se justifier que par un taux d'utilisation important (parc à thèmes) ou par une utilisation professionnelle indispensable (formation, etc.).

11.5 INTERFACE À SIMULATION DE CLIMAT

Dans un certain nombre de cas où il est nécessaire d'évaluer des comportements humains des équipements ou des matériels dans un climat particulier, on peut faire appel aux chambres climatiques. Il est possible de donner à ces ambiances atmosphériques les caractéristiques physiques désirées, que l'on peut contrôler de façon continue. Les chambres climatiques peuvent avoir des températures variant de -70 à $+70^\circ\text{C}$, des taux d'humidité de 30 à 95 %, des vitesses de vent de 0 à 70 Km/h et des rayonnements inférieurs à 1200 W/m^2 .

L'ETAS a réalisé un Générateur d'Environnement Agressif pour le Travail (GEPAT). Le système est constitué d'une cabine représentant sensiblement le volume d'un poste de pilotage de véhicule qui peut être soumise à des mouvements sinusoïdaux. La cabine dispose d'un dispositif de climatisation, de moyens auditifs de communication et de transmission du bruit, de moyens d'observation du candidat et de moyens de mesure de l'environnement. L'ensemble est géré par un PC qui permet de définir les configurations de tests ainsi que les fichiers concernant les sujets et les enregistrements des résultats. Il est schématisé sur la Figure 11.16.

Ce moyen permet de faire réaliser des tests au sujet dans un environnement contraignant : température : 15 à 35°C , hygrométrie : 50%, mouvement : 1 m/s^2 à 10 Hz ou 1 m/s^2 à $0,6 \text{ Hz}$, bruit et confinement. Une des applications est l'entraînement de médecins aux gestes élémentaires de réanimation à l'aide d'un patient virtuel. Il s'agit de donner au médecin l'illusion d'ambiances agressives qui ne sont pas une simulation d'une situation réelle.

Dans un esprit plus «réalité virtuelle», quelques réalisations ont été faites pour simuler par exemple l'écoulement du vent sur la peau, soit à l'aide de systèmes pneumatiques installés dans un blouson, soit de systèmes vibratoires se déplaçant par vague sur le corps.

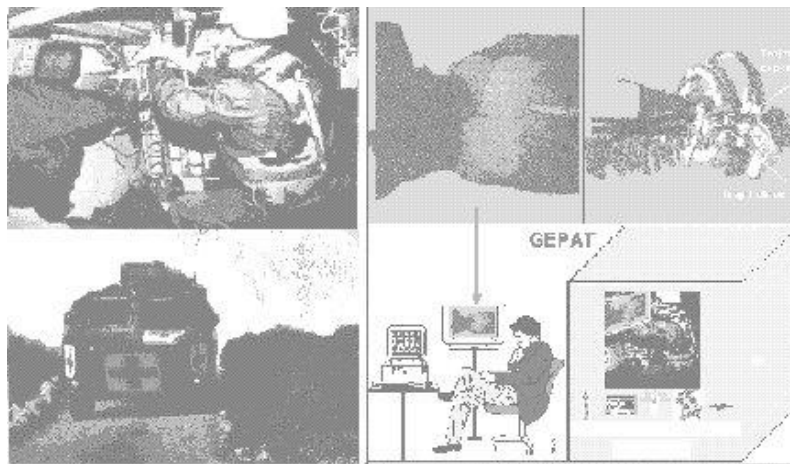


Figure 11.16: Le Générateur d'Environnement Agressif pour le Travail, GEPAT

11.6 RÉFÉRENCES BIBLIOGRAPHIQUES

Adresses Web des entreprises :

Siège de simulation pour mono-utilisateur

Accutone Services	http://www.acesim.com
Amusements international	http://www.gyrostore.com
August Design Inc.	http://www.august-design.com
Dreamality Technologies	http://www.dreamalitytechnologies.com
Flogiston	http://www.flogiston.com
Flyit	http://www.flyit.com
Illusion Inc.	http://www.illusioninc.com
MotionBase	http://www.motionbase.com
Panoram technologies	http://www.panoramtech.com
Servos & Simulation	http://www.servos.com
Simtec	http://www.simtec-gmbh.de
ViRtogo	http://www.ViRtogo.com
VR-Atlantis	http://www.vr-atlantis.com

Cabine de simulation pour multi-utilisateurs

CAE	http://www.cae.ca
Corys	http://www.corys.com
Institut de l'Image	http://www.ai.cluny.ensam.fr
Moog	http://www.moog.com
MotionBase	http://www.motionbase.com
SEOS Displays LTD	http://www.seos.co.uk
Servos & Simulation	http://www.servos.com
ViRtogo	http://www.ViRtogo.com

12 LES INTERFACES OLFACTIVES

Philippe Fuchs

12.1 INTRODUCTION

Nous avons présenté dans le chapitre sur les sens et les réponses motrices humaines le sens de l'odorat. Celui-ci, comme le goût, est stimulé chimiquement et non physiquement. Les études les plus récentes font penser que les odeurs sont spécifiques et non décomposables en quelques odeurs primaires. Ceci implique que les odeurs ne peuvent être créées par combinaison de stimuli élémentaires, comme la vision qui est stimulé par la combinaison de trois couleurs primaires. Il faut donc se contenter pour des simulations virtuelles à émettre l'odeur spécifique réelle dont on veut immerger l'utilisateur. Et si l'on souhaite avoir tout un panel d'odeur, celui-ci doit être préparé et inclus dans l'interface olfactive. Celle-ci n'est donc qu'un *émetteur commandable* d'odeurs préalablement stockées en petites quantités.

Rappelons qu'il est nécessaire pour qu'un produit puisse être senti qu'il présente un degré moyen de volatilité. Les molécules du produit volatil véhiculées dans le courant respiratoire passent alors dans les fosses nasales. Elles y sont absorbées au contact d'une région de la muqueuse revêtue d'une couche de cellules nerveuses. Les interfaces olfactives doivent donc émettre précisément la quantité d'odeurs nécessaire dans le volume d'air adéquat. Le dispositif peut être posé sur un support ou peut être porté autour du cou pour être proche du nez, ce qui facilite l'inhalation des odeurs. Dans le cas d'une interface posée sur un support, pour que les odeurs arrivent plus facilement au nez de l'utilisateur, la circulation de l'air doit être étudiée et si possible contrôlable.

Peu de systèmes sont proposés commercialement, le marché étant naissant et limité dans les applications, quelles soient de réalité virtuelle ou non. On peut mettre à part les applications en parfumerie où le mélange de fragrances est en soit la fonctionnalité recherchée. La plupart des autres applications n'exploitent le sens de l'odorat que pour des fonctionnalités secondaires. Les récents travaux et connaissances sur les odeurs, sur leurs perceptions humaines et sur les interfaces olfactives sont présentés dans l'article [Barfield, 1996] et sur le site Internet du Laboratoire Neurosciences et Olfaction de l'Université Claude Bernard Lyon1 [http1].

12.2 LES INTERFACES OLFACTIVES COMMERCIALISÉES

12.2.1 L'INTERFACE OLFACTIVE DE OLFACOM

OlfaCom est une interface olfactive développée par A.C.2.i qui permet de diffuser des arômes et parfums grâce à des cartouches amovibles. L'interface est pilotée par ordinateur par l'intermédiaire d'un port série. L'interface diffuse régulièrement dans son intégralité une composition parfumée complexe. La durée de fonctionnement d'une cartouche est de 3 mois normalement pour une utilisation continue. La diffusion par sublimation s'effectue par passage d'air dans l'échangeur sphérique à obturation pivotante de la cartouche amovible (dimensions 60 x 60 x 47 mm). Le dispositif peut

fonctionner entre 0 et 70° Celsius. La partie active de l'échangeur sphérique est composée du mélange maître aromatique, qui a été incorporé naturellement dans des granulés du type Polyether Block Amide (PEBA). La dose incorporée se fait à un certain pourcentage selon l'effet de la restitution désirée. L'incorporation en l'état du mélange maître sans additif et sans transformation chimique ou thermique, assure une restitution intégrale sans distorsion olfactive (Figure 12.1).

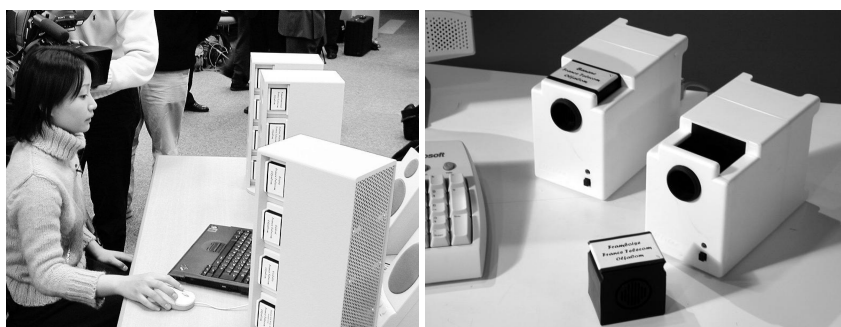


Figure 12.1 : L'interface olfactive fixe d'Olfacom sous forme de tour ou de petit diffuseur, copyright photo Olfacom

Ce type d'interface olfactive est exploité dans différentes applications, de réalité virtuelle ou non. L'interface a été employée dernièrement à Paris chez Sephora pour une présentation olfactive d'un nouveau parfum. L'armée américaine, division des chars d'assaut, utilise le système OlfaCom dans leur salle d'immersion. Les odeurs complètent l'image et le son, pour recréer des conditions plus exactes : aux odeurs de diesel s'ajoutent celles des sous-bois et de la poudre.

12.2.2 L'INTERFACE OLFACTIVE DE RUETZ

En coopération avec l'école supérieure du cinéma de Munich et un réalisateur de théâtre, la société Ruetz s'est lancée dans la conception d'un diffuseur personnel de senteurs, le Sniffman, piloté à distance et en synchronisme avec une scène présentée à un spectateur ou avec une page Web. Ruetz Technologies a travaillé en partenariat avec France Télécom et L'ISIPCA de Versailles (école spécialisée dans la parfumerie) pour des applications multimédia et de télévision. Des essais d'association de production d'odeurs en rapport avec une scène particulière présentée sur Internet ou à la télévision ont été réalisés.

Le Sniffman est une interface qui diffuse des odeurs, mais celles-ci ne sont pas, bien sûr, transportées par câble d'un site à un autre. Les diffuseurs dans chaque lieu doivent être chargés avec des capsules odoriférantes correspondant aux stimuli nécessaires à la simulation olfactive. Le début, la durée et la périodicité des diffusions ainsi que l'intensité des odeurs sont paramétrés de manière à éviter au maximum les phénomènes de rémanence pour l'utilisateur. Le Sniffman s'apparente à un baladeur très léger qu'on porte autour du cou, autorisant ainsi à chacun une liberté de mouvements et permettant aux sources olfactives d'être proches du nez. Ce modèle contiendra d'ici un an près de 200 senteurs micro dosées et fidèlement restituées.

Pour la télévision, les diffuseurs sont dissimulés de chaque côté de l'ordinateur. Destinée à une utilisation plus collective autour d'un téléviseur, le dispositif illustre quant à lui les usages possibles d'une chaîne TV interactive odorante : météo, visite virtuelle d'appartements, offres promotionnelles de produits et de voyage, etc. Ces prototypes doivent s'améliorer par une miniaturisation et par le développement de nouveaux conditionnements des parfums sous formes solides et concentrés. Cela permettra de stocker 200 odeurs sur des cartes de format «carte de crédit». Le développement de la télévision numérique et de la technologie Bluetooth permettra une communication sans fil. Ruetz Technologies et France Télécom assurent être prêts à sortir des modèles commercialisables dans moins d'un an.

12.2.3 L'INTERFACE OLFACTIVE DE AROMAJET

Pinoke est une interface olfactive développée par AromaJet qui permet de diffuser des arômes grâce à dispositif au cou ou installé sur différents supports : visiocasque, siège de salle de spectacle, etc. Un système de micro diffuseurs, développé par MicroFab Technologies pour les micro injections dans diverses industries, est exploité dans les interfaces de AromaJet. Avec ces dispositifs, on peut créer différentes odeurs à partir de 16 ingrédients aromatiques de base (dans le futur, il y aura 32 ingrédients). Le mélange est programmé par l'ordinateur et chaque ingrédient est dosé par incrément de 1% (Figure 12.2).



Figure 12.2 : L'interface olfactive portable Pinoke développée par AromaJet, copyright photo AromaJet

Différentes applications existent : avec le dispositif Kiosk, l'utilisateur peut proposer à tout client de sentir en quelques secondes différents parfums avec diverses odeurs :

rose, orange, jasmin, etc. Après acceptation du mélange par le client, le système génère le parfum testé.

D'autres applications peuvent être réalisées par l'interface de AromaJet dans le domaine du jeu, du spectacle, de la réalité virtuelle et d'Internet. Dans ce dernier domaine, on peut communiquer des odeurs d'un site à un autre en transmettant seulement des données numériques (les concentrations des seize ingrédients), (Figure 12.3).

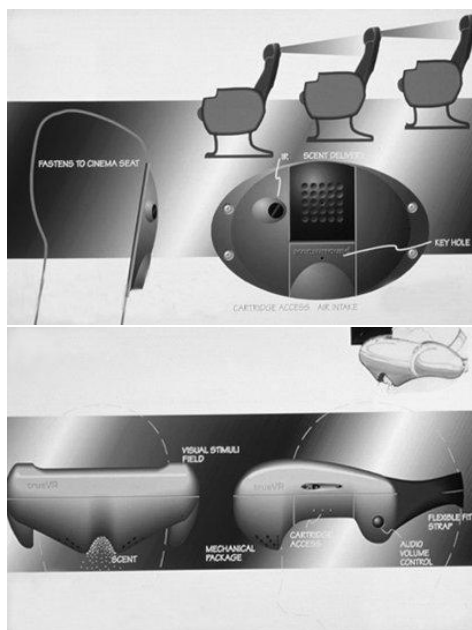


Figure 12.3 : L'interface olfactive d'AromaJet incorporée dans un siège ou dans un visiocasque, copyright photo AromaJet

12.2.4 L'INTERFACE OLFACTIVE DE AEROME

La société Aerome propose son produit «ScentCartridge» pour diffuser des odeurs. Il est composé de plusieurs tubes de verre qui sont remplis de granulés pour stocker les molécules des odeurs. L'interface est capable de garder les produits sur de longues périodes (au moins une semaine). En version standard, six tubes sont utilisés. La diffusion des odeurs est commandée par des valves précises, contrôlées par un microprocesseur. Les jets d'air passent par les tubes et sont émis vers l'utilisateur en une fraction de seconde (Figure 12.4).

12.2.5 L'INTERFACE OLFACTIVE DE DIGISCENTS

Digiscents a développé un ensemble complet matériel et logiciel pour l'émission d'odeur. L'interface olfactive, iSmell, peut créer jusqu'à mille odeurs sur demande. Ses logiciels sont adaptés à la simulation d'odeur sur les pages Web d'Internet, pour les jeux,

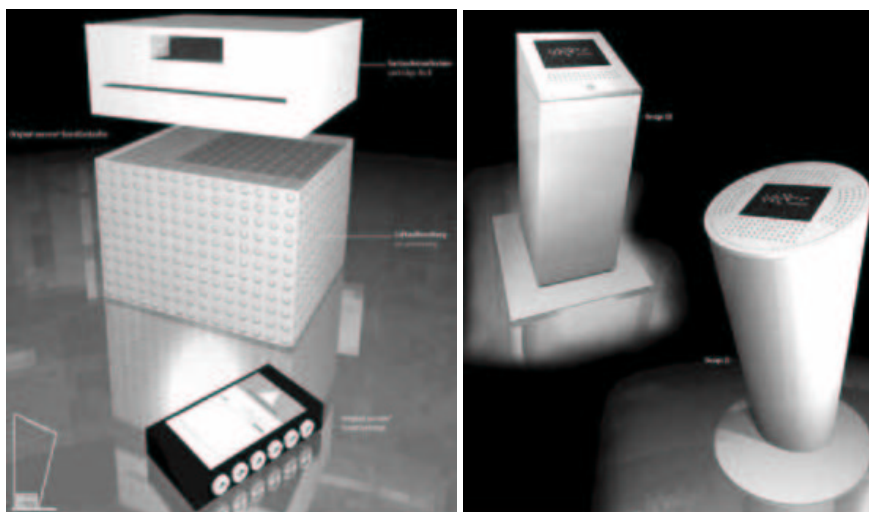


Figure 12.4 : Le diffuseur d'odeur ScentCartridge et les bornes de diffusion d'odeur de Aerome, copyright photo Aerome

le cinéma, les applications multimédia, etc. L'interface iSmell est similaire aux autres interfaces olfactives.

12.3 CONCLUSION

Les interfaces olfactives ont été développées dans la perspective du domaine multimédia et surtout d'Internet. Mais si les produits existent et sont fiables techniquement, rien n'est moins sûr qu'en à l'essor commercial de ces interfaces. Elles ne sont pas indispensables pour l'internaute et elles ne seront intéressantes dans ce secteur que si une masse critique d'ordinateurs seront équipés. Par contre, pour les applications de réalité virtuelle, le précédent problème d'équipement ne se pose pas, les interfaces olfactives commercialisées étant normalement peu onéreuses par rapport aux autres interfaces. Ce n'est donc pas un blocage technico-économique qui empêche quiconque d'exploiter ce sens, perçu «secondaire» par l'homme. La question de l'exploitation de ce sens est donc plus orientée sur son intérêt par rapport aux objectifs de l'application. Doit-on simuler des odeurs, en nombre limité ? Quelles immersions sensorielle et mentale peut-on en attendre ?

12.4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Barfield, 1996] W. Barfield, E. Danas, «Comments on the use of olfactory displays for virtual environments», *Presence*, vol. 5, N° 1, 109-121, MIT, Massachusetts, 1996.
- [http1] <http://olfac.univ-lyon1.fr/olfac>, Laboratoire Neurosciences et Olfaction de l'Université Lyon1.

Adresse web des entreprises :

A.C.2.i	http://www.olfacom.com
Aerome	http://www.aerome.com
AromaJet	http://www.aromajet.com
Digiscents	http://www.digiscents.com
Ruetz	http://www.ruetz.de
France Telecom	http://www.francetelecom.com/vfrance/actualite

13 LES PRIMITIVES COMPORTEMENTALES VIRTUELLES

Philippe Fuchs

Ce chapitre n'expose qu'une première analyse incomplète des solutions envisageables pour les Primitives Comportementales Virtuelles. Les PCV sont les objectifs du niveau d'Immersion et d'Interaction fonctionnelles. S'il peut être réaliste de bien définir leurs caractéristiques intrinsèques après analyse détaillée du niveau I² fonctionnelles, il peut être très aisé ou au contraire très compliqué de concevoir des solutions pour les réaliser, à cause principalement des études multidisciplinaires qu'elles induisent. Il s'agit donc dans ce chapitre que d'une modeste contribution à un vaste problème ouvert dans le champ de la réalité virtuelle.

13.1 RAPPEL SUR NOTRE DÉMARCHE EN RÉALITÉ VIRTUELLE

13.1.1 CONCEPTS DE BASE

Nous avons présenté le fondement de notre approche de la réalité virtuelle qui est basé sur un schéma hiérarchique en trois niveaux et aussi sur un découpage transversal entre le sujet et le monde virtuel : parallèlement aux différents niveaux d'immersion et d'interaction (I²) sensori-motrices et mentales pour la personne, nous avons deux niveaux de fonctionnement logiciel pour le monde virtuel. Symétriquement aux I² sensori-motrices, l'ordinateur doit gérer la partie logicielle temps réel (noyau temps réel et drivers pour les interfaces matérielles), permettant une réalisation physique du monde virtuel. Face aux I² mentales, la partie logicielle spécifique de réalité virtuelle doit gérer la modélisation comportementale du monde virtuel. Cette partie logicielle doit fournir des *Aides Logicielles Comportementales* (ALC) de «motricité et de perception programmées», associées aux schèmes exploités, pour faciliter les I² mentales de la personne, comme nous l'avons explicité précédemment.

Au niveau I² fonctionnelles, par rapport à l'application et ses objectifs, il faut se poser la question suivante : quelles sont les activités que l'utilisateur doit exécuter ? Dans toutes les applications de réalité virtuelle, les activités du sujet sont toujours décomposables en quelques comportements de base que nous avons appelés les «Primitives Comportementales Virtuelles» (PCV). Il faut donc au niveau des I² fonctionnelles bien définir les PCV et leurs spécificités. Quelque soit l'application, nous avons déjà précisé que les PCV peuvent être regroupées en quatre catégories :

- observer le monde virtuel ;
- se déplacer dans le monde virtuel ;
- agir sur le monde virtuel ;
- communiquer avec autrui.

Dans la dernière catégorie, le sujet peut aussi avoir besoin de communiquer avec le système informatique, pour modifier sa configuration, pour donner des ordres, etc. Mais

dans ce cas, le sujet n'est plus en immersion dans un environnement virtuel. On se situe plus alors dans le domaine classique de dialogue homme-machine des IHM.

Nous allons développer dans les paragraphes suivants les solutions envisageables pour chaque classe et sous-classes de PCV en précisant les interfaces comportementales exploitables, les processus cognitifs mis en jeu (schèmes, métaphores...) et les Aides Logicielles Comportementales (ALC) associées de motricité et de perception programmées. Les ALC apportent des aides dans la tâche de l'utilisateur. Le concepteur doit y réfléchir pour proposer celles qui aident efficacement l'utilisateur, tout en respectant l'objectif des I² fonctionnelles. Mais avant de développer sur les PCV et les ALC, nous rappelons notre schéma de référence de la réalité virtuelle qui nous sert de canevas dans notre démarche de conception. Il permet de resituer les différents concepts de notre approche, Figure 13.1.

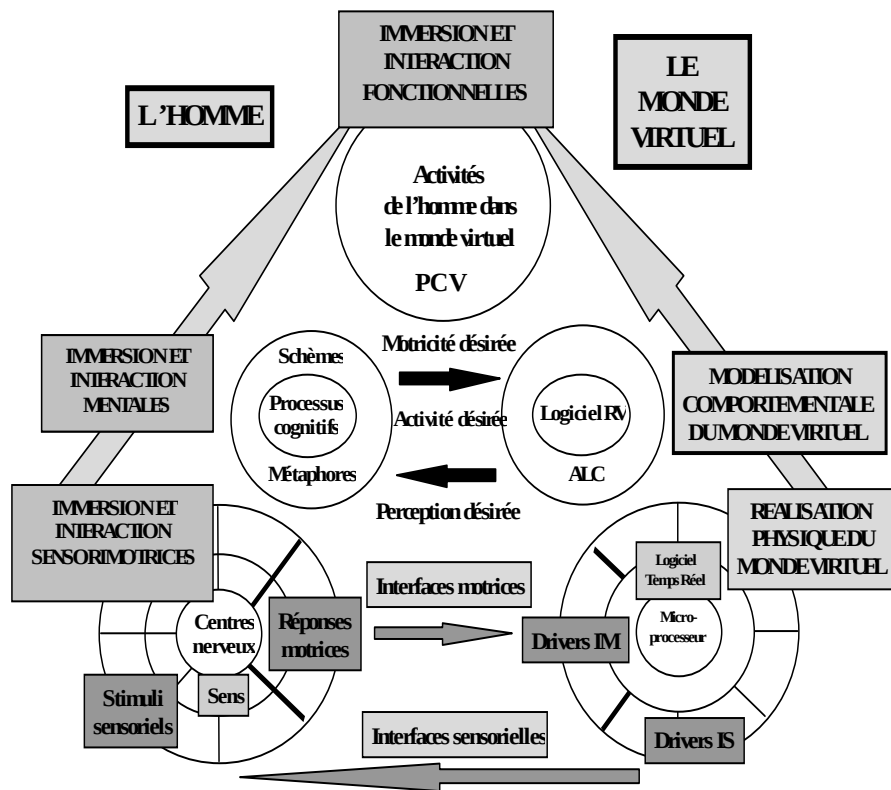


Figure 13.1 : Schéma technocentrique de référence en RV

Avec ce schéma pluridisciplinaire, nous avons clarifié les notions d'immersion et d'interaction. Dans l'optique de la conception d'un dispositif RV, au lieu de chercher à reproduire le plus fidèlement possible la réalité, nous avons indiqué que nous recherchons l'**optimum d'I² fonctionnelles** liées à l'application et par analyse et déduction aux optimums des I² mentales et sensori-motrices. Pour cela, au niveau I² fonctionnelles, suivant la finalité de l'application, les PCV sont à déterminer précisément. Ensuite, le concepteur doit déterminer les solutions les plus efficaces. Nous avons indiqué

la chronologie de notre démarche de conception que nous rappelons :

- étape 1 : détermination des I² fonctionnelles ;
- étape 2 : détermination des PCV requises à partir des I² fonctionnelles ;
- étape 3 : conception des Interfaces Comportementales à partir des PCV, (Schèmes, substitution sensori-motrice ou métaphores – artefacts - sens et réponses motrices sollicités) ;
- étape 4 : conception des ALC à partir des IC. ;
- étape 5 : conception du logiciel RV (programmation des ALC, des drivers des IC...).

Après cette conception, il faut vérifier la faisabilité des choix au niveau I² sensori-motrices : caractéristiques de l'interface suffisante pour la PCV, ordinateur assez puissant, coût acceptable, etc. Ensuite, l'efficacité des choix au niveau I² mentaux doit être vérifiée par des tests avec les utilisateurs finaux. Deux types de tests sont à envisager : tests d'utilisabilité de l'interface et tests de réalisation optimum de la PCV par rapport à l'application. A partir des PCV déterminées en étape 2, nous allons indiquer les solutions envisageables pour la conception des Interfaces Comportementales, (étape 3). Mais avant, nous précisons sous quels critères ont pu établir une taxonomie des PCV.

13.1.2 TAXONOMIE DES PCV

Nous pouvons faire différentes sous catégories de PCV à partir des 4 PCV de base pour étudier les différentes solutions. Nous privilégions deux principaux critères :

- L'importance de la PCV par rapport aux I² fonctionnelles : est-elle sujet d'étude ou est-elle simplement indicative d'une activité à entreprendre par le sujet ? Dans le premier cas, la PCV doit être a priori la plus similaire du schème sensori-moteur en environnement réel. Dans le deuxième cas, elle peut être plus proche d'une métaphore virtuelle ou d'un schème par substitution sensori-motrice ;
- Les caractéristiques spatio-temporelles de la PCV et de l'environnement virtuel sont à prendre en compte pour le choix des solutions des interfaçages comportementaux. Par exemple, l'étendu du champ visuel d'observation, les degrés de liberté du déplacement, les dimensions de l'espace virtuel, etc.

Un autre critère sera pris en compte pour la PCV observation : les sens humains mis en jeu pour l'observation de l'environnement.

Les 4 catégories de PCV peuvent être associées deux à deux et cela influence les choix à faire pour les interfaces comportementales. On peut avoir souvent la PCV «observer» associée à la PCV «déplacer». On peut avoir, mais rarement, la PCV «déplacer» associée à la PCV «agir». La PCV «agir» implique presque souvent en même temps la PCV «observer». Le fait que la PCV «communiquer» se fasse seule ou en association avec les autres PCV, n'intervient pas a priori dans les différentes solutions pour cette PCV. L'association de deux types de PCV peut interdire certaines solutions, comme par exemple l'emploi d'un visiocasque (PCV «observer») avec un tapis roulant (PCV «déplacer») difficilement envisageable sans risque de chute. C'est au concepteur de prendre en compte ces incompatibilités que nous ne précisons pas en général dans

les différentes solutions, listées dans les prochains paragraphes. Toujours au niveau du couplage de plusieurs PCV, le fonctionnement des sens, des réponses motrices et du cerveau doivent aussi être analysés et pris en compte. Par exemple, dans le déplacement naturel à la marche, l'observation visuelle n'est pas indépendante : lors d'un changement de direction, c'est la tête qui tourne en premier par anticipation du mouvement [Berthoz, 1998]. Action et perception étant souvent liées, le couplage de PCV est un problème assez fréquent qui complexifie les choix des interfaces comportementales. Mais par manque de recul et aussi parce qu'il est plus simple en première analyse d'étudier une PCV isolément, les couplages entre PCV ne seront pas développés en tant que tels. Ils seront juste mentionnés dans des cas de solutions contradictoires pour chaque PCV. Les études et les résultats des physiologistes sur le fonctionnement du cerveau au niveau de la perception et de l'action seront primordiales pour obtenir des immersions et des interactions efficaces dans un environnement virtuel.

13.2 PRIMITIVES COMPORTEMENTALES VIRTUELLES D'OBSERVATION

13.2.1 CLASSIFICATION

Comme nous venons de l'indiquer, les PCV «observer» sont différentes suivant les sens humains mis en jeu pour l'observation de l'environnement. Les sens sont choisis a priori au niveau des I^2 fonctionnelles. Il y a donc trois sous PCV «observer» : observation visuelle, observation auditive et observation tactile. Nous pourrions aussi parler d'observation kinesthésique et proprioceptive mais en général ce ne sont que des sous tâches dépendant des PCV «agir» ou «déplacer». En conséquence elles ne sont pas étudiées dans ce paragraphe. Mais le lecteur peut toujours trouver des contre-exemples : quand vous allez au marché, vous soupesez le melon pour observer sa qualité de mûrissement. Les trois sous PCV «observer», visuelle, auditive et tactile, peuvent être associées par combinaison de sens.

Au niveau des I^2 fonctionnelles, on peut aussi étudier et déterminer si l'observation a pour but de s'orienter dans l'espace, de comprendre l'environnement virtuel ou de chercher quelque chose ou quelqu'un dans cet environnement. On doit se poser la question de savoir s'il faut reconstruire mentalement l'espace tridimensionnel 3D. L'utilisateur doit-il mémoriser tout l'espace visité ou seulement là où il se trouve ? Une reconstruction de la topologie bidimensionnelle (2D) de l'espace suffit-elle ? En exemple, sur un terrain, même vallonnée, l'observateur s'en fait a priori une représentation bidimensionnelle. Par contre pour un bâtiment à étage, la représentation est tridimensionnelle, même si elle est souvent une extension d'une représentation bidimensionnelle et référencée à la verticale du lieu (en station spatiale, cela doit être moins évident).

Le nombre d'observateurs simultanés est un critère qui doit être pris en compte car il n'y a que certaines interfaces visuelles qui soient multi-utilisateurs. Les dimensions métriques et angulaires de l'espace observé sont à déterminer car elles influencent directement le choix de l'interface comportementale. Les caractéristiques spatiales de la PCV «observer» sont prises en compte pour la classification.

13.2.2 OBSERVATION VISUELLE

Une observation visuelle se fait sur différentes zones spatiales. Nous les avons classées en quatre cas, mais nous avons un continuum entre ces cas. Les zones spatiales d'observation ne tiennent pas compte de la PCV de déplacement qui peut augmenter la zone d'observation. C'est donc avec la restriction d'un **observateur immobile dans l'environnement virtuel** que l'on étudie les principales solutions :

- 1 le point de vue est fixe, donc le champ de vision est limité à une zone conique assez étroite ;
- 2 les mouvements de la tête sont horizontalement limités à moins de 180° et verticalement limités à moins de 90° ;
- 3 les mouvements de la tête sont horizontalement et verticalement plus amples que le cas précédent ;
- 4 l'observation est centrée sur une petite zone fixe de l'environnement virtuel, le point de vue de l'observateur pouvant tourner autour de cette zone.

Dans le premier cas, un écran d'ordinateur suffit si l'espace est petit ou, si l'espace est grand et doit être vu à l'échelle 1, un grand écran avec vidéoprojection.

Dans le deuxième cas, s'il n'y a qu'un observateur, plusieurs interfaces visuelles sont possibles : triple écrans, petit dôme visuel, visiosalle (salle immersive) ou visiocube. S'il y a plusieurs observateurs, seuls les deux derniers types d'interfaces sont possibles.

Dans le troisième cas, l'interface à employer est le visiocasque, sauf si la résolution n'est pas suffisante ou si le port du visiocasque est contraignant pour l'observateur. Si on ne peut pas employer un visiocasque, l'emploi d'un visiocube est possible si le coût n'est pas rédhibitoire. Autrement, l'emploi d'un seul écran (moniteur ou grand écran) est exploitable avec un traqueur lié à la tête de l'observateur. Une ALC pour augmenter le champ de vision doit être bien choisie pour être efficacement exploitable. Par exemple, pour l'avoir testée au Centre de Robotique de l'Ecole des Mines de Paris, quand l'utilisateur tourne sa tête d'un angle suffisant par rapport à la normale à l'écran, le point de vue tourne aussi (le centre de rotation est la tête de l'utilisateur). La vitesse de rotation est soit fixe, soit proportionnelle à l'angle de rotation. Ceci n'est exploitable qu'après un apprentissage qui n'est pas trop long. Mais l'orientation dans le monde virtuel peut poser problème.

Dans le quatrième cas, une fenêtre immersive peut bien convenir, si les mains de l'utilisateur sont libres. Autrement, on peut exploiter un écran avec un traqueur sur la tête. On doit réaliser cette observation avec une ALC similaire au type précédent, si ce n'est que le point de vue tourne autour de la zone fixe d'observation. Un visiocasque peut aussi convenir si la résolution exigée n'est pas trop grande ou si le champ de vision, tête immobile, doit être assez grand, sans être immense car dans un visiocasque, c'est hélas techniquement limité.

Remarques :

Il faut se poser la question de savoir si la verticale du lieu virtuel restera toujours identique, donc stable, à la verticale du lieu réel où se situe l'observateur. Dans ce cas, une ALC doit contraindre la verticale du lieu virtuel à rester identique si on craint que

l'utilisateur incline involontairement sa tête traquée devant un écran ou dans un visiocasque. Dans le cas contraire, en vol d'avion par exemple, l'observation de l'espace est plus difficile, surtout pour s'orienter. Une autre question à se poser concerne la position des yeux de l'observateur : resteront-ils toujours sur un même plan horizontal virtuel ? Si ce n'est le cas, par exemple maintenance de dispositifs mécaniques complexes (cas d'observation N° 4), seul le visiocasque peut convenir. D'une façon plus générale, il faut bien analyser les mouvements obligatoires¹ du repère lié à la tête de l'observateur par rapport à un repère fixe dans le monde virtuel.

13.2.3 OBSERVATION ACOUSTIQUE

L'observation acoustique est assez simple à analyser à partir des I² fonctionnelles. S'il s'agit de fournir un son monophonique ou stéréophonique de l'environnement virtuel, il n'y a pas de difficultés technico-économiques ou ergonomiques. Par contre, s'il s'agit de fournir un son spatial ou 3D, c'est à dire des sons provenant de sources sonores localisées dans l'espace 3D, les solutions techniques existent avec du matériel spécifique et avec l'appui de spécialistes. Le nombre de sources sonores influence directement le prix du matériel. Si le son spatial ne peut être réalisé, l'observation visuelle et auditive simultanément de la source sonore, couplée à un son monophonique ou stéréophonique, peut être un palliatif intéressant. L'effet Doppler associé à un objet mobile peut être aussi exploité comme ALC.

13.2.4 OBSERVATION TACTILE

L'observation tactile est peu utilisée car il n'y pas beaucoup d'offres d'interfaces efficaces. La détermination des caractéristiques d'un contact (état de surface, vibrations, pression et température) peut être aidée par l'observation visuelle associée. Des métaphores virtuelles peuvent être aussi employées, telles que des sons variables de différentes caractéristiques en fonction des types de contacts. Par exemple, la variation spatio-temporelle de l'intensité sonore est fonction de l'état de surface et du glissement du point de contact.

Si l'observation tactile fait partie des I² fonctionnelles sans que les observations proprioceptives (retour d'effort) en fassent partie, les solutions d'interfaces tactiles existent. Mais s'il faut en même temps des retours d'effort, ceux-ci vont prendre le dessus mécaniquement et les retours tactiles seront difficilement contrôlables pour être bien simulés.

13.3 PRIMITIVES COMPORTEMENTALES VIRTUELLES DE DÉPLACEMENT

13.3.1 PROBLÉMATIQUE D'UN DÉPLACEMENT DANS UN MONDE VIRTUEL

Rappelons que la sensation de déplacement, chez l'homme, est donnée principalement par le système visuel et le système vestibulaire. Plus précisément, les sens réceptifs pour la marche sont les capteurs proprioceptifs des jambes, les récepteurs tactiles aux plantes des pieds, la vision et le système vestibulaire pour le sens kinesthésique. Dans

¹Obligatoires par rapport aux I² fonctionnelles.

un système de réalité virtuelle, en général, le sujet change de lieu virtuel sans avoir à déplacer son corps réel sur une zone de plus d'un mètre : il y a donc un problème technique de base pour la conception de l'interface de déplacement de l'utilisateur qui en général ne se déplace pas ou très peu dans le monde réel et se déplace sur une grande distance dans le monde virtuel.

On peut, au plus simple, facilement donner une impression de déplacement en stimulant la vue : c'est le principe de base de bien des dispositifs d'affichage. Le système vestibulaire dans l'oreille interne apporte les informations relatives aux accélérations, vitesses angulaires et forces d'inertie subies par la tête. En stimulant ce système et la vision, comme dans les simulateurs de transport à cabine mobile, on procure une sensation plus forte de mouvement. Si physiquement la personne marche sur un tapis roulant devant un écran, ce sont les capteurs proprioceptifs des jambes, les récepteurs tactiles aux plantes des pieds et la vision qui sont stimulés, mais pas le système vestibulaire, la tête étant (presque) immobile [Burkhardt e.a., 1999]. Donc, dans tous les cas précédents, le déplacement virtuel ne stimule pas tous les récepteurs sensoriels sollicités pendant un déplacement à pieds dans le monde réel. Si on veut la totale correspondance, il faut que le déplacement à pieds se fasse obligatoirement sur les mêmes distances que dans le monde réel, ce qui est rarement possible pour des distances supérieures à quelques mètres. Mais le cas de la marche sur tapis, les réponses motrices dans le déplacement virtuel sont proches de celles dans un déplacement virtuel, ce qui est intéressant, par exemple quand on souhaite que le sujet ait une fatigue similaire à celle dans un monde réel.

Pour trouver l'interface comportementale de déplacement la plus efficace, on doit bien analyser la PCV au niveau des I² fonctionnelles. Est-on obligé de faire concrètement marcher la personne ? Des déplacements commandés par le regard ou par des mouvements de main sont à considérer et à comparer aux interfaces spécifiques de locomotion. Différentes questions sont à se poser avant de faire le choix de la solution :

- le déplacement est-il à 1 degré de liberté (la trajectoire fixée), à 2 ddl (déplacement sur une surface sans rotation du corps), à 3ddl (déplacement sur une surface avec rotation du corps par rapport à la normale de la surface), à 4 ddl, à 5 ddl ou à 6 ddl ?
- le déplacement peut-il être guidé par différents types de contraintes géométriques ou cinématiques ? Peut-on réduire le nombre de ddl en répondant toujours aux objectifs des I² fonctionnelles ?
- la trajectoire est-elle importante ou est-ce seulement le lieu d'arrivée ?
- le déplacement doit-il se faire en continu, peut-il se faire pas à pas ou directement du début à la fin du déplacement ?
- le déplacement doit-il se faire à vitesse identique à celle supposée dans un monde réel simulé ?
- si le déplacement se fait pour plusieurs utilisateurs, y a-t-il un seul utilisateur qui commande le déplacement ?
- si le déplacement doit être similaire physiquement à celui du monde réel, s'agit-il de la marche, de la course, de monter des escaliers, etc. ?

13.3.2 LES DIFFÉRENTES INTERFACES COMPORTEMENTALES EXPLOITABLES

Sans essayer de faire une liste exhaustive, nous pouvons donner la plupart des solutions exploitées à ce jour pour l'interfaçage comportemental de déplacement. Nous les présentons en fonction de la partie du corps mis en œuvre et en mettant à part le cas des déplacements par l'intermédiaire de moyens de transport.

Au niveau des jambes, l'utilisateur peut se déplacer dans un monde virtuel :

- en ayant des traqueurs sur les jambes qui lui permettent d'indiquer les directions de déplacement souhaitées, la vitesse de déplacement étant fixe ou proportionnelle à l'écartement des pieds par rapport à une position neutre à la verticale du bassin du sujet ;
- en ayant des patins à roulettes instrumentés aux pieds, avec le corps suspendu à un portique [Iwata e.a. , 1994] ;
- en exploitant un tapis roulant 1D, avec un dispositif de changement de direction (guidon, par exemple) [Lourdeaux, 2001] ;
- en exploitant un tapis roulant 2D ;
- en utilisant une trottinette avec un tapis roulant instrumenté, mesurant la course virtuelle de la trottinette ;
- en utilisant des moyens de glisse : ski, luge, planche à voile ou surf.

Au niveau des mains, l'utilisateur peut se déplacer dans un monde virtuel en donnant des ordres binaires de direction (gauche, droite, avant, etc.) ou des ordres progressifs relatifs. Comme nous l'avons déjà mentionné, ces ordres servent soit à indiquer la direction souhaitée, soit le déplacement relatif du sujet par rapport à sa position actuelle, soit la «téléportation» de son corps virtuel. Ces commandes peuvent être données :

- par indication de la direction en pointant avec une main instrumentée ;
- avec une manette de jeux ;
- avec une chaise roulante ;
- en déplaçant relativement un chariot de magasin ;
- par métaphore en pointant la direction, le lieu d'arrivée sur un plan, etc. ;
- par un volant.

Au niveau de la tête, comme pour les mains, l'utilisateur peut se déplacer dans un monde virtuel en donnant des ordres binaires de direction ou des ordres progressifs relatifs :

- en pointant le regard dans la direction du mouvement, la vitesse étant fixe ou réglée par une commande manuelle. Cette solution est très utilisée avec l'emploi d'un visiocasque ;
- en déplaçant sa tête dans un plan parallèle à l'écran ;
- par commande vocale, mais celle-ci est astreignante pour l'utilisateur.

Au niveau de tout le corps, l'utilisateur peut se déplacer en exploitant :

- un delta-plane ;
- un parachute.

Le déplacement peut se faire par un moyen de transport :

- en utilisant un vélo d'appartement instrumenté ;
- en utilisant un véhicule de transport : voiture, train, bateau ou par avion.

Avec toutes ces solutions, il faut étudier les ALC que l'on doit programmer pour permettre l'exploitation d'une interface comportementale efficace et transparente. Suivant le type de dispositif, il est facile au lecteur de déterminer les ALC utiles, qui sont principalement des contraintes géométriques et cinématiques à programmer. Il reste ensuite à tester l'exploitation de l'interface, qui est perturbée par les incohérences sensorimotrices, toujours présentes, puisque tous les récepteurs sensoriels pour un déplacement ne sont pas stimulés ensemble.

13.4 PRIMITIVES COMPORTEMENTALES VIRTUELLES POUR AGIR SUR LE MONDE VIRTUEL

13.4.1 INTRODUCTION

Dans le cas d'action sur le monde virtuel, nous avons évidemment plusieurs sous catégories de PCV : manipuler un objet en translation (3ddl), orienter un objet en rotation (3ddl), les deux actions associées, déformer un objet, assembler des objets, etc. Nous pouvons établir la liste suivante, qui regroupe la plupart des sous PCV «agir» :

- manipuler un objet en translation ;
- orienter un objet en rotation ;
- manipuler et orienter un objet ;
- assembler des objets ;
- déformer un objet ;
- agir sur un être vivant.

Les sens et les réponses motrices principalement mis en action sont la vision, la sensibilité cutanée, les récepteurs proprioceptifs et les muscles des membres supérieures. La PCV «agir» ne peut donc être solutionnée sans prendre en compte les solutions de la PCV «observer» et la PCV «déplacer» éventuelle. Ceci complexifie la recherche de la solution optimale. Dans le monde réel, toute PCV «agir» est fonction de l'action motrice mais aussi de la vision par l'opérateur de l'espace correspondant. En conséquence, l'interface visuelle associée à l'interface motrice est à concevoir spécifiquement pour cette action et joue un rôle important pour la réussite de la PCV «agir» dans le monde virtuel. Si l'interface visuelle sert aussi à la PCV «observer», elle doit y être compatible, ce qui n'est pas forcément aisé à réaliser techniquement.

Que pouvons nous dire de général pour toutes les sous PCV «agir»? Une des principales questions pour l'interfaçage concerne l'exploitation ou non d'une simulation des retours d'effort et(ou) tactiles? Ceux-ci sont-ils absolument indispensables, en ayant bien conscience des difficultés à les réaliser? Peut-on simuler les retours d'effort ou tactiles par une substitution sensorielle, auditive ou visuelle? Si les retours d'effort et(ou) tactiles ne servent que d'information pour l'utilisateur de son action sur l'EV, il n'est pas obligatoire a priori de les simuler. Par exemple, comme nous l'avons indiqué pour la formation sur une infrastructure ferroviaire, le formé doit décrocher un téléphone. Le retour d'effort n'est pas utile car la formation n'a pas pour but d'apprendre à téléphoner. Par contre, si les retours d'effort ou tactiles sont sujets d'études (donc exigés par les I² mentales), il faut les simuler d'une façon réaliste, par exemple pour la formation d'actions physiques : dans une opération chirurgicale, le chirurgien doit ressentir les efforts sur ses mains, dans une formation à l'assemblage de mécanismes, etc. Par ailleurs, l'exploitation de métaphore virtuelle est-elle envisageable pour simplifier la réalisation, physiquement ou mentalement, de l'action par l'utilisateur?

13.4.2 MANIPULER ET ORIENTER UN OBJET

Manipuler un objet en translation ou en rotation pose déjà le problème pour l'utilisateur d'avoir une représentation spatiale adéquate. Les mouvements de translation étant commutatifs posent moins de problèmes que les mouvements de rotation autour de trois axes perpendiculaires qui ne sont pas commutatifs. Les interfaces matérielles exploitables sont principalement les manettes, les souris 3D, les traqueurs, les gants de données et les interfaces à retour d'effort. Nous avons déjà indiqué dans le chapitre des interfaces manuelles motrices que la préhension effective d'un l'objet peut être souvent simpliste : l'objet est pris ou l'objet est lâché. C'est le cas le plus élémentaire de la PCV «agir». Dans ce cas, un simple traqueur suffit pour indiquer les mouvements de l'objet et sa prise et son lâcher se font par une simple commande binaire. Cet ordre binaire peut être donné par un bouton ou par le simple rapprochement visuel à l'écran entre un curseur virtuel et l'objet à manipuler. Si la PCV «agir» est sujet d'étude, l'emploi d'un gant peut être utile et permet de plus se rapprocher du schème sensori-moteur de la préhension.

Pour la manipulation en orientation de l'objet, il y a spécifiquement deux cas à considérer : l'orientation se fait-elle dans toutes les directions (par exemple, pour l'observation d'une boîte sous toutes ses faces) ou est-elle limitée à celle permise par la rotation de l'objet avec une seule main (par exemple pour décrocher un combiné de téléphone)? Dans le premier cas, l'utilisation de deux gants de données est relativement onéreuse et peu pratique à cause des câbles électriques. Un traqueur, ayant une étendue de mesure dans toutes les directions (traqueur électromagnétique principalement) peut être exploité sans difficulté majeure pour l'utilisateur. Si de plus, le traqueur est au centre de masse d'un objet réel de forme identique à l'objet virtuel, l'orientation spatiale est plus aisée. Nous avons pu le vérifier pour la manipulation de boîtes parallélépipédiques avec un traqueur électromagnétique mis dans un cube en bois. L'inconvénient de ce type de manipulation est qu'elle n'est exploitable que si les objets ont des formes prédéfinis et similaires.

Nous avons indiqué dans le chapitre sur les traqueurs que ceux-ci peuvent servir :

- soit à manipuler un objet virtuel ;

- soit à diriger un curseur dans un environnement virtuel.

Dans le premier cas, l'utilisateur agit naturellement sur le monde virtuel, le traqueur se contentant de suivre le mouvement de sa main. Dans le deuxième cas, il s'agit en général de commander une action spécifique : par exemple, commander symboliquement l'ouverture d'une porte, etc. Dans ce cas, les traqueurs peuvent être remplacés par des interfaces de commande classiques : manettes ou souris 3D. Celles-ci, posées en général sur une table, permettent à l'utilisateur d'agir sur le monde virtuel, mais d'une façon relativement éloignée du schème sensori-moteur. La manipulation d'un objet devient délicate si elle se fait sur beaucoup de degrés de liberté, avec les problèmes de couplage translation-rotation et de la non-commutativité des rotations.

Les interfaces à retour d'effort sont techniquement toujours en même temps des capteurs de localisation. Elles sont comparables pour cette fonction à des traqueurs mécaniques (d'où problème si les rotations se font dans toutes les directions). Mais intrinsèquement, grâce aux stimuli proprioceptifs qu'elles procurent, elles permettent de plus se rapprocher du schème sensori-moteur de manipulation. Si dans la simulation virtuelle, l'action se fait par l'intermédiaire d'un outil travaillant dans un espace limité, la solution mécanique du retour d'effort est en général réalisable techniquement, comme nous l'avons déjà indiqué. Par contre, pour une manipulation sans intermédiaire d'outil et de plus sur un espace vaste et complexe, il existe rarement de solution mécanique fiable et efficace pour l'interface à retour d'effort. Dans ce cas, si les retours d'effort sont vraiment indispensables pour les I^2 fonctionnelles, on peut arriver à une impossibilité de concevoir un dispositif RV. C'est le problème ouvert pour la recherche actuelle en réalité virtuelle. Cette difficulté se rencontre particulièrement dans la sous PCV «agir» : assemblage d'objets, présentée plus loin.

Dans tous les cas précédents, si les *objets virtuels* sont à portée de *la main réelle*, les correspondances spatiales des gestes dans les deux environnements réel et virtuel sont assimilables par l'utilisateur. Au mouvement de sa main réelle, on associe visuellement le mouvement de la représentation virtuelle de sa main. La correspondance géométrique n'a pas besoin d'être parfaitement identique, pourvue qu'elle puisse être perçue correcte par l'utilisateur. A ce propos, il est souhaitable que les spécialistes des sciences cognitives nous indiquent quels sont les écarts maximums admissibles entre les mouvements réel et virtuel. Nous avons testé ce cas dans la prise de produits posés sur des étagères d'un magasin. Le devant des étagères et des produits étant au niveau de l'écran placé face à l'observateur, ce dernier tend vers l'écran sa main, instrumentée par un traqueur, mais sans toucher l'écran. Nous avons pu vérifier que ces faibles écarts entre les mouvements réel et virtuel ne posent pas problème d'interprétation à l'utilisateur.

Par contre si l'objet à manipuler est hors de portée de sa main réelle, il faut prévoir des ALC. On peut employer un rayon pour pointer à distance l'objet et le manipuler à distance [Bolt, 1980]. Si l'objet est loin et en conséquence difficile à pointer, le rayon peut être remplacé par un cône d'angle fixe [Liang e.a. , 1994]. Si l'utilisateur doit en plus amener l'objet vers soi, il peut envoyer à l'ordinateur un ordre pour réaliser le rapprochement [Bowmann e.a. , 1999]. D'autres ALC sont envisageables quand l'objet est hors de portée de la main réelle [Coquillart e.a. , 1999].

13.4.3 ASSEMBLER DES OBJETS

L'assemblage de deux objets pose des problèmes mécaniques complexes. S'il s'agit de vérifier que deux objets peuvent s'assembler, des actions sans retour d'effort peuvent être envisagées si on remplace les efforts mécaniques entre les deux objets par des ALC adéquates : par programmation, il faut contraindre l'objet manipulé à avoir un mouvement par rapport à l'autre similaire à ce qu'il se produirait dans un monde réel. Il faut donc prévoir des «attracteurs» entre les objets et bien les ajuster. Par exemple, s'il s'agit de poser une pièce à fond plat sur une table, dès que celle-ci est proche et parallèle à la table, les deux objets peuvent être liés par 3 ddl. Les zones d'attraction doivent être réglées pour un assemblage réaliste, ce qui n'est pas toujours simple. Si on étudie le cas d'un assemblage d'un tiroir dans un meuble, les ALC (contraintes cinématiques) ne sont pas forcément réalistes pour étudier l'assemblage. Dans ce cas, une interface avec retour d'effort devient indispensable, bien qu'elle soit très difficile à réaliser. Si, de plus, au cours de l'assemblage, on souhaite étudier les contraintes géométriques imposées par les mains de l'opérateur, l'interfaçage est encore plus difficile à réaliser. La transmission des efforts sur toutes les parties des mains (doigts et paumes) est très complexe à concevoir, sauf si on modifie la partie simulée : on peut avoir l'objet réel et non virtuel (par exemple une copie par stéréographie). Le retour d'effort, fourni par l'interface, est alors appliqué sur la copie réelle de l'objet. L'opérateur tient en main cette copie, ce qui lui permet de ressentir sur ses mains des efforts réalistes. Mais ceci au détriment d'une potentialité de la réalité virtuelle : la forme de l'objet ne peut pas être modifiée numériquement, donc rapidement et sans coût complémentaire.

13.4.4 DÉFORMER UN OBJET

Dans cette situation de déformation de la surface d'un objet, il est plus délicat de se passer de retour d'effort. Le dispositif de base est donc une interface à retour d'effort, avec tous les problèmes que cela pose comme précédemment. De plus, les efforts sont plus difficiles à calculer en temps réel car il faut faire intervenir les lois de résistance des matériaux. Une substitution sonore peut remplacer la perception des retours d'effort, mais cette métaphore donne des indications approximatives sur les déformations qui doivent être visualisés en complément s'ils sont assez importants pour être perçus visuellement. S'ils ne sont pas assez grands, le concepteur doit se poser la question de savoir si l'exagération des déformations est incompatible ou non avec les I² fonctionnelles pour résoudre le problème.

13.4.5 AGIR SUR UN ÊTRE VIVANT

Une interaction avec un être vivant (clone, avatar) est délicate à réaliser si elle doit être réaliste. Le comportement de l'être virtuel doit être suffisamment naturel. Par exemple, agir sur un mannequin virtuel pour des études ergonomiques n'est pas simple à réaliser tout en imposant un comportement humain du mannequin (voir le chapitre sur les *comportements des objets du monde virtuels* pour les difficultés de modélisation). Des ALC peuvent être prévues mais en général elles ne peuvent pas être seulement purement géométriques et mécaniques, d'où les problèmes complexes qu'elles génèrent pour l'informaticien.

13.5 PRIMITIVES COMPORTEMENTALES VIRTUELLES DE COMMUNICATION AVEC AUTRUI

Dans le cas de la communication, nous avons plusieurs sous catégories de PCV : il est envisageable de communiquer avec d'autres utilisateurs, directement ou par l'intermédiaire de clones les représentant. Il est possible de dialoguer avec des personnages virtuels (avatars). Une dernière sous PCV est de permettre à l'utilisateur de communiquer avec le système informatique pour configurer autrement l'application : cela est utile si c'est l'utilisateur final qui choisit lui-même la configuration. Par exemple dans une revue de projet d'un nouveau produit, l'ingénieur, qui teste un produit en s'immergeant, doit pouvoir en changer les caractéristiques pour observer plusieurs variantes de son produit. Mais cette sous PCV est hors du champ de la finalité de la réalité virtuelle : le sujet n'est plus vraiment en immersion pseudo naturelle puisque mentalement il sort de l'environnement virtuel pour configurer celui-ci.

La communication peut se faire sur le site même ou à distance par l'intermédiaire de réseau. La communication se fait en général par la parole mais elle est souvent aussi importante, même si elle n'est pas explicite, par l'observation des gestes et du visage de l'autre utilisateur. Ceci pose un vrai problème lorsque la communication se fait à l'aide d'un avatar, dont les expressions gestuelles et faciales sont difficiles à simuler et à maîtriser en temps réel.

Sans développer dans ce chapitre, tous les cas de communication possibles en réalité virtuelle, nous précisons les solutions du cas particulier du dialogue avec le système informatique. Ce cas peut être considéré partiellement comme le complémentaire en 3D de la communication IHM en 2D (Interface Home Machine). Comment donner des informations au système informatique, à l'application et comment celui-ci peut-il répondre à l'utilisateur ?

Le mode oral est possible malgré la fatigue qu'il peut engendrer si les commandes sont fréquentes. Si les commandes sont complexes et multiples, la mémoire et l'attention de l'opérateur peuvent être très sollicitées. Des menus (déroulants) peuvent être une aide efficace pour compléter l'interfaçage. L'interprétation d'ordre complexes peuvent être techniquement difficiles à interpréter.

Il y a la possibilité d'agir manuellement pour donner des informations au système informatique ou à l'application. D'une façon similaire aux menus 2D, il y a souvent deux opérations à faire : sélectionner et commander une fonction d'un sous-menu. Pour sélectionner, nous pouvons exploiter les solutions similaires que celles exposées pour la sous PCV de manipulation dans le cas d'objets hors de portée de la main. Le menu, permettant la communication avec le système, peut être représenté en 2D dans l'environnement virtuel 3D. Dans ce cas, l'interfaçage comportemental n'est pas forcément adapté au 2D. La sélection n'est pas toujours aisée. Il faut prévoir des ALC pour faciliter les actions de l'opérateur. Le menu 2D est soit fixe par rapport au monde virtuel 3D, soit fixe par rapport au point de vue de l'observateur, par exemple avec un visio-casque. On peut aussi prévoir la sélection sur une surface physique comme support du menu 2D. Le support est fixe dans le monde réel (par exemple, une petite zone à un coin de l'écran du bureau immersif), soit tenu en main comme une palette de peintre [Coquillart e.a. , 1999]. Un menu 3D est envisageable, mais il n'apporte pas intrinsèquement d'amélioration, presque tous les utilisateurs étant habitués aux menus déroulant 2D.

13.6 CONCLUSION

Ce chapitre n'expose qu'une première analyse partielle des solutions techniquement réalisables pour les Primitives Comportementales Virtuelles. Notre contribution est incomplète pour deux raisons : nous n'avons pas fait un état de l'art exhaustif des solutions proposées dans toutes les applications RV actuelles et il y a encore bien des solutions qui n'ont pas été réalisées, testées et validées sérieusement par la communauté des chercheurs. Cela reste un vaste problème ouvert dans le champ de la réalité virtuelle.

13.7 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Berthoz, 1998] A. Berthoz. *Le sens du mouvement*. Odile Jacob, Paris (1998). ISBN 2-7381-0457-6.
- [Bolt, 1980] R. A. Bolt. Put-that-there : Voice and gesture at the graphics interface. *Computer Graphics*, 14(3) :262–270 (1980).
- [Bowmann e.a. , 1999] D. Bowmann and L. Hodges. An evaluation of techniques for grabbing and manipulating remote objets in immersive virtual environments. *The Journal of Visual Languages and Computing*, 10(1) :37–53 (1999).
- [Burkhardt e.a., 1999] J.-M. Burkhardt, D. Lourdeaux, and P. Fuchs. Conception d'un système de rv pour la formation des agents de conduite aux opérations en milieu ferroviaire. In *Actes des journées Réalité Virtuelle et Cognition* (1999), pages 123–132, Paris.
- [Coquillart e.a. , 1999] S. Coquillart and G. Wesche. The virtual palette and the virtual remote control, a device and an interaction paradigm for the responsive workbench. In *Proceedings of IEEE Virtual Reality'99* (1999), pages 213–216.
- [Iwata e.a. , 1994] H. Iwata and K. Matsuda (1994). Haptic walkthrough simulator. Technical Report 305, Institute of Engineering Mechanics, Université de Tsukuba, Japon.
- [Liang e.a. , 1994] J. Liang and M. Green. Geometric modeling using six degrees of freedom input devices. *Computers and Graphics*, 18(4) (1994).
- [Lourdeaux, 2001] D. Lourdeaux (2001). *Réalité Virtuelle et Formation : Conception d'Environnements Virtuels pédagogiques*. Thèse de doctorat, Ecole des Mines de Paris, Paris.

Sixième partie

Modélisation des mondes virtuels

14 MODÈLES GÉOMÉTRIQUES DES ENVIRONNEMENTS VIRTUELS

Guillaume Moreau

14.1 INTRODUCTION

Dans un environnement virtuel, l'observateur évolue dans un monde qui est une représentation d'une partie de la réalité ou d'un environnement imaginaire [Fuchs, 1996]. Dans la plupart des cas, la représentation qui est faite à l'observateur est en premier lieu visuelle. Cette représentation visuelle est formée par la projection d'un modèle géométrique 3D de l'environnement virtuel sur une ou plusieurs images planes 2D par l'intermédiaire de caméras virtuelles dont le mouvement peut être lié à celui de l'observateur.

Cet environnement n'est bien sûr pas uniquement constitué d'éléments de nature géométrique ; prenons l'exemple de ce qu'on appelle la *ville virtuelle*. Une ville virtuelle se compose au moins d'une représentation du bâti et du réseau routier ainsi que d'acteurs évoluant dans cette ville (cf. figure 14.1 représentant une simulation de tramway dans un carrefour nantais [Donikian e.a., 1998]). Les représentations géométriques du bâti et du réseau routier sont insuffisantes pour permettre cette simulation [Donikian, 1997]. La seule intégration d'un tramway dans le cadre de cette étude a nécessité la création d'un modèle géométrique et cinématique d'un tramway, mais aussi la définition du comportement du couple conducteur-tramway au passage de certains éléments remarquables du réseau routier (feux, intersections, arrêts de tramway). Dès lors, la modélisation du réseau routier devient non seulement géométrique, mais aussi topologique (orientation dans un réseau) et sémantique (signification du feu rouge) [Thomas, 1999].



Figure 14.1 : Etude d'impact sur l'intégration d'un tramway dans un carrefour nantais (image IRISA/Siames).

Si nous avons rapidement montré que la modélisation d'un environnement virtuel n'est pas uniquement géométrique, celle-ci constitue néanmoins une étape importante du processus de construction d'un environnement virtuel. Ce chapitre et le chapitre suivant

y seront consacrés. Dans un premier temps, nous ferons un tour des modèles théoriques associés à la description géométrique, puis nous aborderons les outils de la modélisation. Les aspects non géométriques seront abordés ensuite au chapitre 15.

14.1.1 TYPES D'OBJETS

Avant de modéliser des objets géométriques, il faut pouvoir répondre à la question : que cherche-t-on à modéliser ? il est évident que la modélisation d'un solide indéformable est différente de celle d'un fluide compressible. L'objet le plus simple est bien sûr le solide rigide non déformable. Cet objet ne peut subir que des transformations solides $T = (\mathbf{R}, \mathbf{t})$, où $\mathbf{R}$ est une rotation dans l'espace et $\mathbf{t}$ une translation. Le cube de la figure 14.2 est un exemple de solide rigide indéformable. Le principal avantage des solides rigides est de n'avoir à calculer que leur position solide dans l'espace avant d'effectuer leur rendu. Les modèles non-rigides peuvent se décomposer quant à eux en plusieurs catégories :

- les solides articulés : ce sont des solides rigides indéformables reliés entre eux par des liaisons mécaniques (par exemple une liaison pivot). Ils sont utilisés pour représenter par exemple le mouvement d'un bras relié à un tronc ou les roues d'une voiture qui sont liées au châssis, en première approximation bien sûr. C'est aussi le cas du pendule de la figure 14.3. Les liaisons entre les solides sont représentées par leurs paramètres (ici les angles θ_1 et θ_2) ;
- les solides déformables dont les déformations suivent les lois de la mécanique (ou pas, dans un monde imaginaire). Ce peut être la peau d'un être vivant, une balle se déformant au contact du sol. Ces modèles sont régis par leurs paramètres mécaniques. Il est donc nécessaire de calculer l'évolution de leur forme avant d'effectuer leur rendu visuel. Les vignettes de la figure 14.4 présentent l'évolution de la lampe Luxo se déformant sous l'effet de son propre poids [Cozot, 1996a] ;
- en dernier les modèles non-solides, qui représentent tout ce qu'on ne peut pas associer à la notion d'objet. On peut citer les fluides, les modèles particuliers qui associent plusieurs sous-modèles associés mécaniquement entre eux. La création d'un jet de lave par exemple utilise des modèles particuliers.

Il existe par ailleurs des modèles qui ne rentrent pas en compte dans le cadre de la réalité virtuelle à utilisation professionnelle ; ils sont associés aux techniques dites de *rendu non photo-réaliste*. Ces techniques ont des résultats souvent spectaculaires mais elles ne peuvent être classées en familles simples comme celles que nous venons de citer. De plus leur utilisation en réalité virtuelle semble marginale. Elles ne seront donc pas abordées dans cet ouvrage.

La contrainte majeure de la réalité virtuelle en terme de rendu graphique porte naturellement sur la nécessité absolue d'un rendu temps réel du modèle 3D. Il est donc évident qu'on cherche à privilégier dans les environnements virtuels les objets les plus simples (solides rigides et articulés) au détriment des plus complexes (fluides visqueux par exemple). On se limite donc essentiellement dans les paragraphes qui suivent à des solides rigides représentant des objets manufacturés ou des éléments naturels.

14.1.2 PROPRIÉTÉS DES MODÈLES

Selon Michael Mortenson [Mortenson, 1985], «un modèle géométrique est une représentation mathématique non ambiguë et complète de la forme d'un objet physique afin

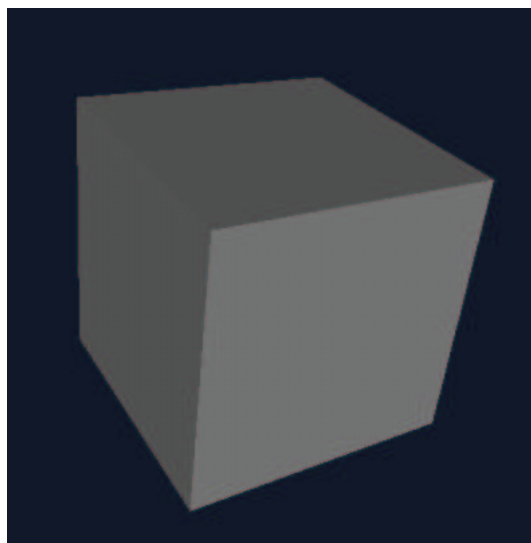


Figure 14.2 : Solide rigide : un cube.

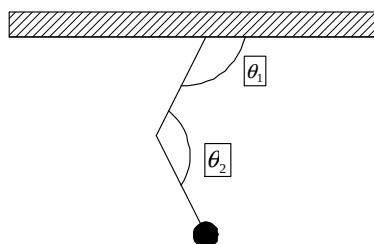


Figure 14.3 : Solide articulé : un bi-pendule.

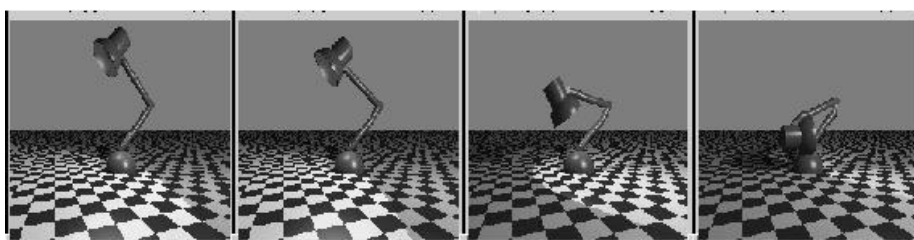


Figure 14.4 : Objet déformable : la lampe Luxo se déformant sous son propre poids (image IRISA/Siames).

de la traiter sur ordinateur». Comme il existe une variété de modèles, il est utile de pouvoir comparer les différentes représentations d'un même objet entre elles. Les critères utilisés sont les suivants :

- *le domaine* : il s'agit de l'ensemble des objets qu'une représentation peut modéliser ;
- *la complétude* : c'est la capacité théorique d'une représentation à pouvoir répondre à certaines questions géométriques simples comme l'aire d'une surface, la position d'un point par rapport à un volume... ;
- *l'unicité de la représentation d'un objet* : ce critère est important lorsqu'on cherche à déterminer l'égalité de deux objets, à un couple rotation/translation près ;
- *la facilité de manipulation* : c'est la base de la création et de la modification du modèle ;
- *les performances techniques* : la précision, la concision de la structure, mais aussi la rapidité des algorithmes associés (détection de collision, rendu visuel...).

Il existe deux techniques fondamentales pour représenter des objets en trois dimensions :

- la représentation volumique : elle consiste naturellement à déterminer une représentation informatique de leur volume ;
- la représentation surfacique : elle provient du constat que nous nous situons dans un espace euclidien à trois dimensions où tout objet fini régulier a une frontière non dégénérée. On peut donc déterminer le volume de l'objet comme étant l'intérieur de sa frontière. En résumé, on cherche donc ici à modéliser la sphère plutôt que la boule. Si cela peut paraître plus complexe en théorie, cette représentation est en réalité plus utile, puisque dans le cas de solides non transparents, le rendu n'est effectué pour la seule surface de l'objet.

Pour les représentations volumiques comme surfaciques, il existe plusieurs modèles mathématiques vérifiant chacun des propriétés différentes. Nous verrons dans les sections suivantes les principaux types de modèles tant volumiques que surfaciques puis l'algorithmique associée aux conversions de modèles. Enfin nous verrons les améliorations que nous pouvons apporter à une modélisation pour l'utiliser dans un environnement virtuel, avec les contraintes de temps réel qui s'y rattachent.

14.2 MODÈLES VOLUMIQUES

Les modèles dits volumiques sont les modèles qui représentent le volume d'un objet. Deux familles de techniques sont utilisées :

- *l'énumération spatiale* : elle consiste à découper l'espace en un nombre important de petites cellules qui appartiendront totalement, en partie, ou pas du tout au volume de l'objet. La modélisation consiste alors à trouver un découpage de l'espace et une fonction qui détermine le *remplissage* d'une cellule lorsque seule une partie de celle-ci est occupée par l'objet à modéliser ;
- *la CSG* : La CSG (Constructive Solid Geometry) est un arbre binaire de construction où les feuilles sont des primitives géométriques et les nœuds des opérateurs booléens.

14.2.1 ÉNUMÉRATION SPATIALE

On cherche donc à construire une partition de l'espace. On utilise des axes de découpage orthogonaux aux axes du repère global pour simplifier. Chaque cellule est ainsi

localisée par rapport à l'objet à représenter : si elle est pleine, on dit qu'elle est à l'intérieur de l'objet, si elle est vide on dit qu'elle est à l'extérieur de l'objet. Si elle n'est que partiellement remplie, on fournira une fonction de test d'appartenance. La figure 14.5 présente un maillage grossier d'un lapin en deux dimensions. La seconde vignette prouve qu'il est impossible de reconnaître visuellement un lapin à partir d'un maillage si large. Le problème de la précision suffisante du maillage est donc posé. La notion de maillage régulier se généralise aisément en trois dimensions.

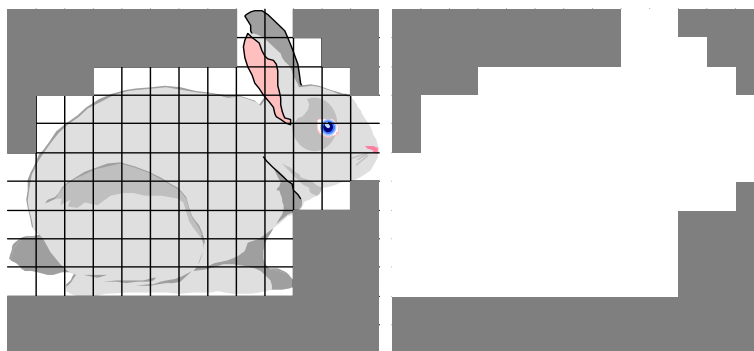


Figure 14.5 : Maillage régulier en deux dimensions d'un lapin.

Le choix du maillage revêt une grande importance puisqu'il est très directement lié à la concision de la représentation. Lorsqu'on travaille dans un espace à 3 dimensions avec un maillage uniforme, le nombre de cellules croît avec le cube du niveau de subdivision. Pour un volume formé dans un cube de 10 cm de côté découpé avec une précision de 1 mm, on obtient déjà 1 million de cellules élémentaires. Cette technique est très utilisée en imagerie médicale où les images scanner ont une résolution de 256 pixels par 256 (voire 512 par 512 maintenant). Chaque coupe scanner utilise à elle seule plus de 65000 cellules, nombre qu'il faut encore multiplier par le nombre de couches. Pire encore, quelque soit le niveau de subdivision de l'espace, on ne pourra jamais représenter de façon exacte une simple droite oblique (et à plus forte raison un cercle ou une sphère).

On pense alors logiquement à utiliser un maillage adaptatif, c'est-à-dire grossier sur les grands volumes de l'intérieur de l'objet et plus fin vers ses frontières. L'utilisation d'un maillage adaptatif n'augmentera pas le domaine des objets modélisables de façon exacte, mais en permettra une représentation plus compacte. On construit donc une représentation arborescente d'un objet selon un algorithme descendant : si une cellule est uniforme, une feuille terminale est créée ; dans le cas inverse, 2^d (d étant la dimension de l'espace considéré, en 2D on parle de *quadrees*, en 3D d'*octrees*) cellules filles partitionnant régulièrement la cellule mère sont créées et l'algorithme est appliqué à chacune d'entre elles [Jackins e.a. , 1980]. Bien entendu on fixe une limite maximale de subdivision de façon à ce que l'algorithme puisse se terminer. La figure 14.6 montre un exemple de calcul d'un quadtree à partir d'une image (qui est elle-même par définition un maillage régulier de l'espace en deux dimensions). Le passage dans un espace de dimension 3 est tout aussi évident.

Il existe aussi des techniques de subdivision plus élaborées puisqu'elles ne découpent pas les cellules suivant les axes, mais suivant des schémas simples ; elles sont utiles pour découper de façon exacte des polyèdres. On parle alors de *polytree* ou de *PM tree*

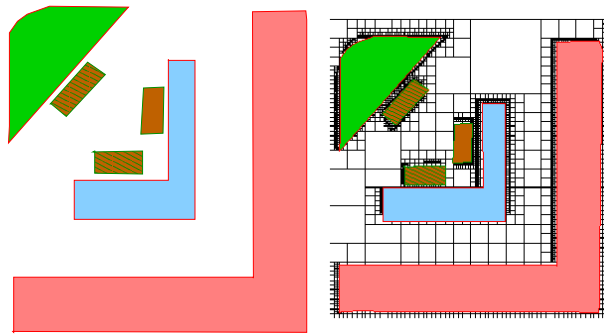


Figure 14.6 : Construction d'un quadtree à partir d'un maillage régulier d'un espace à deux dimensions (dans cet exemple, on a modélisé l'espace libre, c'est-à-dire le complémentaire des objets).

(PM signifiant alors Polygonal Map).

La représentation informatique d'une représentation par énumération d'un objet est simple : une grille régulière peut-être représentée par un tableau à trois dimensions de booléens. Les structures d'octree (généralisation du quadtree dans un espace à trois dimensions) sont des arbres octaux que l'on construit récursivement.

L'utilisation de structures d'énumération spatiale (régulières ou adaptatives) est très simple d'un point de vue algorithmique, les complexités sont souvent assez faibles, même les opérateurs géométriques booléens sont faciles à mettre en œuvre. Les problèmes se posent lors des changements d'échelle (dont le facteur n'est pas un multiple de 2) ou lors des rotations (d'angle différent de 90 degrés).

Du point de vue des propriétés énoncées en 14.1.2, on peut dire que :

- le domaine des objets représentés est très large si on ne fixe pas une marge d'erreur strictement nulle. En revanche, le domaine de représentation exacte est particulièrement limité. Toutefois, la précision de modélisation est connue à l'avance ;
- le modèle est bien adapté aux algorithmes simples ;
- l'unicité des modèles n'est pas vérifiée à cause de la sensibilité des modèles à l'orientation de l'objet ;
- dans l'ensemble, les modèles sont faciles à manipuler, même si les rotations et les mises à l'échelle peuvent poser de gros problèmes ;
- les structures de données sont particulièrement lourdes puisque la précision est une racine cubique du nombre de cellules utilisées. Cette complexité peut être diminuée en utilisant des structures arborescentes comme les *octrees*.

Comme on l'a évoqué précédemment, les grilles sont très utilisées dans le domaine médical (IRM, scanner), mais aussi en géophysique (l'Institut Français du Pétrole par exemple pour représenter des cubes de données sismiques). La figure 14.7 montre une reconstruction du sous-sol de la région de Morges dans les Alpes à l'aide d'un modèle numérique de terrain et d'un certain nombre de coupes. On utilise donc essentiellement les grilles pour représenter le réel (données acquises) plutôt que pour des activités de création. Les structures adaptatives ne sont généralement pas utilisées en tant que modèle mais en tant qu'intermédiaires de calcul, notamment pour les calculs d'intersec-

tion. Le rendu visuel des modèles volumiques passe soit par une conversion en modèle surfacique soit par des techniques particulières comme le logiciel OpenGL Volumizer [http1].

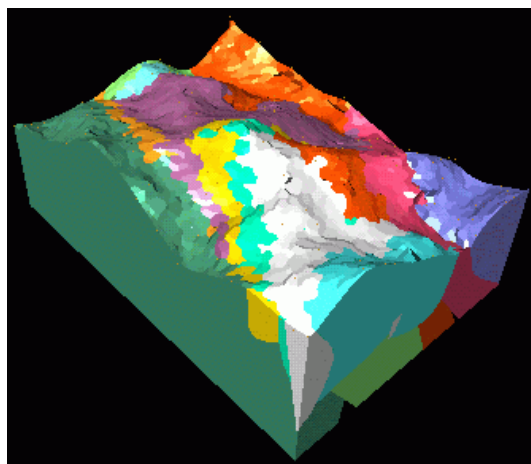


Figure 14.7 : Reconstruction du sous sol de la région de Morges (image INRIA/Prisme).

14.2.2 CONSTRUCTIVE SOLID GEOMETRY

La CSG est une technique de modélisation particulière puisqu'elle ne préjuge pas de la représentation réelle. On peut utiliser la CSG en commun avec des primitives surfaciques, même si la surface résultant de l'intersection d'une sphère avec un cube sera difficile à calculer. On s'intéresse ici à la CSG en tant que représentation d'un objet en trois dimensions. On considère donc un ensemble E de primitives géométriques, constitué par exemple de boules, de pavés, de tores... On peut aussi inclure dans E des volumes obtenus par extrusion ou par balayage, cela n'a aucune incidence sur le modèle théorique. Le domaine des volumes constructibles est donc l'ensemble $\mathcal{D}$ engendré par $E \cup \{0\}$ et les opérateurs ensemblistes $\cap$, $\cup$ et $-$.

Les primitives CSG sont très utilisées dans le domaine de la CAO pour créer des formes de pièces. Si la plupart des modeleurs 3D du commerce supportent la CSG, le logiciel **povray** [http2] du domaine public est des rares logiciels de rendu à l'exploiter de façon native. On peut très facilement créer des modèles d'objets CSG à l'aide du langage de script. L'exemple suivant montre la création d'un pavé percé par trois trous dont l'un est rempli par un cylindre ¹. L'arbre CSG associé au fichier est représenté figure 14.8.

¹Dans le langage de povray, lorsque deux primitives géométriques se suivent, elles sont implicitement reliées par l'opérateur union.

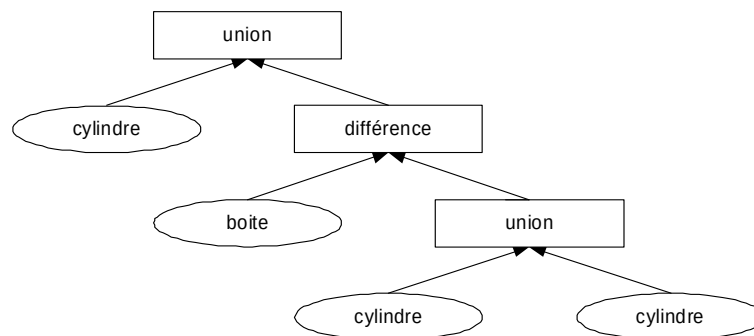


Figure 14.8 : Exemple d'arbre CSG.

```

cylinder {
  <0,-4,0>
  <0,4,0>
  0.2
  pigment { Yellow }
}
difference {
  box {
    <-1.5,-0.5,-0.5>
    <1.5,0.5,0.>
  }
  cylinder {
    <-0.8,0,-3>
    <-0.8,0,3>
    0.1
  }
  cylinder {
    <0.8,0,-3>
    <0.8,0,3>
    0.1
  }
  pigment { Red }
}
  
```

En y ajoutant une caméra et une source de lumière, on obtient l'image de la figure 14.9.

En termes plus informatiques, la représentation d'un arbre CSG est une simple structure d'arbre où les feuilles sont des primitives géométriques et les nœuds des opérateurs booléens. Certains algorithmes de traitement sont extrêmement simples à mettre en œuvre, d'autres ne sont pas prévus par la structure (par exemple il est impossible d'effectuer un calcul de volume dans le cas général). Le problème fondamental inhérent à la représentation CSG est celui de la représentation visuelle des primitives élémentaires et la détermination du résultat de l'application des opérateurs sur ces structures.

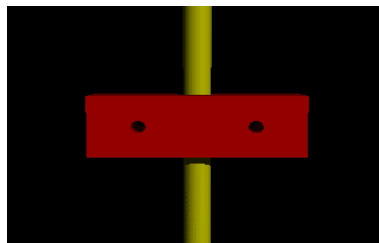


Figure 14.9 : Exemple de pièce modélisée par CSG avec povray

En termes de propriétés, nous pouvons dire que :

- l'ensemble des objets représentés dépend essentiellement des primitives de base utilisées, mais peut être très étendu ;
- le modèle est très adapté à certains calculs tout en rendant des opérations, comme un simple calcul de volume, complètement impossibles ;
- l'unicité n'est évidemment pas vérifiée, l'ordre de construction influe sur les primitives utilisées et sur l'ordre d'application des opérateurs ;
- les modèles CSG sont très faciles à utiliser et sont donc généralement implantés dans tous les outils de modélisation 3D, même si la structure interne de ces derniers est plutôt de nature surfacique ;
- la précision et la concision des modèles dépendent essentiellement des primitives utilisées et de leur adéquation à l'objet à modéliser.

Les modèles CSG sont naturellement la clé de voûte des outils de CAO mécanique, à la fois pour des questions d'ergonomie et de fabrication des objets conçus. Toutefois, leur adéquation à la réalité virtuelle est plus que discutable comme nous le verrons par la suite. En matière de rendu visuel, la CSG est plutôt utilisée dans des logiciels non temps réel comme les logiciels de lancer de rayon. En effet les techniques de lancer de rayon peuvent s'affranchir du calcul explicite du volume de l'objet modélisé.

14.3 MODÈLES SURFACIQUES

Nous arrivons maintenant à la représentation surfacique où nous allons chercher à modéliser les frontières de l'objet. Pour cette raison, la modélisation surfacique est aussi appelée Brep (Boundary Representation). Le problème est différent en ce sens que nous déterminons cette fois une surface, c'est-à-dire un hyperplan de l'espace à trois dimensions. Il faut alors chercher à représenter la surface frontière sous une forme qui la rende utilisable pour des traitements algorithmiques. Les modèles surfaciques sont classés suivant le type de surface utilisé.

14.3.1 UTILISATION DE SURFACES PLANES

C'est le modèle le plus généralement utilisé pour le rendu temps réel. On représente la surface frontière d'un objet à l'aide d'un ensemble de faces polygonales connectées entre elles. Une frontière se représente donc à l'aide de trois ensembles : un ensemble de sommets S , un ensemble d'arêtes A et un ensemble de faces F . Le modèle doit respecter certaines contraintes d'intégrité :

- les faces ne s'intersectent qu'en leurs sommets ou arêtes communes ;
- une arête appartient à deux faces ;
- tous les sommets adjacents à un sommet donné forment dans l'espace un polygone simple (pas d'intersection entre les arêtes, si possible convexe).

On utilise généralement des triangles comme polygones, pour simplifier au maximum les calculs, le triangle étant la surface polygonale fermée la plus simple. Elle est de plus forcément planaire. Tout polygone pouvant être réduit à un ensemble de triangles, les problèmes sont donc équivalents. La figure 14.10 présente un exemple de pièce modélisée à l'aide d'un maillage triangulaire. La précision du maillage n'est limitée que par le nombre de triangles qu'on peut afficher. Aujourd'hui, les cartes graphiques 3D sont capables d'afficher un grand nombre de triangles par seconde, mais en contrepartie elles ne peuvent afficher que des triangles !

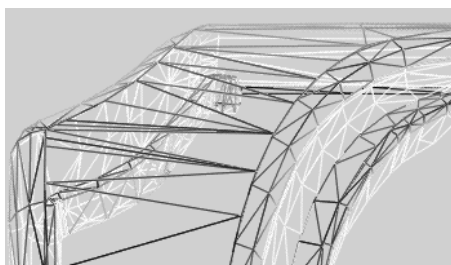


Figure 14.10 : Exemple de pièce modélisée à l'aide d'un maillage triangulaire.

14.3.2 UTILISATION DE SURFACES NON PLANES

Le modèle proposé au paragraphe précédent se généralise sans problème à des surfaces non-planes. Les contraintes d'intégrité sont plus difficiles à faire respecter. Les surfaces non planes présentent généralement l'avantage de la compacité de la représentation. Elles sont plus faciles à manipuler pour des opérateurs de CAO puisqu'elles nécessitent moins de paramétrages. Leur aspect visuel est plus agréable à quantité de données égale, elles permettent de représenter plus facilement et plus précisément des surfaces *lisses*. En règle générale, les modelleurs utilisent des modélisations à base de surfaces gauches, qu'on cherche ensuite à trianguler afin de simplifier les calculs de rendu. La triangulation, qui est un problème complexe abordé dans la section 14.4, est nécessairement un compromis entre la précision du modèle triangulé, le temps nécessaire au rendu et le temps consacré à la triangulation elle-même.

Les modelleurs actuels font la part belle aux surfaces NURBS pour la modélisation d'objets. C'est le cas de CATIA pour la CAO, mais aussi de 3DS Max pour l'infographie. Les formats de fichiers dédiés à la réalité virtuelle (VRML par exemple) permettent d'ailleurs l'incorporation directe de surfaces NURBS, même si elles seront triangulées pour la visualisation temps réel.

14.3.2.1 Les surfaces NURBS

La plupart des attributs des surfaces NURBS sont décrits dans la signification des initiales :

- *NU* : les surfaces NURBS sont non-uniformes ;
- *R* : les équations qui les décrivent sont rationnelles ;
- *BS* : ce sont des surfaces splines de Bézier.

En résumé, les NURBS sont des surfaces paramétriques rationnelles construites à partir de deux paramètres u et v , d'un ensemble de points de contrôle, d'une séquence de nœuds et du poids affecté à chacun des points de contrôle. Elles permettent de représenter avec beaucoup de précision les surfaces utilisées en CAO, ainsi que de prendre en compte les différents types de continuité entre surfaces d'un même objet. Les surfaces NURBS sont des généralisations des B-splines non-rationnelles qui proviennent elles-mêmes des courbes de Bézier rationnelles. On peut aller encore plus loin en ajoutant que ces dernières sont une généralisation des courbes de Bézier...

Reprenons depuis le début : une courbe de Bézier de degré n est une courbe paramétrique $C(u)$ définie par :

$$C(u) = \sum_{i=0}^n B_{i,n}(u) \mathbf{P}_i \quad 0 \leq u \leq 1 \quad (14.1)$$

Les $\mathbf{P}_i$ sont appelés les points de contrôle de la courbe. Les fonctions de base $B_{i,n}$ sont les classiques polynômes de Bernstein de degré n définis par :

$$B_{i,n}(u) = \frac{n!}{i!(n-i)!} u^i (1-u)^{n-i} \quad (14.2)$$

Les courbes de Bézier sont très utilisées pour le dessin interactif en deux dimensions, on en trouve dans toutes les applications Windows par exemple. Typiquement, on utilise une courbe de degré 3 (qui possède par conséquent 4 points de contrôle), comme celle de la figure 14.11. Malheureusement, les courbes de Bézier ne permettent pas de représenter les coniques. Une conique (par exemple un cercle) peut être représentée à l'aide d'une fonction rationnelle (le quotient de deux polynômes), soit en dimension 3 :

$$x(u) = \frac{X(u)}{W(u)} \quad y(u) = \frac{Y(u)}{W(u)} \quad z(u) = \frac{Z(u)}{W(u)} \quad (14.3)$$

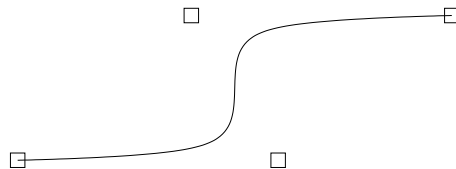


Figure 14.11 : Courbe de Bézier de degré 3, avec ses 4 points de contrôle.

A partir de cette observation, on définit les courbes de Bézier rationnelles :

$$\mathbf{C}(u) = \frac{\sum_{i=0}^n B_{i,n}(u) w_i \mathbf{P}_i}{\sum_{i=0}^n B_{i,n}(u) w_i} \quad 0 \leq u \leq 1 \quad (14.4)$$

où les $\mathbf{P}_i$ (respectivement les B_i) sont toujours les points de contrôle (respectivement les polynômes de Bernstein). Les w_i sont des scalaires qu'on appelle les poids des points de contrôle. Leur rôle consiste à *attirer* ou au contraire *repousser* la courbe des points de contrôle.

Généralement, l'utilisation d'un nombre réduit de points de contrôle ne permet pas de dessiner facilement la forme adéquate. Si on augmente le nombre de points de contrôle, on augmente par conséquent le degré des polynômes, ce qui les rend moins intéressants numériquement (instabilité numérique, coûts de calcul), mais aussi du point de vue de l'utilisation interactive (un point de contrôle a une influence sur toute la courbe ce qui limite les possibilités de modifications très localisées). On va donc créer des courbes de Bézier jointes bout à bout, en y ajoutant des contraintes de continuité. Pour ceci, on va utiliser certains points de contrôle plusieurs fois.

La courbe $\mathbf{C}(u)$ est alors définie sur $u \in [0, 1]$ et est composée des segments $\mathbf{C}_i(u)$, $1 \leq i \leq m$. Ces segments sont joints à leurs extrémités $u_0 = 0 < u_1 < u_2 < u_3 < u_4 = 1$ avec un certain niveau de continuité. On dit que la courbe est C^k continue à une extrémité u_i si $\mathbf{C}_i^{(j)}(u_i) = \mathbf{C}_{i+1}^{(j)}(u_i)$ pour tout $0 \leq j \leq k$, où $\mathbf{C}_i^{(j)}$ représente la j -ème dérivée de $\mathbf{C}_i$. Pour rendre les équations compatibles avec une implantation informatique, on utilise la définition suivante :

$$\mathbf{C}(u) = \sum_{i=0}^n N_{i,p} \mathbf{P}_i \quad a \leq u \leq b \quad (14.5)$$

où les $\mathbf{P}_i$ sont les points de contrôle et $N_{i,p}$ les fonctions B-splines de base de degré p . Il y a plusieurs approches pour définir ces fonctions B-splines de base. Nous nous contenterons de l'approche récurrente, bien adaptée à une implantation informatique. La B-spline a des points de contrôle de jonction entre ses segments qu'on appelle des nœuds. Une séquence de ces nœuds est appelée le vecteur nœud et s'écrit $U = (u_0, \dots, u_m)$. U est une séquence non-décroissante de nombre réels. La fonction B-spline de base est alors définie par la formule de récurrence suivante :

$$\begin{aligned} N_{i,0}(u) &= \begin{cases} 1 & \text{si } u_i \leq u \leq u_{i+1} \\ 0 & \text{sinon} \end{cases} \\ N_{i,p}(u) &= \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u) \end{aligned} \quad (14.6)$$

L'équation (14.6) peut résulter en un quotient $0/0$ qu'on définit alors à 0. Le vecteur nœud d'une courbe B-spline non-uniforme est alors un vecteur non-périodique et non-uniforme de la forme :

$$U = (\underbrace{a, \dots, a}_{p+1}, u_{p+1}, \dots, u_{m-p+1}, \underbrace{b, \dots, b}_{p+1}) \quad (14.7)$$

On appelle polygone de contrôle d'une courbe B-spline le polygone obtenu en joignant tous les points de contrôles $\mathbf{P}_i$.

On généralise le résultat obtenu aux courbes rationnelles et on définit les courbes B-splines rationnelles et non-uniformes, c'est-à-dire les courbes NURBS :

$$\mathbf{C}(u) = \frac{\sum_{i=0}^n N_{i,p}(u) w_i \mathbf{P}_i}{\sum_{i=0}^n N_{i,p}(u) w_i} \quad a \leq u \leq b \quad (14.8)$$

où les $\mathbf{P}_i$ sont les points de contrôle, les w_i les poids et les N_i les fonctions B-splines de base définies sur le vecteur nœud non-périodique et non-uniforme U de l'équation (14.7). On peut représenter une courbe rationnelle à n dimensions comme une courbe polynomiale à $n+1$ dimensions en utilisant les coordonnées homogènes. Les points de contrôle en coordonnées homogènes s'écrivent alors $\mathbf{P}_i^w = (w_i x_i, w_i y_i, w_i z_i, w_i)$ dans un espace à 4 dimensions où $w_i \neq 0$. Cette opération est équivalente à une projection perspective sur l'origine. On peut alors réécrire l'équation (14.8)

$$\mathbf{C}^w(u) = \sum_{i=0}^n N_{i,p}(u) \mathbf{P}_i^w \quad (14.9)$$

C'est cette dernière équation que nous allons utiliser pour définir les surfaces NURBS, en coordonnées homogènes :

$$S^w(u, v) = \sum_{i=0}^n \sum_{j=0}^m N_{i,p}(u) N_{j,q}(v) \mathbf{P}_{i,j} \quad (14.10)$$

où $\mathbf{P}_{i,j}^w$ forme un réseau bidirectionnel de points de contrôle, $N_{i,p}$ et $N_{j,q}$ sont les fonctions B-splines de base définies sur les vecteurs nœuds :

$$\begin{aligned} U &= \{\underbrace{a, \dots, a}_{p+1}, u_{p+1}, \dots, u_{r-p+1}, \underbrace{b, \dots, b}_{p+1}\} \\ V &= \{\underbrace{c, \dots, c}_{q+1}, v_{q+1}, \dots, v_{s-q+1}, \underbrace{d, \dots, d}_{q+1}\} \end{aligned} \quad (14.11)$$

avec $r = n + p + 1$, $s = m + q + 1$, les limites $[a, b]$ et $[c, d]$ étant généralement fixées à $[0, 1]$.

14.3.2.2 Avantages et inconvénients des NURBS

Après cet intermède théorique, il me semble important de discuter des avantages et inconvénients d'utiliser des NURBS comme surfaces de modélisation. Les avantages sont indéniablement :

- le domaine modélisable de façon exacte est très vaste, la modification interactive est aisée grâce à l'utilisation de points de contrôle et de nœuds. Les NURBS sont capables de modéliser des coniques, ce qui est impossible avec des surfaces de Bézier ;

- le coût calculatoire pour la représentation n'est pas énorme, la plupart des calculs sont numériquement stables ;
- les NURBS sont invariantes par les transformations géométriques courantes (rotation, translation, projection perspective). Il suffit d'appliquer ces transformations aux points de contrôle ;
- les interprétations géométriques sont aisées, les NURBS étant de simples généralisations des courbes et surfaces de Bézier ;
- la représentation est particulièrement compacte pour des formes complexes, qui peuvent ne même pas être représentables par d'autres méthodes.

Malheureusement, ceci ne va pas sans quelques inconvénients :

- les NURBS contiennent énormément d'informations géométriques, la représentation de courbes et de surfaces simples traditionnelles peut présenter des inconvénients en terme de stockage ; ainsi, la représentation d'un cercle dans l'espace ne requiert normalement que le rayon, le centre du cercle et une normale au plan du cercle, soit 7 nombres réels. Si on utilise des NURBS, il faut 7 points de contrôle homogènes et 10 nœuds, soit 38 nombres réels. Cette remarque dépend considérablement du type de surface à représenter ;
- la manipulation des poids peut mener à des constructions aberrantes ;
- les opérations courantes (intersection de surfaces) sont extrêmement difficiles à mener, ce qui peut s'avérer particulièrement gênant en réalité virtuelle ;
- certains algorithmes fondamentaux sont instables numériquement (calcul des paramètres (u, v) d'un point (x, y, z) appartenant à une surface).

14.4 GÉOMÉTRIE ALGORITHMIQUE

Dans cette section, nous nous proposons d'aborder quelques notions de géométrie algorithmique pour la conversion de modèles. Pour des questions de rapidité, les modèles utilisés en réalité virtuelle sont des modèles surfaciques à base de maillages polygonaux. Les objets représentés à l'aide d'autres modèles doivent donc être convertis en objet représentés par des maillages polygonaux. De façon plus générale, nous présentons aussi quelques algorithmes permettant de construire des surfaces à partir d'un nuage de points non structurés. Ces algorithmes sont tout aussi utiles lorsqu'on numérise des objets réels.

14.4.1 TRANSFORMATION D'UN VOLUME EN SURFACE

Quand on a affaire à des données médicales comme des coupes scanner, on utilise d'abord des techniques de segmentation pour extraire les contours intéressants qui caractériseront les volumes à reconstruire, puis on effectue une mise en volume, c'est-à-dire qu'on construit un modèle volumique tel qu'on l'a décrit en 14.2. Pour visualiser ce modèle en temps réel, il faut donc le convertir en un modèle surfacique à base de triangles. La technique de référence est l'algorithme des Marching Cubes, conçu par Lorensen et Cline [Lorensen e.a. , 1987] en 1987. Cet algorithme n'est pas parfait, mais reste universellement reconnu. La suite de ce paragraphe est donc consacré à la présentation de l'algorithme des Marching Cubes.

Il s'agit de découper l'espace en une grille régulière de cubes, de déterminer la topologie de la surface à l'intérieur du cube, de construire les triangles associés et enfin de

passer au cube suivant. On construit donc un cube à partir de quatre points d'une coupe et quatre points d'une coupe adjacente. Il faut ensuite valuer les sommets du cube ; on assigne à chaque sommet 1 si la valeur de la coupe en ce point est supérieure au seuil et 0 sinon. Chaque cube est donc valué par 8 nombres égaux à 0 ou 1, ce qui donne $2^8 = 256$ cas possibles. En mettant à profit toutes les symétries du cube et du problème, on peut limiter l'étude à 14 cas dont 2 sont décrits figure 14.12. En utilisant une interpolation linéaire pour déterminer les points d'intersection de la surface recherchée et des arêtes du cube, on construit ainsi de 0 à 4 triangles par cube. L'algorithme possède diverses améliorations, mais le principe reste celui-ci. Les limitations des Marching Cubes portent sur la génération possible de triangles aberrants.

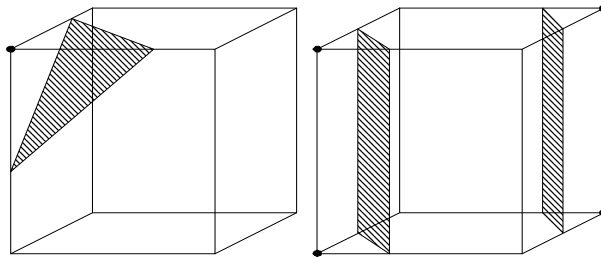


Figure 14.12 : Exemple de construction des triangles à partir de la valuation des sommets des cubes (les sommets valués à 1 sont indiqués par un point noir).

La figure 14.13 présente une coupe scanner d'une mâchoire humaine numérisée dans le cadre du projet Visimplant. Les coupes scanner sont segmentées afin de séparer les différents matériaux présents. L'ensemble des coupes scanner est utilisée pour reconstruire un modèle en trois dimensions de la mâchoire à l'aide de l'algorithme des Marching Cubes modifié pour éviter la génération de triangles aberrants. La figure 14.14 montre une vue de cette reconstruction [Dutreuil, 2001].

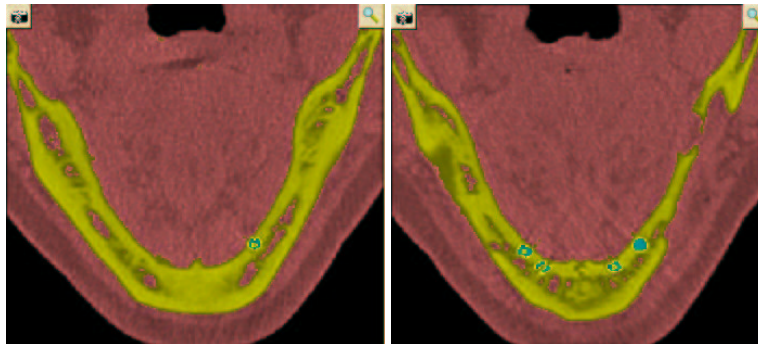


Figure 14.13 : Coupes scanner d'une mâchoire humaine (image Ecole des Mines de Paris/Centre de Robotique)

14.4.2 MAILLAGE POLYGONAL D'UN NUAGE DE POINTS

Lorsqu'on utilise un scanner laser, on obtient en sortie un nuage de points non structuré. Il va falloir construire un maillage polygonal (de préférence de triangles) à partir de ce

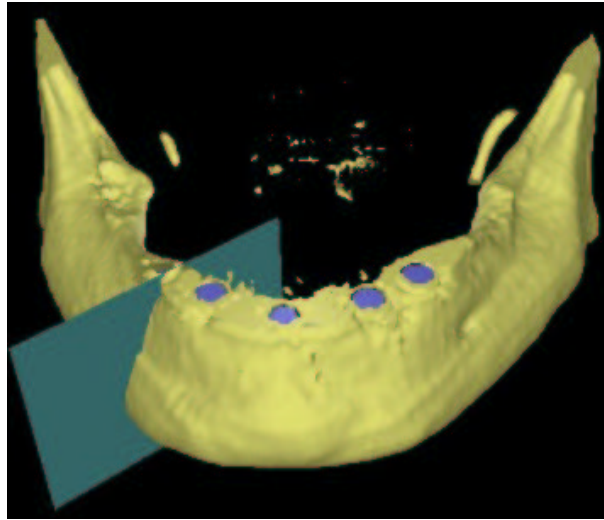


Figure 14.14: Reconstruction 3D de la mâchoire à l'aide de l'algorithme des Marching Cubes (image Ecole des Mines de Paris/Centre de Robotique)

nuage. Le problème du maillage d'un nuage de points consiste donc à déterminer la surface S construite à partir d'un ensemble de points $p \in P$ de la surface réelle S , vérifiant :

$$\forall p \in P, \exists x \in S, x = p \text{ ou } x \text{ proche de } p \quad (14.12)$$

La suite de ce paragraphe est inspirée de l'excellent état de l'art présenté par Mencl et Müller à Eurographics'98 [Mencl e.a., 1998] et complété entre autres par le rapport du MPI dirigé par Leif Kobbelt [Kobbelt e.a., 2000]. Avant d'effectuer une reconstruction de surface, il faut se poser quelques questions : de la qualité de l'échantillonnage dépend le résultat. Si l'échantillonnage n'est pas dense, certaines méthodes ne fonctionneront pas. Si celui-ci n'est pas assez précis, la reconstruction sera de faible qualité, tandis que s'il est trop précis, le volume de données à gérer sera trop important et il faudra mettre en œuvre des techniques de décimation. Il faut aussi analyser les propriétés de la surface à reconstruire : cherche-t-on à conserver les propriétés topologiques de la structure ? en dernier lieu, le type de surface a aussi son importance : nous nous intéressons ici aux maillages polygonaux, mais il existe d'autres méthodes. Les maillages polygonaux conservent les propriétés topologiques mais ne fournissent pas un aspect très lisse. On peut dénombrer 4 types de méthodes, que nous allons brièvement décrire. Un dernier type de problème ne sera pas abordé car il constitue un domaine de recherche à part entière qui dépasse largement le cadre de ce livre : généralement, les scanners laser sont utilisés plusieurs fois sur l'environnement ou l'objet à modéliser. On obtient de la sorte plusieurs nuages de points 3D dont le recalage par rapport à un repère absolu peut présenter des défauts. Sans aborder le problème du recalage proprement dit, les algorithmes de construction de surfaces doivent prendre en compte les erreurs de recalage des nuages de points.

14.4.2.1 Méthodes de subdivision spatiale

Le principe de ces méthodes consiste à subdiviser la boîte englobante de P en un ensemble de cellules disjointes (grilles régulières, octrees, tétraèdres). La seconde étape consiste à sélectionner les cellules traversées par des surfaces, puis à calculer une surface à partir des cellules sélectionnées.

La méthode la plus simple consiste donc à diviser la boîte englobante de P en voxels (ce sont les éléments d'une grille régulière en 3 dimensions). La seconde étape consiste à déterminer les voxels qui contiennent au moins un point de P . La construction de la surface se fait alors en utilisant les quadrilatères extérieurs aux voxels sélectionnés. On peut alors découper les quadrilatères en deux triangles et utiliser différentes méthodes de lissage pour obtenir un aspect visuel plus agréable. On peut éventuellement déformer la surface obtenue pour la rapprocher des points de P . On peut aussi utiliser l'algorithme des Marching Cubes décrit en 14.4.1.

Edelsbrunner et Mücke proposent une méthode utilisant les α -shapes. On commence par construire une tétraédrisation de Delaunay [Boissonnat e.a., 1995] de P , puis on efface des tétraèdres, des triangles et des arêtes en utilisant des α -boules : tous les tétraèdres, triangles ou arêtes dont la sphère englobante minimum n'est pas contenue dans la boule d'effacement de rayon α sont effacés. Le résultat est appelé une α -shape. La troisième étape consiste à extraire les triangles en utilisant les règles suivantes : on considère les deux sphères de rayon α passant par les trois points d'un triangle. Si au moins une des deux sphères ne contient aucun autre point de P , le triangle appartient à la surface. Le principal problème de cette méthode réside dans le choix du α . C'est un paramètre global, un α trop petit peut créer des trous dans les surfaces et déconnecter certains éléments.

Il est aussi possible d'utiliser une décomposition en cellules orientée-volume, où l'on élimine les cellules qui n'appartiennent pas au volume délimité par la surface. Ces méthodes sont essentiellement fondées sur la tétraédrisation de Delaunay. Boissonnat propose par exemple d'éliminer certains tétraèdres (il commence par ceux qui ont 2 faces, 5 arêtes et 4 points). Cette approche ne permet pas de reconstruire des surfaces à trous, Isselhard propose donc de modifier la procédure d'effacement des tétraèdres, en permettant l'élimination de nouveaux types de tétraèdres. De façon identique Veltkamp utilise des γ -indicateurs (une fonction qui permet de classer les triangles vis-à-vis d'une sphère) et élimine les triangles par γ -indicateurs décroissants. Cette technique s'adapte très bien aux différences de densité du nuage de points, mais comme l'approche de Boissonnat, elle ne permet pas reconstruire les surfaces contenant des trous.

14.4.2.2 Méthodes à fonction de distance

Une fonction de distance en géométrie algorithmique est une fonction dont la valeur absolue est la plus courte distance entre un point de l'espace et un point de la surface. Elle est négative lorsque le point se trouve à l'intérieur du volume délimité par la surface et positive autrement. Le point clé de ce type d'approche réside dans le calcul de la fonction de distance.

Hoppe [Hoppe e.a., 1992] suggère l'idée suivante : à chaque point $\mathbf{p}_i \in P$, on associe

un plan tangent estimé. Cette estimation est obtenue en prenant k points de P dans le voisinage de $\mathbf{p}_i$ et calculant le meilleur plan approximant au sens des moindres carrés. La distance est calculée en déterminant le plan tangent le plus proche du point dont on cherche la distance à la surface. Le problème est alors de trouver une orientation cohérente des plans tangents pour déterminer le signe de la fonction de distance. On utilise pour ceci le graphe de Riemann².

Comment utiliser une fonction de distance ainsi définie ? La méthode de Hoppe *et al* est une méthode à sélection de cellules qui utilise une fonction de distance : on commence par une décomposition régulière en voxels. Les voxels sélectionnés sont ceux dont les sommets sont de signe opposés. La surface est alors obtenue en utilisant l'algorithme des Marching Cubes (cf. section 14.4.1).

14.4.2.3 Méthodes de déformation

Il s'agit cette fois de déformer une surface initiale jusqu'à ce qu'elle donne une bonne approximation de la forme, c'est-à-dire qu'elle constitue une bonne approximation des points de P . Ces méthodes sont particulièrement utiles lorsqu'on dispose d'une première approximation grossière de la surface. On distingue trois types de méthodes assez différentes :

- purement géométrique : elle consiste à déformer tout l'espace (et donc l'objet qui s'y trouve) en fournissant des points sources et des points de destination (typiquement des points de la surface), et en interpolant entre ces points ;
- approche physique : c'est tout simplement une approche masses-ressorts telle que décrite en 15.3.4. Les points du maillage sont reliés aux points de P par des ressorts ;
- approche informatique : on utilise ici un réseau de neurones (feature map de Kohonen) que l'on va entraîner pour faire coller la surface aux points de contrôle. La difficulté réside dans le fait qu'il faut arriver à lier les poids et les points de la surface. L'intérêt est de pouvoir modéliser a priori n'importe quel type de surface.

14.4.2.4 Méthodes à croissance de surfaces

L'idée de la construction incrémentale de surface est de construire les surfaces à partir des points et de leurs propriétés orientées surface. Ceci peut se faire de plusieurs façons très différentes.

On peut commencer par relier deux points proches à un endroit de P et ajouter des triangles dont les sommets sont de proches voisins des points déjà reliés. C'est ce que propose notamment Boissonnat : on commence par relier les deux points les plus proches de P . Pour attacher un nouveau triangle à ces points (et plus tard aux arêtes de la frontière), on estime localement un plan tangent grâce à un voisinage. Les points sur le voisinage de la frontière sont projetés sur le plan tangent ; le nouveau triangle est

²Le graphe de Riemann est le graphe dont les sommets sont les centres o_i des plans tangents définis comme les centroïdes des k points utilisés pour les estimer. Deux plans sont connectés ssi le centre de l'un d'entre est le k -voisinage de l'autre. On affecte comme poids aux arcs une valeur égale à 1 moins la valeur absolue du produit scalaire entre les normales aux plans. L'orientation est alors déterminée en propageant une orientation de référence par un arbre couvrant minimal.

obtenu en connectant un de ces points à la frontière (celui qui maximise l'angle sous lequel il voit les deux autres points du triangle).

Mencl et Müller [Mencl e.a., 1998] proposent une approche qui utilise la reconnaissance de caractéristiques dans les nuages de points. Cette technique se décompose en plusieurs étapes :

- calcul d'un EMST, c'est-à-dire un graphe couvrant euclidien minimal, qui servira de base au graphe de description de surface ;
- ce graphe est étendu aux feuilles de l'arbre, en reliant chacune des feuilles à ses voisins les plus proches ;
- reconnaissance de caractéristiques : l'idée est d'intégrer différents algorithmes de reconnaissance dans l'algorithme principal tout en conservant la cohérence structurelle d'un graphe de description de la surface. Cela permet de s'adapter aux différences de densité des nuages de points, aux différents types de surfaces rencontrés, voire même de reconstruire un ruban de Möbius.

14.4.3 DÉCIMATION DE MAILLAGES

Afin d'être utilisés dans une application de réalité virtuelle, les maillages polygonaux doivent comporter un nombre de polygones *raisonnable*. Les modèles fournis, par maillage d'un volume ou d'un nuage de points comportent assez régulièrement trop de polygones. Plus simplement, les maillages issus de logiciels de CAO comportent un niveau de détail trop important pour être visualisés en temps réel. On cherche alors à réduire le nombre de triangles d'un modèle. Cette technique s'appelle la décimation. Notons que cette technique s'applique tout aussi bien pour la génération des niveaux de détail (cf. 14.5.2, page 426).

De façon générale, on peut dire que la décimation est la création d'une approximation polygonale d'un modèle complexe à la précision suffisante pour l'application considérée. Cela induit quelques problèmes généraux :

- quelles sont les données de départ ? Il existe différentes approches pour créer des maillages de taille réduite, certaines d'entre elles ne procèdent pas à partir d'un maillage, mais proposent une décimation directement basée sur les nuages de points. Hoppe [Hoppe e.a., 1993] propose par exemple de minimiser une fonction de coût sur un maillage à partir d'un nuage de point. Le maillage initial fournit simplement un point de départ pour l'algorithme d'optimisation ;
- quelles sont les propriétés géométriques des modèles qui doivent être considérées ? Cherche-t-on la précision de l'éclairage, le contact entre les surfaces, l'aspect rugueux d'une surface ?
- la notion d'approximation nécessite la définition d'une erreur d'approximation ;
- une fois définie cette erreur, comment définir la précision *suffisante* à une application donnée ?
- en dernier lieu, les maillages actuels issus des systèmes de numérisation 3D sont véritablement énormes (le projet Digital Michelangelo obtient des modèles à plusieurs milliards de polygones [Levoy e.a., 2000]) et ne peuvent même pas être chargés en mémoire d'un seul tenant. La structuration des maillages et l'accès aléatoire à leurs éléments est donc une des nouvelles composantes de la recherche dans le domaine de la décimation.

Il existe une large variété de techniques de décimations de maillage dans la littérature, pas forcément dans uniquement dans le domaine de l'informatique graphique d'ailleurs. Ainsi, les techniques d'éléments finis qui impliquent le maillage de la surface ou du volume d'un objet mettent aussi en œuvre la décimation.

Historiquement, on peut distinguer deux types d'approches : celles qui sont un sous-ensemble du maillage existant, et celles qui recréent un nouveau maillage. L'approche la plus simple en apparence est celle de Schröder [Schröder e.a., 1992]. Elle consiste à choisir un des sommets du maillage, à l'enlever, puis retriangler la surface autour de ce sommet. Au contraire, Turk [Turk, 1992] préfère générer de nouveaux sommets dans les régions de la surface, le nombre de sommets générés dépendant de la courbure locale de la surface. Les nouveaux points sont retriangulés. Cette méthode garantit la génération de triangles ayant de bonnes propriétés, mais ne permet guère de garantir une bonne approximation du maillage initial.

14.4.3.1 Algorithmes incrémentaux

L'algorithme incrémental de Schröder [Schröder e.a., 1992] est paramétré par une borne d'erreur ϵ qui donne non pas un maximum, mais un minimum ! Le schéma général de l'algorithme est le suivant :

Répéter

trouver une région dont l'erreur est inférieure à ϵ
 enlever des triangles dans la région
 retriangler la région

Jusqu'à aucune région admissible

Malheureusement, cette méthode ne donne pas de borne sur la distance entre le maillage initial et le maillage final. De plus, les transformations appliquées sont locales pour résoudre un problème global. Il existe des extensions de cet algorithme qui prennent en compte ce problème : il s'agit alors d'utiliser des mécanismes de file d'attente pour hiérarchiser les transformations en fonction de leur intérêt. Là encore, aucune propriété de minimisation globale n'est assurée. La question du choix des triangles à éliminer reste posée. Schröder distingue plusieurs types de triangles, comme le montre la figure 14.15. Il pose des conditions sur la connectivité des sommets et des arêtes. Les triangles considérés par son algorithme sont les deux premiers de la figure, les seconds ne sont pas considérés comme admissibles. Les algorithmes de décimation, à l'instar de celui de Schröder ne traitent que très rarement les triangles non admissibles.

14.4.3.2 Opérateurs

Nous avons dit dans le paragraphe précédent qu'on enlevait des triangles dans les régions sélectionnés. Les algorithmes se limitent généralement à un petit nombre d'opérateurs, si possible inversibles (ce type d'algorithme peut aussi être utilisé dans le cas où l'on cherche à raffiner un maillage). Le premier opérateur consiste simplement à enlever un sommet, comme le montre la figure 14.16. Il reste cependant trop de degrés de liberté sur cet opérateur : la retriangulation proposée par la figure n'est qu'une

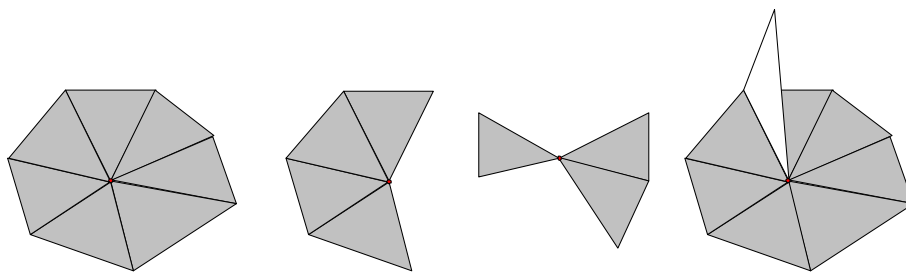


Figure 14.15 : Types de configurations de triangles.

triangulation parmi d'autres.

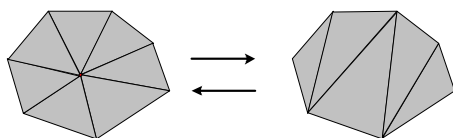


Figure 14.16 : Suppression d'un sommet dans un maillage.

Pour pallier cet inconvénient, il existe un autre type d'opérateur, appelé *edge collapse*, schématisé par la figure 14.17. De sorte, la seule liberté sur l'opérateur reste l'emplacement du nouveau sommet, qui peut se trouver n'importe où sur l'arête qui a été supprimée. Une solution classique consiste à choisir l'un des deux sommets de cette arête, cela s'appelle le *restricted edge collapse*. Ces opérateurs ne modifient pas la topologie du maillage, les régions non connexes du maillage le resteront.

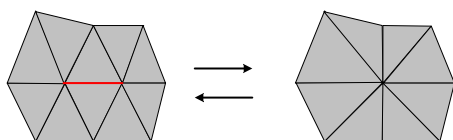


Figure 14.17 : Suppression d'une arête dans un maillage, méthode dite du *edge collapse*.

Il existe aussi des opérateurs qui modifient la topologie du maillage, soit en connectant des parties du maillage qui ne l'étaient pas à l'origine, soit en intervertissant des arêtes, ce qui peut rendre possible l'application des opérateurs précédents dans certains cas.

Tous ces opérateurs ne s'appliquent qu'à des sommets intérieurs aux maillages. Dans le cas des maillages qui ne représentent pas des objets solides, les sommets frontières doivent être traités avec des précautions particulières : il s'agit cette fois d'approximer le contour d'un polyèdre. Les algorithmes classiques peuvent être utilisés, par exemple en conservant le périmètre du polyèdre.

14.4.3.3 Métriques d'erreur

Dans ce paragraphe, nous abordons rapidement la notion d'erreur ou de précision de la décimation. La classification des métriques d'erreur se fait selon trois critères :

- les propriétés mesurées : topologie, géométrie, attributs ;
- la portée de la métrique : l'erreur est-elle évaluée d'étape à étape, ou y-a-t-il une comparaison avec le maillage original ?
- l'interprétation : la métrique est-elle interprétable de façon intuitive ?

La première idée consiste à mesurer la distance qui sépare un maillage du maillage de l'étape précédente [Schröder e.a., 1992]. Malheureusement cette étape est purement locale, ce qui rend l'estimation d'une erreur globale difficile. Cela ne permet pas de disposer d'une borne supérieure de l'erreur fiable et précise, comme la CAO le requiert par exemple. La plupart des métriques cherchent donc à évaluer une borne supérieure garantie de l'erreur. On peut utiliser pour ceci la *distance de Hausdorff*. La distance de Hausdorff entre deux formes A et B est définie comme étant le maximum des distances entre un point de A à son point le plus proche dans B . L'inconvénient de la distance de Hausdorff (qui n'est pas une distance au sens mathématique du terme) outre qu'elle n'est pas symétrique, est qu'elle est particulièrement difficile à calculer. Une alternative consiste à calculer uniquement la distance entre les points du sommet initial aux triangles du maillage décimé. On peut aussi utiliser une toute autre métrique, par exemple en cherchant la conservation du volume.

Les approches modernes visent plutôt à définir des métriques de qualité qui incorporent les notions de précision géométrique des métriques d'erreur, mais aussi des notions comme la qualité visuelle, l'absence de rugosité des surfaces, les limites de courbure, la conservation de la topologie. Plus généralement, les critères sont liés à l'application visée.

14.5 OPTIMISATION DES MODÈLES POUR LA RÉALITÉ VIRTUELLE

Comme l'ont montré les paragraphes précédents, nous disposons d'une variété de modèles plus ou moins adaptés aux contraintes d'une application de réalité virtuelle. La conversion entre ces modèles est bien entendu possible, mais celle-ci n'est pas forcément suffisante : un modèle surfacique à base de triangles ou de polygones peut être trop complexe pour la machine qui va devoir le faire tourner. De ce fait, il existe aujourd'hui un certain nombre de techniques qui visent à diminuer le nombre de triangles affichés à l'écran. Certaines de ces techniques sont utilisables hors-ligne avant la simulation, d'autres sont intégrables à l'intérieur même de l'application. D'autres encore, comme les niveaux de détail, proposent des solutions mixtes, où une partie des calculs est effectuée hors-ligne, tandis que le reste se fait lors de la simulation.

Toutes les techniques citées dans ce qui suit, visent à diminuer le nombre de triangles du modèle ou de la partie du modèle affichée à l'écran. Bien sûr, les algorithmes décrits dans le paragraphe 14.4.3 font partie des techniques d'optimisation des modèles.

14.5.1 TEXTURATION

14.5.1.1 Introduction

Intuitivement, la texture d'un objet représente son état de surface. En informatique graphique, c'est une image plane qu'on va plaquer sur une surface en trois dimensions, qui sera ensuite projetée sur l'écran. Pour connaître la couleur d'un point de l'écran, il suffit donc de connaître la transformation géométrique qui permet de passer du plan de la texture à un point de la surface. Cette fonction constitue ce qu'on appelle les coordonnées de texture. Reste alors à déterminer un algorithme de plaquage de texture.

Les techniques de plaquage de texture diffèrent en fonction de la surface qui reçoit la texture. Idéalement, le plaquage de texture consiste à trouver un isomorphisme entre une image (surface plane définie par des coordonnées $(u, v) \in [0, 1]^2$) et une surface quelconque définie par une fonction $f(x, y, z) = 0$. Sur des surfaces planes, on se contente généralement d'une double interpolation. Cet isomorphisme n'existe évidemment pas pour tous les types de surface. La texturation d'une sphère (par exemple pour représenter la terre) est impossible par des simples algorithmes d'interpolation.

14.5.1.2 Avantages et inconvénients des textures

L'utilisation de textures permet de diminuer le nombre de polygones nécessaires pour représenter un objet. Reprenons l'exemple de la terre du paragraphe précédent. Si on réalise cet objet sans texture, le nombre de polygones nécessaires pour représenter la nature du sol (terre, mer) à laquelle il faut ajouter les polygones représentant les nuages, la modélisation devient plus fastidieuse et requiert un nombre de polygones beaucoup plus important. Ici, quelques photos satellites et une sphère (décimée en quelques centaines de triangles) ont suffi à représenter l'image de la figure 14.18.



Figure 14.18 : Représentation de la terre à l'aide de textures.

Les cartes graphiques 3D récentes incorporent en plus des algorithmes d'affichage 3D pré-câblés et des fonctions de texturation des objets 3D. Les images constituant les

textures sont stockées dans une mémoire spécifique appelée mémoire de texture embarquée sur la carte graphique. L'accès est donc beaucoup plus rapide que pour la mémoire centrale et un processeur spécialisé se charge du plaquage de textures déchargeant alors d'autant le processeur principal. En contrepartie, les textures fournissent naturellement une qualité d'affichage inférieure (dans l'exemple précédent, la perception ne s'améliore pas si on se rapproche de la terre). L'utilisation de cartes graphiques à mémoire spécialisée est une technique relativement coûteuse puisque la mémoire utilisée a des temps de réponse extrêmement faibles : celle-ci est donc généralement disponible en quantité limitée.

14.5.2 NIVEAUX DE DÉTAILS

La technique des niveaux de détail consiste à diminuer le nombre de polygones d'un objet affiché à l'écran en fonction de la façon dont il est perçu par l'utilisateur : un objet lointain ou petit nécessite moins de détail qu'un objet proche ou gros. Ceci est vrai pour deux raisons principales :

- le centre d'intérêt de l'image est généralement au centre de l'écran et de taille respectable ;
- l'image est constituée d'une suite de pixels discrets, les détails dont la projection est inférieure à la taille du pixel ne sont donc pas représentables.

L'utilisation de niveaux de détails dans une application suppose de disposer d'une part d'un certain nombre de modélisations d'un même objet à des précisions différentes et d'une fonction de détermination du modèle à utiliser. Dans les fichiers de données 3D classiques représentant les modèles 3D par des arborescences, un nœud niveau de détail (Level Of Detail soit LOD en anglais) se présente comme sur la figure 14.19.

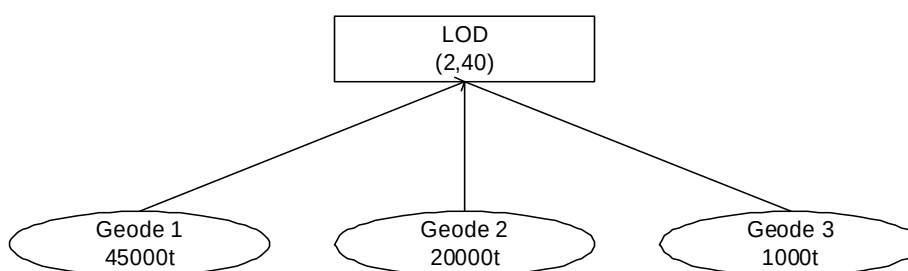


Figure 14.19 : Représentation des niveaux de détail dans l'arborescence d'un fichier de données 3D

S'il est possible de modéliser trois fois le même objet à des précisions différentes, il est évidemment plus intéressant d'effectuer une seule fois la modélisation avec une bonne précision et d'utiliser ensuite des outils de décimation automatique. Ce calcul se fait classiquement hors-ligne, mais peut s'effectuer aussi en ligne. La fonction de choix du niveau de modélisation à visualiser est calculée en ligne.

14.5.2.1 Commande de transition

Il existe plusieurs critères de transition d'un niveau de détail à un autre. Le principal paramètre utilisé est bien sûr la distance entre la caméra virtuelle et l'objet considéré. On utilise une table de décision qui donne le niveau de détail à utiliser en fonction de cette distance. Cette technique est insuffisante puisqu'elle ne tient pas compte de la taille des objets : un gros objet éloigné peut nécessiter plus de détail qu'un petit objet à proximité de la caméra. Plus généralement, le niveau de détail requis par un objet dépend de la taille à l'écran de l'objet plus que de sa taille réelle. L'angle sous lequel est vu l'objet a donc aussi son importance.

En dehors de ces critères, il existe d'autres données à prendre en compte. J. Nouvel [Nouvel, 1998] les explicite pour le cas du pilote de chasse dans son simulateur de vol :

- l'importance relative de l'objet dans son environnement : un arbre au milieu d'une plaine peut constituer un repère de vol, alors qu'il est parfaitement anodin au milieu d'une forêt ;
- l'intérêt fonctionnel de l'objet : toujours pour le pilote de chasse, un avion présente un intérêt largement supérieur à un arbre. Cette technique, appelée «technique des aires d'intérêt» peut s'appliquer à l'ensemble de la base de données 3D ;
- la vitesse linéaire de l'objet par rapport à l'observateur : plus l'objet se déplace vite, moins ses détails sont visibles ;
- la vitesse angulaire du modèle : plus un objet tourne vite sur lui-même, moins ses détails seront visibles ;
- la position de l'objet sur l'écran : l'œil est plus précis en son centre qu'à sa périphérie, les objets situés dans l'axe du regard nécessitent donc plus de détails. De plus, le champ de vision a tendance à se rétrécir avec la vitesse. On nuancera ce propos par le fait qu'il n'existe aujourd'hui que des prototypes de dispositifs permettant d'effectuer le suivi des mouvements de yeux ;
- la visibilité dans l'environnement : les objets dans le brouillard disparaissent plus vite ;
- enfin, la charge du générateur d'images : on peut choisir d'éliminer des objets (ou d'utiliser des niveaux de détail plus grossiers) lorsque la machine graphique n'arrive plus à générer le nombre d'images suffisant.

Ces critères sont bien sûr à adapter pour les autres applications, mais ils donnent les trois idées de base ; on réduit les détails en fonction de la distance entre l'observateur et l'objet, de la charge machine et de critères fonctionnels liés au domaines de l'application.

Une fois la décision de transition prise, il faut l'appliquer au modèle. La solution la plus naturelle consiste à remplacer purement et simplement un modèle de l'objet par un autre. Cette technique est évidemment très peu coûteuse, mais peut provoquer des artefacts visuels gênants (clignotements des objets éloignés qui se rapprochent) lors du changement et perturber l'attention de l'utilisateur. À l'inverse le fondu-enchaîné (ou *blending*) consiste à afficher les deux modèles en même temps en rendant l'un de plus en plus transparent tandis que l'autre devient de plus en plus opaque. Les artefacts disparaissent, mais au détriment du coût de calcul : il faut effectuer le rendu des deux modèles, mais aussi effectuer des calculs de transparence. Les techniques d'interpolation 3D constituent alors un bon compromis entre la commutation et le *blending* : il

s'agit de déterminer hors-ligne des relations entre les sommets des niveaux de détail et d'effectuer une interpolation entre ces sommets lors de la simulation.

14.5.2.2 Génération des niveaux de détail

La méthode générale pour calculer des niveaux de détail se décompose en deux phases :

1. calcul de l'importance : il s'agit de calculer l'intérêt d'un sommet pour la reconnaissance du modèle auquel il appartient ;
2. construction : une fois que les sommets ont été triés en fonction de leur importance, on génère le nouveau niveau de détail lié à un seuil d'importance, par exemple en éliminant les sommets dont l'importance est faible. L'approche par raffinement existe aussi dans la littérature (approche multi-résolution par ondelettes [Eck e.a., 1995]) où on construit le modèle complet à partir d'une forme simple en ajoutant au fur et à mesure les sommets du modèle le plus raffiné.

14.6 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Boissonnat e.a. , 1995] J.-D. Boissonnat and M. Yvinec. *Géométrie algorithmique*. Ediscience International (1995).
- [Cozot, 1996] R. Cozot (1996). *Environnement de simulation des systèmes physiques*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Donikian, 1997] S. Donikian. Multilevel modeling of virtual urban environments for behavioral animation. In *Computer Animation* (1997), pages 127–133, Genève, Suisse.
- [Donikian e.a., 1998] S. Donikian, G. Moreau, and G. Thomas (1998). Multimodal driving simulation in realistic urban environments. In S. G. Tzafestas and G. Schmidt, editors, *Progress in System and Robot Analysis and Control*. Springer-Verlag.
- [Dutreuil, 2001] J. Dutreuil (2001). *Modélisation 3D et robotique médicale pour la chirurgie : développement d'un protocole d'implantologie dentaire assistée par robotique et étude d'un cas clinique*. Thèse de doctorat, Ecole des Mines de Paris, Paris.
- [Eck e.a., 1995] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetzle. Multiresolution analysis of arbitrary meshes. In *Computer Graphics* (1995), pages 173–181.
- [Fuchs, 1996] P. Fuchs. *Les interfaces de la réalité virtuelle*. Les Presses de l'Ecole des Mines de Paris (1996).
- [Hoppe e.a., 1992] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. In *SIGGRAPH'92*, volume 26 of *Computer Graphics* (1992), pages 71–78.
- [Hoppe e.a., 1993] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Mesh optimization. In *Computer Graphics (SIGGRAPH'93 Proceedings)* (1993), pages 19–26. ACM.
- [Jackins e.a. , 1980] C. L. Jackins and S. L. Tanimoto. Oct-trees and their use in representing three-dimensional objects. *Computer Graphics Image Processing*, 14 :249–270 (1980).
- [Kobbelt e.a., 2000] L. Kobbelt, S. Bischoff, K. Kähler, R. Schneider, M. Botsch, C. Rössl, and J. Vorsatz (2000). Geometric modeling based on polygonal meshes. Rapport de recherche MPI-I-2000-4-002, Max Planck Institut für Informatik, Sarrebruck, Allemagne.

- [Levoy e.a., 2000] M. Levoy, S. Rusinkiewicz, M. Ginzton, J. Ginsberg, K. Pulli, D. Koller, S. Anderson, J. Sade, B. Curless, L. Pereira, J. Davis, and D. Fulk. The digital michelangelo project : 3d scanning of large statues. *Proceedings of SIGGRAPH 2000*, pages 131–144 (2000).
- [Lorensen e.a. , 1987] W. E. Lorensen and H. E. Cline. Marching cubes : A high resolution 3d surface construction algorithm. *Computer Graphics (Proceedings of SIGGRAPH 87)*, 21(4) :163–169 (1987), Held in Anaheim, California.
- [Mencl e.a. , 1998] R. Mencl and H. Müller. Interpolation and approximation of surfaces from three-dimensional scattered data points. In *Eurographics State of The Art Report* (1998), Lisbonne, Portugal.
- [Mortenson, 1985] M. Mortenson. *Geometric Modeling*. John Wiley & Sons (1985).
- [Nouvel, 1998] J. Nouvel (1998). *Niveaux de détail dynamiquement adaptatifs pour objets et terrain*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Schröder e.a., 1992] W. Schröder, J. Zarge, and W. Lorensen. Decimation of triangle meshes. In *Computer Graphics (SIGGRAPH'92 Proceedings)*, volume 26 (1992), pages 65–70.
- [Thomas, 1999] G. Thomas (1999). *Environnements virtuels urbains : modélisation des informations nécessaires à la simulation de piétons*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Turk, 1992] G. Turk. Re-tiling polygonal surfaces. In *Computer Graphics (SIGGRAPH'92 Proceedings)*, volume 26 (1992), pages 289–298.

Adresses des sites Web

- [http1] <http://www.sgi.com/software/volumizer>
- [http2] <http://www.povray.org>

15 COMPORTEMENTS DES OBJETS DU MONDE VIRTUEL

Guillaume Moreau

15.1 INTRODUCTION

Avant de parler d'animation par ordinateur, nous pouvons très rapidement parler de l'origine de l'animation. Il s'agit évidemment des films d'animation, à commencer par les dessins animés, type Tex Avery ou Disney. C'est Disney qui a posé les principes de l'animation en couches ou *cell animation*. Différents calques représentant différents plans de la scène sont superposés. Cette technique permet de ne redessiner qu'un minimum d'objets. Les inconvénients restent naturellement le dessin manuel des différentes couches mobiles, et les limitations intrinsèques du modèle : Ce modèle à plusieurs plans ne permet que des animations 2D ou des travellings horizontaux. Dans le cadre de la réalité virtuelle, nous nous intéressons à une animation d'objets 3D sans restriction sur les mouvements de caméra qu'on peut assimiler à un objet animable. La problématique est double : il s'agit d'un part de paramétrer les descriptions géométriques des modèles (2D ou 3D), puis de contrôler l'évolution de ces paramètres dans le temps.

On s'intéresse à un système continu qui évolue au cours du temps. Une animation est une séquence d'images, où chaque image est une observation du système à un instant t . Les images d'une animation sur l'écran d'un ordinateur constituent une discrétisation à la fois temporelle (fréquence d'affichage) et spatiale (résolution de l'écran) des états du système. On considère donc une animation comme étant une suite $E_t = E_{t_0+n\delta t}$ d'états du système pris à chaque pas de temps. L'animation consiste donc d'une part à décrire les relations entre les E_t (c'est-à-dire les mouvements et les déformations) et d'autre part à les évaluer. Les modèles de contrôle du mouvement sont donc le cœur de tout système d'animation.

B. Arnaldi et G. Hégron [Hégron e.a. , 1992] donnent une classification à trois niveaux des modèles de contrôle du mouvement fondée sur le niveau d'abstraction de la description du mouvement. Le plus bas-niveau est nommé *niveau descriptif* et reproduit les effets décrits sans se préoccuper de ce qui les a produits ; par contre, le niveau immédiatement supérieur, dit *niveau générateur*, s'intéresse aux causes physiques du mouvement. Mais la description du mouvement uniquement à l'aide de la physique s'avère rapidement compliquée à mettre en œuvre pour la simulation des organismes vivants : un niveau d'abstraction supplémentaire est nécessaire, il s'agit du *niveau comportemental*. Le développeur cherche par exemple à donner un ordre concret à son personnage de synthèse : «gare-toi derrière la pendule de la gare». Cet ordre induit un comportement, qui se traduit par des actions physiques (le déplacement d'une voiture). Ce déplacement physique est représenté à l'écran par un ensemble d'images décrites les unes après les autres par le simulateur physique.

Dans ce chapitre, nous allons décrire ces trois types de modèles de mouvement, en nous intéressant plus particulièrement aux modèles comportementaux.

15.2 MODÈLES DESCRIPTIFS

15.2.1 MODÈLE

Les modèles descriptifs s'attachent à reproduire les effets des mouvements sans se préoccuper de leurs causes. L'utilisateur cherche essentiellement à exprimer E_t sous une forme explicite. Les E_t représentent alors une trajectoire spatio-temporelle dans l'espace de description du système à animer [Arnaldi, 1994], appelé aussi espace d'états. Il existe deux façons de lier E à t :

- uniquement en fonction du temps :

$$E_t = F(t) \quad (15.1)$$

- à la fois en fonction du temps et des autres E_t :

$$E_t = F(t, E_{t-\delta t}, E_{t-2\delta t} \dots) \quad (15.2)$$

Comme il s'agit d'automatiser le processus de création des images, l'animateur va chercher à en dessiner le moins possible, tout en faisant calculer le reste à l'ordinateur. Son travail consiste à fournir un certain nombre de positions clé (à la fois dans l'espace et dans le temps). Ensuite il reste à effectuer une interpolation spatiale ou spatio-temporelle entre les points clé pour calculer l'animation. Cette technique s'appelle le *keyframing*.

15.2.2 ANIMATION PAR KEYFRAMING

Comme le montre la figure 15.1, le keyframing consiste à définir des images clés, puis une méthode d'interpolation entre ces images clés. La figure évoque aussi un des problèmes fondamentaux des méthodes d'interpolation : si les deux images clé sont l'image de gauche et l'image de droite, le mouvement vers le haut est occulté lors de l'interpolation. Le choix des images clé repose alors uniquement sur l'animateur.

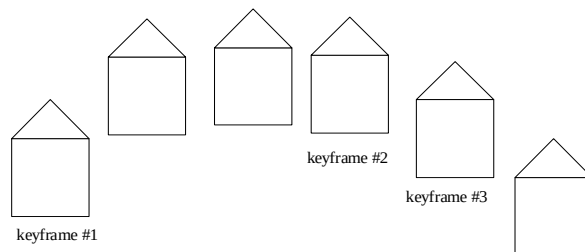


Figure 15.1 : Interpolation entre des images clé

Pour un simple point matériel animé, le keyframing pose encore un certain nombre de questions auxquelles il faut pouvoir répondre :

- Comment choisir le paramétrage ?

- Quel type d'interpolation utiliser ?
- Comment garder le contrôle sur l'animation ?

La définition de trajectoires de façon paramétrique plutôt qu'implicite semble évidente. Pour tout objet qui n'est pas un point, nous pouvons choisir d'autres types de paramétrage : on peut paramétrer l'orientation dans l'espace d'un solide (attention cependant aux problèmes de continuité, voir [Watt, 2000] pour une description des méthodes d'interpolation de rotations), sa taille, sa couleur, les différents éclairages d'une scène, la forme générale d'un objet. En dernier lieu, il ne faut pas oublier que le mouvement de la caméra s'apparente aussi au mouvement d'un objet solide. Cela reste un problème difficile puisqu'il faut exprimer mathématiquement ce que l'animateur veut obtenir à l'écran !

La technique d'interpolation la plus simple est bien entendu l'interpolation linéaire. Les positions clés sont données par l'animateur, puis l'ordinateur interpole linéairement entre ces positions. L'interpolation entre formes ayant le même nombre de sommets est possible grâce à cette technique, même si les formes intermédiaires n'ont pas de raison de ressembler à quoi que ce soit de réaliste. L'autre problème qui survient est que le mouvement d'un objet réel est extrêmement rarement un mouvement linéaire. En ajoutant des positions clé, l'animateur augmente son contrôle sur le mouvement mais diminue l'intervention de l'ordinateur. En limitant le nombre de positions clé, le mouvement peut avoir l'air saccadé. Une interpolation non-linéaire apporte un effet bénéfique à l'aspect des trajectoires. L'utilisation de courbes splines ou de simples cubiques de Bézier s'est généralisé dans les logiciels d'animation. Les courbes sont décrites sous forme paramétrique. Les problèmes classiques de l'influence globale des points de contrôle sur des courbes locales reste posé, mais le problème principal n'est pas ici : la vitesse du point animé dépend du type d'interpolation mis en œuvre.

On considère un point matériel animé sur une trajectoire interpolée $Q(t)$. On note $s = A(t)$ la longueur de l'arc parcouru à l'instant t . Pour conserver une vitesse constante, on doit reparamétriser non pas directement par le temps, mais par la longueur de l'arc. La trajectoire devient alors $Q(A^{-1}(u))$. Cette technique s'appelle reparamétrisation des arcs (on la trouve aussi sous le nom de redéfinition du temps [Arnaldi, 1994]). De la même manière la reparamétrisation peut être faite directement par une courbe de vitesse, si l'animateur cherche une vitesse non constante.

On remarque que le keyframing s'adapte à tout paramétrage défini par l'animateur. Il s'applique notamment aux paramètres articulaires d'un solide articulé. Prenons l'exemple d'un bras de pelleuse que l'animateur cherche à amener d'une position de repos à une position particulière (cf. Figure 15.2). L'intérêt de l'animateur n'est pas ici de paramétrer les angles articulaires du bras, mais plutôt de spécifier les positions initiale et finale de la pelle, et de faire calculer à l'ordinateur les valeurs articulaires. C'est ce que se propose de faire la cinématique inverse.

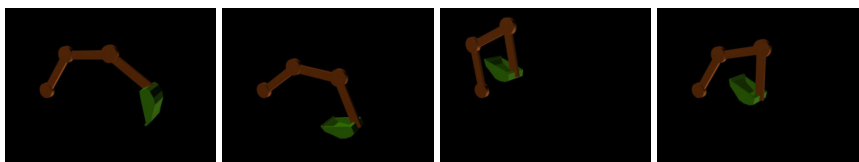


Figure 15.2 : Animation d'un bras de pelleuse

15.2.3 CINÉMATIQUE DIRECTE ET INVERSE

Le keyframing et la cinématique inverse représentent avec le *motion capture* (cf. section 15.10, page 461) les techniques les plus utilisées aujourd'hui dans le domaine du cinéma d'animation ou des jeux vidéo (qui ont les mêmes contraintes de temps réel que la réalité virtuelle).

Le lien entre le point terminal et les valeurs articulaires d'un système articulé comme le bras de pelleteuse de l'exemple précédent s'écrit de façon générale comme indiqué par l'équation (15.3).

$$\mathbf{x} = f(\theta) \quad (15.3)$$

où $\mathbf{x}$ représente les coordonnées de l'effecteur, θ les valeurs articulaires. f est alors appelé le modèle géométrique direct. Le problème précédent consiste à évaluer θ en fonction de $\mathbf{x}$:

$$\theta = f^{-1}(\mathbf{x}) \quad (15.4)$$

Bien sûr, dans le cas général, f n'est pas inversible, la plupart des méthodes de résolution [Watt, 2000] s'orientent vers une résolution itérative du problème. On écrit d'abord l'équation de la cinématique directe :

$$\frac{d\mathbf{x}}{d\theta} = J(\theta) \quad (15.5)$$

$J(\theta)$ est la jacobienne de f . L'algorithme de cinématique inverse consiste à résoudre de façon itérative le problème (15.6).

$$d\theta = J^{-1}(\theta)(d\mathbf{x}) \quad (15.6)$$

Le schéma général de l'algorithme s'écrit donc ainsi :

Répéter

$d\mathbf{x}$ = petit mouvement dans la direction de $\mathbf{x}$

$d\theta = J^{-1}(\theta)(d\mathbf{x})$

$\mathbf{x} = f(\theta + d\theta)$

$J = \frac{d\mathbf{x}}{d\theta}$

inverser J

$\mathbf{x} = \mathbf{x} + d\mathbf{x}$

Jusqu'à but atteint

Le problème n'est pas résolu pour autant puisque la jacobienne J n'a pas de raisons d'être inversible. Elle n'est même pas forcément carrée ! Diverses approches existent alors, la plus simple consiste à utiliser la pseudo-inverse de J , d'autres méthodes

cherchent à minimiser l'erreur avec des moindres carrés pondérés. Il n'y a pas de solutions universelles tant les causes de problèmes sont nombreuses :

- il peut ne pas y avoir de solution unique. Par exemple la main humaine peut atteindre la même position de plusieurs façons différentes ;
- il peut ne pas y avoir de solution. C'est le cas dans les systèmes surcontraints ;
- rien ne prouve que le mouvement reproduit à l'aide de la cinématique inverse est un mouvement réaliste, en particulier pour le mouvement humain. Les contraintes habituelles de la commande robotique (minimisation d'énergie par exemple) ne s'appliquent que très mal aux mouvements de l'être humain [Multon e.a., 1999]. Une approche fondée sur la minimisation des *jerks* (dérivée de l'accélération) semble plus adaptée.

15.3 MODÈLES GÉNÉRATEURS

15.3.1 MODÈLE

Les modèles descriptifs ne permettent pas à l'objet animé d'agir ou de réagir par lui-même, ce qui contredit les principes d'interaction de la réalité virtuelle ; contrairement aux modèles descriptifs, les modèles générateurs permettent de décrire les causes du mouvement. Il s'agit donc d'intégrer les lois de la mécanique comme cause du mouvement [Arnaldi, 1988]. Les intérêts des modèles générateurs par rapport aux modèles descriptifs sont multiples [Cozot, 1996a] :

- **réalisme des mouvements** : les corps animés à l'écran réagissent aux lois de la physique. Cela garantit donc le réalisme de l'animation par rapport à la simulation ;
- **réduction du nombre de degrés de liberté du système** : l'utilisation des modèles générateurs peut permettre de réduire le nombre de paramètres du mouvement puisque ceux-ci sont régis par les lois de la mécanique. Un corps en chute libre n'est plus soumis qu'à l'action de la gravité, l'animateur n'a donc pas à se préoccuper de sa trajectoire ;
- **paramétrage naturel** : le paramétrage d'une animation s'effectue à l'aide de grandeurs physiques (masse, élasticité..) et ne dépend donc pas du savoir-faire de l'animateur.

Là encore, on retrouve les mêmes distinctions que celles que nous avons faites sur les types de modèles géométriques en 14.1.1. Dans le cas de la réalité virtuelle, nous nous intéressons d'abord principalement aux solides rigides et aux mécanismes rigides articulés, même si les modèles déformables sont de plus en plus utilisés dans les systèmes mettant en œuvre le retour d'effort. Nous commencerons cette partie sur les modèles générateurs par quelques rappels de mécanique des solides, puis nous décrirons rapidement le formalisme proposé par R. Cozot [Cozot, 1996a] pour unifier les différents modèles de matière pour la simulation.

15.3.2 RAPPELS SUR LA MÉCANIQUE DES SOLIDES

15.3.2.1 Cinématique du solide rigide

Le solide rigide [Bamberger, 1981] se caractérise par une absence totale de déformation quels que soient les efforts appliqués. Son mouvement est donc régi par 6 paramètres (3 en translation et 3 en rotation). On peut donc écrire que le mouvement d'un point M d'un solide S en fonction du temps t :

$$M_t = G(M_0, t) = B_t + R_t M_0 \quad (15.7)$$

où B_t est un vecteur de translation et R_t une matrice de rotation. En dérivant l'équation (15.7) par rapport au temps, on obtient

$$V(M, t) = \frac{\partial}{\partial t} B_t - \Omega_s \times (B_t - M_t) \quad (15.8)$$

Le champ de vitesse des points d'un solide s'écrit alors

$$V(M, t) = V(M', t) + \Omega_s \times M' M \quad M, M' \in S \quad (15.9)$$

En redérivant, on obtient l'accélération en tout point M d'un solide

$$\Gamma(M, t) = \frac{\partial^2}{\partial t^2} B_t + \frac{\partial}{\partial t} \Omega_s \times R M_0 + \Omega_s \times (\Omega_s \times R M_0) \quad (15.10)$$

15.3.2.2 Dynamique du solide rigide

La dynamique d'un solide rigide est donnée par deux formalismes équivalents : la *seconde loi de Newton* (aussi appelée principe de la dynamique) et le *principe des travaux virtuels*.

La seconde loi de Newton exprime que la variation de quantité de mouvement d'un solide est égale à la somme des forces extérieures qui agissent sur le solide, soit

$$\frac{d}{dt} \int_S \rho(M) V(M, t) d\tau = \int_S f(M, t) d\tau \quad (15.11)$$

où $\rho(M)$ représente la masse volumique de S en M et $f(M, t)$ les forces qui s'appliquent en M à l'instant t . A l'aide des propriétés de la cinématique du solide rigide et des définitions des moments d'inertie, on obtient alors les deux *relations fondamentales de la dynamique*

$$\begin{aligned} m \Gamma_a &= F_a \\ I_a \dot{\Omega}_s + \Omega_s \times I_a \Omega_s &= C_a \end{aligned} \quad (15.12)$$

pour $a \in S$, m masse de S , F_a résultante des forces en a , I_a la matrice d'inertie définie par

$$I_a^{ij} = \int_S (e_i \times aM) \cdot (e_j \times aM) d\tau \quad (15.13)$$

et C_a la résultante des couples en a définie par

$$C_a = \int_S f(M) \times AM d\tau \quad (15.14)$$

Le principe des travaux virtuels est un formalisme énergétique : on calcule les énergies mises en jeu lors du mouvement, on écrit la loi de conservation de l'énergie qu'on projette dans l'espace des paramètres. Le principe des travaux virtuels exprime alors que pour tout déplacement virtuel d'un corps S , la somme des travaux virtuels est nulle, soit

$$\delta\mathcal{W}_d + \delta\mathcal{W}_j = 0 \quad (15.15)$$

où $\delta\mathcal{W}_d$ est le travail virtuel des actions données et $\delta\mathcal{W}_j$ le travail virtuel des actions d'inertie. Dans l'espace des paramètres, l'équation (15.15) s'écrit

$$\mathcal{D}_i + \mathcal{I}_i = 0 \quad i \in \{1, \dots, n\} \quad (15.16)$$

avec $\mathcal{D}_i$ représentant les actions généralisées données relatives au i -ème paramètre et $\mathcal{I}_i$ les actions généralisées d'inertie relatives au i -ème paramètre. Les équations de Lagrange (exprimées dans l'espace des paramètres q_i) s'écrivent alors ainsi :

$$\mathcal{I}_i = -\frac{d}{dt} \left(\frac{\partial C}{\partial \dot{q}_i} \right) + \frac{\partial C}{\partial q_i} \quad i \in \{1, \dots, n\} \quad (15.17)$$

C étant l'énergie cinétique définie par

$$C = \int_S \rho(M) V^2(M, t) d\tau \quad (15.18)$$

15.3.3 MÉCANISMES RIGIDES ARTICULÉS

Les mécanismes rigides articulés sont construits à partir des solides en ajoutant des liaisons. On considère maintenant un mécanisme rigide articulé composé d'un ensemble fini de n solides rigides. Le mécanisme est donc décrit par $6n$ paramètres. Mais une liaison est une limitation aux libertés de mouvement des solides entre eux. Cette dépendance s'exprime par une expression de la forme de l'équation (15.19).

$$f(\mathbf{q}, \dot{\mathbf{q}}) \geq 0 \quad (15.19)$$

Il existe des cas particuliers de liaisons :

- si $f(\mathbf{q}, \dot{\mathbf{q}}) = 0$, on parle de liaison bilatérale ;
- si l'expression de la liaison ne fait pas intervenir les paramètres $\dot{\mathbf{q}}$, on parle de liaison holonome ;
- si un mécanisme est décrit par p liaisons bilatérales, indépendantes et holonomes, il peut s'exprimer en fonction de $n - p$ paramètres. Les $n - p$ paramètres minimaux permettant de décrire la configuration géométrique complète du système sont appelés les degrés de liberté du système.

La *troisième loi de Newton*, appelée *principe d'action-réaction* exprime que si un solide S_i exprime une force $f_{i \rightarrow j}$ sur un solide S_j , alors S_j exerce sur S_i une force $f_{j \rightarrow i} = -f_{i \rightarrow j}$. La dynamique globale des systèmes articulés se calcule alors à partir du formalisme de Newton-Euler ou du principe des travaux virtuels. Dans le premier cas, le formalisme de Newton-Euler s'écrit

$$\begin{aligned} m_i \Gamma_{G_i} &= F_i + f_{i-1 \rightarrow i} - f_{i \rightarrow i+1} \\ I_{G_i} \dot{\Omega}_i + \Omega_i \times I_{G_i} \Omega_i &= C_{G_i} + c_{i-1 \rightarrow i} - c_{i \rightarrow i+1} \end{aligned} \quad (15.20)$$

avec G_i centre de gravité du solide S_i , I_{G_i} la matrice d'inertie de S_i en G_i , C_{G_i} le couples extérieurs en G_i s'exerçant sur le solide, $f_{i \rightarrow j}$ les forces de liaison et $c_{i \rightarrow j}$ les couples de liaison entre deux solides. Si on écrit récursivement les équations (15.20), on obtient un système dont on peut éliminer les termes de forces de couples de liaison.

De la même manière, on écrit le principe des travaux virtuels :

$$\delta \mathcal{W}_d + \delta \mathcal{W}_l + \delta \mathcal{W}_j = 0 \quad (15.21)$$

où $\delta \mathcal{W}_l$ représente le travail des actions de liaison. Après projection sur l'espace des paramètres, on obtient que les actions données relatives à $\mathbf{q}_i$ sont données par :

$$\mathcal{D}_i = \int_S F_M \frac{\partial M}{\partial \mathbf{q}_i} d\tau \quad (15.22)$$

pour tout M point du solide. Les actions d'inertie se calculent à partir de l'expression de l'énergie cinétique par les équations de Lagrange :

$$\mathcal{I}_i = -\frac{d}{dt} \left(\frac{\partial C}{\partial \dot{\mathbf{q}}_i} \right) + \frac{\partial C}{\partial \mathbf{q}_i} \quad (15.23)$$

Les actions de liaison sont prises en compte par les multiplicateurs de Lagrange et par la méthode de pénalisation :

$$\mathcal{L}_i = \begin{cases} \lambda \frac{\partial f}{\partial \mathbf{q}_i} & \lambda \text{ multiplicateur de Lagrange, } f \text{ expression de liaison holonome} \\ \lambda \frac{\partial g}{\partial \mathbf{q}_i} & \lambda \text{ multiplicateur de Lagrange, } g \text{ expression de liaison non-holonome} \\ k \frac{\partial f^2}{\partial \mathbf{q}_i} & k \text{ constante de pénalisation, } f \text{ expression de liaison holonome} \\ k \frac{\partial g^2}{\partial \mathbf{q}_i} & k \text{ constante de pénalisation, } g \text{ expression de liaison non-holonome} \end{cases} \quad (15.24)$$

15.3.4 SIMULATION DE SYSTÈMES DÉFORMABLES

Nous n'entrerons pas dans les détails de la dynamique des systèmes déformables, mais nous nous contenterons de décrire rapidement les deux modèles principaux de simulation que sont la méthode des éléments finis et les modèles dits masses-ressorts.

Dans tous les cas, ces modèles sont d'une utilisation assez marginale en réalité virtuelle pour des questions de temps de calcul. Les systèmes de masses-ressorts autorisent des précalculs pour une utilisation en ligne. Cela permet par exemple de simuler l'apprentissage du geste pour la laparoscopie. Le problème du découpage d'une partie de foie malade est plus complexe puisqu'il modifie le maillage. La simulation temps réel est encore possible mais avec de fortes restrictions. Le lecteur trouvera plus de détails dans [Picinbono, 2001].

15.3.4.1 Méthodes des éléments finis

Le principe de la méthode consiste à discrétiser par un ensemble de volumes élémentaires définis à partir de points élémentaires et structurés par un maillage : les points sont les nœuds du maillage. Ce dernier permet donc d'interpoler les grandeurs physiques en tout point du matériau. A chaque instant (pas de simulation), le matériau évolue et se déforme : le maillage suit alors le matériau. On construit la dynamique du système matériel considéré par assemblage des dynamiques des éléments de base.

L'élément fini est un domaine matériel élémentaire. Il est décrit par un ensemble fini de n points matériels qu'on appelle les nœuds de l'élément. On note $\mathbf{q}_i$ le vecteur des paramètres lagrangiens qui décrivent le i -ème nœud de l'élément. Un point M quelconque de l'élément est alors décrit par un vecteur de paramètres v . Le déplacement u de M est donné par :

$$u(v) = \sum_j^n N_j(v) \mathbf{q}_j \quad (15.25)$$

Les $N_j(v)$ sont les fonctions d'interpolation. Chaque fonction N_j est associée à un nœud $\mathbf{q}_j$ de l'élément fini. Si on note $u = (u_1, u_2, u_3)^t$ le vecteur déplacement, le tenseur des déformation ϵ s'écrit par définition :

$$[\epsilon] = \begin{bmatrix} \frac{\partial}{\partial x_1} & 0 & 0 \\ 0 & \frac{\partial}{\partial x_2} & 0 \\ 0 & 0 & \frac{\partial}{\partial x_3} \\ 0 & \frac{\partial}{\partial x_3} & \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} & 0 & \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} & 0 \end{bmatrix} \cdot (u) = [\partial] \cdot (u) \quad (15.26)$$

Pour un matériau solide, le tenseur des contraintes s'écrit :

$$[\sigma] = [C][\epsilon] \quad (15.27)$$

avec σ tenseur symétrique des contraintes et C tenseur des dilatations. L'interpolation sur l'élément fini s'écrit :

$$(u) = [N](\mathbf{q}) = \begin{bmatrix} N_1 & 0 & 0 & \dots & N_n & 0 & 0 \\ 0 & N_1 & 0 & \dots & 0 & N_n & 0 \\ 0 & 0 & N_1 & \dots & 0 & 0 & N_n \end{bmatrix} \cdot (q_1^1, q_1^2, q_1^3, \dots, q_n^1, q_n^2, q_n^3)^t \quad (15.28)$$

On applique le principe des travaux virtuels défini par l'équation (15.29), sur l'élément fini de référence.

$$\int_V \delta u \rho \gamma dv + \int_V [\delta \epsilon][\sigma] dv = \int_V \delta u f dv + \int_S \delta u [\sigma] \cdot n ds \quad (15.29)$$

On obtient :

$$[M_e](\ddot{\mathbf{q}}) + [K_e](\mathbf{q}) - (F_e) = 0 \quad (15.30)$$

où on note

- $[M_e]$, la matrice de masse élémentaire, définie par $[M_e] = \int_V [N]^t \rho [N]$;
- $[K_e]$, la matrice de rigidité élémentaire, définie par $[K_e] = \int_V [N]^t [\delta]^t [C] [\delta] [N]$;
- (F_e) , le vecteur des résultantes aux nœuds défini par $(F_e) = \int_V [N]^t \tilde{f}$.

Le maillage est construit par assemblage d'éléments de référence. Chaque élément de référence est obtenu par translation, orientation et déformation d'éléments de référence. La transformation τ entre l'élément fini de référence réel et l'élément fini de référence est une fonction bijective : un point du domaine réel se transforme en un unique point du domaine de référence et inversement. On peut écrire τ sous la forme :

$$x_{reel}^k = \sum_i \bar{N}_i(u, v) x_i^k \quad (15.31)$$

où les $x_i^k, k \in \{1, 2, 3\}$ sont les coordonnées du nœud $\mathbf{q}_i$. Les fonctions $\bar{N}_i$ sont appelées les fonctions de forme. Si elles sont égales aux fonctions paramétriques N_i , l'élément est dit *isoparamétrique*. Afin de calculer la dynamique de l'élément réel, il est nécessaire de transformer les intégrales sur l'élément réel en intégrales sur l'élément de référence :

$$\int_{V_{reel}} X d\eta d\xi d\psi = \int_{V_{ref}} X \det(J) du dv dw \quad (15.32)$$

avec $\det(J)$ déterminant de la matrice jacobienne J de l'application τ . La matrice de rigidité K_e de l'élément réel s'écrit donc :

$$K_r = \int_{V_{ref}} [N]^t ([J]^{-1})^t [\partial]^t [C] [\partial] [J]^{-1} [N] \det([J]) du dv dw \quad (15.33)$$

De même, on construit la matrice de masse réelle M_r et la résultante réelle des forces extérieures. Les éléments du système dynamique du maillage réel ayant été ainsi construits, la simulation du maillage est effectuée en résolvant le système (15.34) :

$$[M_r](\ddot{\mathbf{q}}) + [K_r](\mathbf{q}) - (F_r) = 0 \quad (15.34)$$

La méthode des éléments finis est très adaptée au calcul des mécanismes solides soumis à de petites perturbations. En effet, le maillage reste constant au cours de la déformation. Dans l'hypothèse où le domaine est soumis à de fortes déformations, les éléments déformés ne constituent pas une bonne approximation du domaine matériel. Ce type de méthode est donc principalement utilisé dans le cadre de calcul de structure. En ce qui concerne les coûts de calcul, la résolution du système (15.34) nécessite l'inversion de la matrice $[M_r]$, dont la taille est égale à la dimension du vecteur de paramètres $(\mathbf{q})$. Le coût de calcul est donc proportionnel au cube de la taille de ce vecteur.

15.3.4.2 Autres méthodes

La méthode alternative la plus connue est la méthode dite *masses-ressorts*. Cette méthode est une approximation de la méthode des éléments finis utilisée pour l'étude des phénomènes vibratoires dans les structures. La matière continue est remplacée par un ensemble de masses ponctuelles m_i , placées aux nœuds $\mathbf{q}_i, i \in \{1, \dots, N\}$ du maillage. La matrice de masse élémentaire $[M_e]$ définie au paragraphe précédent devient alors une matrice diagonale :

$$[\tilde{M}_e] = \text{diag}(m_1, \dots, m_N) \quad (15.35)$$

La matière a donc été modélisée par un maillage de masses ponctuelles, reliées entre elles par des liaisons. Par assemblage, la matrice de masse globale $[M_r]$ reste diagonale. La résolution du système (15.34) ne nécessite plus l'inversion de la matrice de masse globale. Le coût de résolution est alors égal au coût d'évaluation de $[K_r(q)]$. Dans le cas d'études vibratoires, on ne s'intéresse pas à tous les types de déformation. Seuls les déplacements linéaires sont pris en compte. L'élément de matière est alors remplacé par une maille composée de masses et de ressorts. Dans ce cas, la matrice de raideur devient très creuse : les forces internes élastiques sont les forces de rappel des ressorts. Le système dynamique se résout alors linéairement en fonction du nombre de masses ponctuelles.

Il existe d'autres méthodes issues d'une simplification de la méthode des éléments finis : elles consistent à projeter les déformations possibles sur une base de fonctions réduites : on utilise alors un ensemble de modes de déformation a priori connus. Le résultat obtenu est donc l'intensité des déformations considérées. Ces méthodes sont appelées méthodes de synthèse modale. Le mode de résolution est encore cubique en le nombre de paramètres du système, mais la synthèse modale permet de limiter ce nombre de paramètres en utilisant les bons modes de déformation. Le problème consiste à pré-déterminer ces modes de déformation.

On trouve aussi dans la littérature des modèles non structurés où la matière est aussi modélisée par un maillage, mais la topologie du maillage est recalculée à chaque pas de temps : ce modèle est appelé *modèle lagrangien libre*. Le coût de calcul est supérieur puisqu'il faut appliquer la méthode des éléments finis, mais aussi une méthode de maillage automatique (par exemple une triangulation de Delaunay). Le modèle Lagrangien libre-libre englobe quant à lui les méthodes où le milieu continu est décrit par un ensemble de particules en interaction.

15.3.5 DISCUSSION SUR LE RÔLE DES MODÈLES GÉNÉRATEURS

En animation comme en simulation, les modèles utilisés sont les mêmes. L'animation reprend les simplifications usuelles de la simulation. L'utilisation de modèles générateurs dans des environnements virtuels pose la question du choix du modèle : les contraintes de l'animation et de la simulation ne sont pas les mêmes (démonstration dans [Arnaldi e.a. , 1992]). Les modèles de simulation se doivent de fournir une haute précision dans un domaine restreint, alors qu'en animation le but est d'obtenir un résultat visuel crédible. Ce résultat ne passe pas forcément par une simulation extrêmement rigoureuse, mais peut nécessiter pour certaines applications des modèles plus précis. Les mouvements d'un fluide parfait peuvent sembler peu convaincants à un observa-

teur habitué à la complexité du monde réel. Dès lors, R. Cozot [Cozot, 1996a] suggère de concevoir le modèle d'animation en fonction des effets qu'on veut observer. Il propose pour ce faire un modèle général utilisable pour l'animation comme pour la simulation [Cozot, 1996b, Cozot e.a. , 1997], intégrant l'ensemble des modèles de matière que nous avons vus dans les paragraphes précédents. La complexité du système est alors le maximum des complexités des éléments qui composent le système. L'intérêt de cette approche est sa complémentarité vis-à-vis des autres techniques utilisées habituellement : le processus traditionnel consiste à linéariser un système possédant un grand nombre de paramètres, alors qu'ici on cherche plutôt à diminuer le nombre de paramètres, mais sans simplifier le modèle.

Toujours dans un souci d'efficacité de la simulation, R. Cozot propose de découpler le modèle géométrique du modèle mécanique : le modèle géométrique représente la forme de l'objet et se déduit bien sûr du modèle mécanique. On peut cependant considérer un support de déformation sur lequel le modèle géométrique sera appliqué. On évite ainsi une surparamétrisation du système. Les logiciels d'animation moderne comme Maya d'Alias/Wavefront permettent d'effectuer ce qu'on appelle le *skinning* : une peau déformable est appliquée sur un squelette humain possédant nettement moins de degrés de liberté. Cela permet de conserver à la fois des temps de simulation courts et un aspect visuel réaliste.

15.4 MODÈLES COMPORTEMENTAUX

Si l'utilisation d'un modèle générateur peut effectivement procurer un mouvement réaliste pour des objets non vivants, cela ne va pas sans certains inconvénients :

- une phase de modélisation du système physique est requise ;
- les objets à animer possèdent des attributs physiques qui ne sont pas toujours faciles à déterminer ;
- enfin, le contrôle du mouvement est le plus gros problème lié à l'utilisation des modèles générateurs : en effet, une partie des équations du mouvement sont intrinsèquement liées au système à animer, tandis que les autres lient les actions sur le système (la commande) et les paramètres du système. Il faut donc être capable, pour un effet donné, de retrouver la commande qui a engendré cet effet, ce qui est dans le cas général, impossible ou très complexe.

En ce qui concerne les être vivants, le problème est encore plus complexe : l'animateur, ou le développeur d'une application de réalité virtuelle cherche à leur donner une certaine autonomie tout conservant des moyens d'interaction. Les jeux vidéo d'aventure fournissent un exemple très pertinent : le héros interagit avec les autres personnages du jeu. Généralement, les interactions sont particulièrement limitées : combat, dialogue pré-enregistrés... dans un jeu de simulation automobile ou de football, les joueurs qui sont remplacés par l'ordinateur ont un niveau de comportement plus abstrait : on peut dire qu'ils *jouent* leur rôle. Idéalement, le rôle du développeur devrait se limiter à une spécification de ce rôle. Un niveau de contrôle complémentaire sur le mouvement s'avère donc nécessaire : en parallèle de cet objectif, l'animation comportementale propose aussi un niveau d'abstraction supplémentaire en tentant de décrire les réponses du système à des stimuli de l'environnement. On considère que le système à animer est contrôlable par un certain nombre d'actions simples qui vont constituer ses réponses.

Le but de l'animation comportementale est donc de spécifier la *décision* du système : en fonction des stimuli qu'il *perçoit*, il *agit*.

Selon B. Arnaldi [Hégron e.a., 1992], les systèmes d'animation comportementale se divisent en deux grandes familles qui distinguent les organismes qui agissent sur eux-mêmes (modèles de transformation interne) des organismes qui agissent sur leur environnement (modèles de transformations externes).

15.4.1 MODÈLES DE TRANSFORMATION INTERNES

Ces modèles s'attachent à décrire de manière réaliste l'évolution interne d'entités vivantes (plantes, êtres humains...) en fonction de paramètres extérieurs. Il s'agit cependant de rendre compte des modifications de l'enveloppe externe de l'entité à animer (par exemple la croissance d'une plante).

15.4.1.1 Modèles de plantes

Les modèles de croissances des plantes sont des systèmes de réécriture fondés sur un axiome et des règles de production appelé *L-systems*. A. Lindenmayer définit un L-system comme étant un axiome constitué d'une chaîne de symboles paramétrés et temporisés et de règles de production qui spécifient comment les symboles se transforment en fonction du temps [Lindenmayer, 1968]. Les symboles sont associés à des primitives géométriques, ce qui permet de donner une représentation graphique au L-system. Dans [Prusinkiewicz e.a., 1994], P. Prusinkiewicz rappelle la définition la plus générale d'un L-system où une règle de production s'écrit sous la forme de l'équation 15.36 :

$$id : lc < pred > rc : cond \longrightarrow succ : p \quad (15.36)$$

où *id* est le nom de la règle de production, *lc* et *rc* sont les contextes gauche et droite, *pred* la partie gauche de la règle de production, *succ* la partie droite de la règle, *cond* la condition de la réécriture et *p* sa probabilité.

La réécriture se fait en parallèle sur les différents éléments. Il y a deux manières de parvenir à une représentation graphique : d'une part, les symboles réécrits sont associés à une primitive graphique et d'autre part on peut insérer dans la chaîne de production des ordres équivalents à ceux que l'on donne à une tortue Logo. À noter que si la réécriture se fait en parallèle, l'interprétation graphique est complètement séquentielle.

Les règles de production des L-systems sont des règles discrètes : P. Prusinkiewicz définit les *dL-systems* en introduisant un temps continu à la place d'une séquence d'étapes de dérivation discrètes [Prusinkiewicz e.a., 1993]. Les paramètres *w* du système sont déterminés en continu par une équation différentielle ordinaire tant qu'ils restent dans un domaine $\mathcal{D}$. Dès que *w* atteint une frontière de $\mathcal{D}$, une règle de production qui dépend de la position sur la frontière, est appliquée. Ensuite, le système est à nouveau régi par l'équation différentielle. La figure 15.3 présente un exemple simple de L-system.



Figure 15.3 : Exemple de plante générée par un L-system

En biologie, les L-systems sont utilisés pour générer des maquettes tridimensionnelles de plantes. C'est la fonction du logiciel AMAP (Atelier de Modélisation de l'Architecture des Plantes) développé par le CIRAD [[http1](#)] : à partir de mesures effectuées sur le terrain et de systèmes de réécriture stochastiques, il construit des plantes en tenant compte des interactions avec l'environnement (notamment la quantité de lumière traversant chaque voxel de l'espace) lors de sa croissance [Jaeger e.a. , 1992, Blaise, 1991]. LA figure 15.4 présente un exemple d'arbres générés par le logiciel AMAP.



Figure 15.4 : Exemple d'arbre généré par le logiciel AMAP du CIRAD

Dans le cadre de l'animation par ordinateur, H. Noser [Noser e.a. , 1997, Noser, 1997] utilise les *L-systems* pour décrire la croissance des plantes qui servent de décor, mais aussi pour contrôler partiellement un humanoïde de synthèse (voir 15.8.2.1).

En conclusion, les L-systems sont utilisés à la fois en animation par ordinateur et en biologie pour simuler la croissance des végétaux. C'est aussi à la biologie (et notamment à la biomécanique humaine) que nous devons les travaux sur les modèles musculaires du corps humain que nous allons aborder maintenant.

15.4.1.2 Humanoïdes de synthèse

Du point de vue de l'animation, l'être humain se compose essentiellement d'un squelette, de muscles et d'une enveloppe extérieure, la peau. Pour animer un humanoïde, il faut définir un modèle d'un ou plusieurs de ces composants. La structure utilisée

généralement est celle d'une structure osseuse (solide rigide articulé) sur laquelle est plaquée une peau déformable. Deux types de problèmes se posent donc : le mouvement du squelette (par exemple la reproduction de la marche humaine) et la déformation de la peau (par exemple, l'animation faciale).

Dans cette étude, nous allons aborder rapidement le problème de la reproduction de la marche humaine et celui de l'animation faciale.

15.4.1.3 *Reproduction de la locomotion humaine*

Le cas de la locomotion humaine est un cas particulier de locomotion. En effet, contrairement à la nage d'un poisson ou au vol d'un oiseau, la locomotion humaine se distingue par les interactions avec un sol situé à altitude variable. Il n'est donc pas possible d'ignorer complètement les contraintes cinématiques. Différentes techniques sont utilisées [Multon e.a., 1999] :

- Dans le cadre du projet Jack [Phillips e.a. , 1988, Lee e.a., 1989] de l'université de Pennsylvanie, X. Zhao [Zhao e.a., 1996] propose une méthode fondée sur la dynamique et la minimisation d'un critère énergétique. La minimisation de l'effort ou de l'énergie ne correspond pas toujours à une vérité biomécanique, à cause d'une corrélation très forte avec la morphologie du sujet. De plus, l'apprentissage de la marche, ou même la marche au pas des militaires sont des exemples de marche *non-économique* ;
- J. Hodgins [Hodgins, 1996] utilise des techniques classiques de contrôle : il y a un contrôleur différent pour chaque phase de la marche et ces contrôleurs sont reliés entre eux par des automates finis. Ces modèles sont extrêmement coûteux en temps de calcul ;
- Dans le cadre de l'animation interactive, des approches purement cinématiques sont proposées [Boulic e.a., 1990]. Plus récemment, R. Boulic [Boulic e.a. , 1996] a intégré des notions de dynamiques en regroupant les articulations dans un arbre de distribution des masses. Ces approches sont intéressantes du point de vue du coût mais permettent difficilement la personnalisation des démarches obtenues. Une amélioration consiste à utiliser des mouvements acquis à l'aide de systèmes vidéo et à les rejouer [Fua e.a., 1998]. Le coût mémoire est alors plus élevé mais la démarche nettement plus réaliste ;
- F. Multon [Multon, 1996] s'intéresse à la construction de modèles biomécaniques de la locomotion humaine. Le cycle de marche est décomposé en différentes phases reliées par des automates. Pour chacun des cycles des trajectoires articulaires sont interpolées par de simples polynômes de degré 3. L'intérêt de l'approche réside dans le fait que les phases identifiées ne sont pas celles qui sont utilisées habituellement en biomécanique ; trois phases sont utilisées pour l'extension du genou et une pour sa flexion, pour la marche comme pour la course. Cette méthode permet en outre l'identification simple de paramètres naturels comme la longueur ou la fréquence des pas.

15.4.1.4 Animation faciale

L'animation faciale consiste à animer le visage d'un être humain. C'est une tâche complexe qui nécessite la modélisation, l'animation et le rendu de la géométrie du visage mais aussi d'éléments distinctifs du visage comme la peau, les cheveux, les lèvres ou la langue. On peut facilement imaginer le niveau de complexité du visage en se remémorant que la tête ne contient qu'une seule articulation, celle de la mâchoire. Tout le reste doit être représenté par des modèles déformables ! P. Fua [Fua e.a. , 1997] classe les modèles d'animation faciale en deux catégories :

- les modèles à peu de degrés de liberté, qui se concentrent sur les mouvements de la tête. Ils utilisent des modèles assez simplistes et sont donc assez pauvres graphiquement ;
- les modèles plus sophistiqués ; il s'agit en premier lieu d'extraire la géométrie à partir d'images réelles [Essa e.a., 1997] ou de la générer automatiquement [Lee e.a., 1995]. Ensuite, il faut définir des paramètres de déformation pour contrôler les mouvements du visage. Trois types de modèles sont utilisés [Dubreuil e.a. , 1997] : des modèles purement physiques, des modèles purement géométriques ou des modèles anatomiques.

Ces deux types de modèles sont en cours d'intégration dans la future norme MPEG-4 (cf. [Escher e.a., 1998, Ostermann, 1998]).

15.4.2 MODÈLES DE TRANSFORMATION EXTERNES

Les modèles dits de transformation externes définissent le comportement extérieur d'un être, c'est-à-dire ses actions et réactions vis-à-vis de son environnement, que ce soit de façon individuelle ou collective.

La classification des modèles de transformation externes n'est pas simple. Dès lors il existe une classification fondée sur le niveau d'abstraction [Zeltzer, 1991], mais aussi des classifications plus techniques [Donikian, 1994]. Une classification technique paraît plus pertinente ici dans la mesure où les techniques employées sont plus faciles à classer.

15.5 APPROCHE STIMULUS-RÉPONSE

L'approche dite *stimulus-réponse* définit le comportement des objets à partir d'un ensemble de capteurs et d'effecteurs ; ils sont reliés entre eux par un réseau dont les nœuds transforment l'information passée [Wilhelms e.a. , 1990]. Le comportement d'un objet dépend alors à la fois de la perception qu'il a de son environnement et des transformations que va subir l'information perçue.

15.5.1 LES RÉSEAUX SAN

Les réseaux SAN (Sensor Actuator Networks) [van de Panne e.a. , 1993] sont des réseaux de neurones avec en entrée des capteurs binaires et en sortie des commandes

pour les effecteurs. Il existe quatre catégories de capteurs :

- les capteurs de contact qui permettent de tester le contact avec le sol ;
- les capteurs de vue qui déterminent si un point appartient ou non au cône de vision ;
- les capteurs angulaires qui vérifient si un angle appartient ou non à un intervalle $[\theta_1, \theta_2]$;
- et les capteurs de longueur qui vérifient qu'une distance mesurée appartient ou non à un intervalle $[l_1, l_2]$.

De la même manière, il n'existe que deux types d'effecteurs :

- les effecteurs linéaires qui génèrent des forces ;
- et les effecteurs angulaires qui génèrent des moments.

Dans un réseau de type SAN, les capteurs sont connectés à tous les nœuds cachés ainsi qu'à tous les effecteurs comme le montre la figure 15.5. De leur côté, les nœuds cachés et les effecteurs sont complètement interconnectés. Les poids affectés aux connexions sont calculés à l'aide d'un algorithme d'optimisation stochastique.

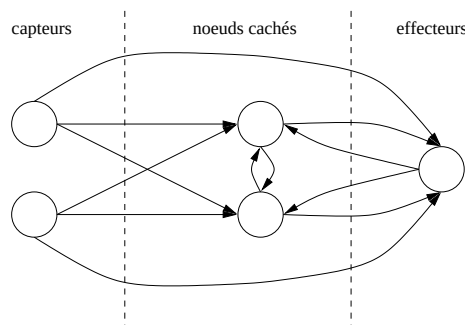


Figure 15.5 : Les réseaux SAN

M. Van de Panne utilise les réseaux SAN pour animer différents types de petits animaux lors de leurs déplacements : certains rampent, d'autres marchent, courent, sautent...

15.5.2 CONTRÔLE STIMULUS-RÉPONSE DE CRÉATURES

J. Thomas Ngo [Ngo e.a. , 1993] propose une approche un peu similaire avec SC (Spacetime Constraints). Son système est composé de la structure physique de la créature à animer, des effecteurs qui contrôlent la configuration interne de la créature et de critères d'évaluation du mouvement. L'algorithme de contrôle du mouvement est un système stimulus-réponse. La créature exécute des *réflexes instinctifs* qui sont déclenchés par des événements qui se produisent dans l'environnement virtuel.

Une *réponse* est un ensemble d'actions sur les effecteurs qui permet d'effectuer des modifications du mouvement de la créature de façon fluide. Les *stimuli* sont introduits à partir de variables de perception : une variable de perception est une fonction à valeurs réelles de l'environnement. J. Thomas Ngo définit :

- des capteurs proprioceptifs (les angles des articulations) ;
- des capteurs tactiles (forces exercées sur les points de contact avec le sol) ;

- des capteurs cinétiques (vitesse du centre de masse) ;
- et des capteurs de position (position du centre de masse).

Une fonction stimulus est donc une combinaison de variables de perception qui s'écrit sous la forme de l'équation (15.37) ; les v_j sont les variables de perception et les λ_j et v_j^0 les paramètres du système.

$$s = 1 - \max_{j=1}^n [\lambda_j (v_j - v_j^0)]^2 \quad (15.37)$$

Les fonctions stimulus sont pondérées par des poids qui sont des fonctions des paramètres λ_j .

Le problème de cette approche, outre le caractère bas-niveau des variables de perception, est bien sûr la détermination des paramètres du système à animer. J. Thomas Ngo utilise un algorithme génétique pour les optimiser. Le domaine d'application favori de ce type de technique est l'animation de personnages articulés (ici mouvements de gymnastique).

15.5.3 SENSE CONTROL ACTION

Les SCA (Sense Control Action) [Granieri e.a., 1995] sont des réseaux comportementaux : chaque comportement est décomposé en sous-comportements implantés sous la forme d'un sous-réseau. Chaque sous-réseau est composé de trois types de nœuds :

- les nœuds *perceptifs* sont des nœuds dont la sortie représente la perception de l'entité à animer. Les capteurs accèdent à la base de données de l'environnement, calculent les distances et les angles de vue d'un objet qui sont les sorties des capteurs. Les capteurs ne diffèrent que par le type d'objet qu'ils sont capables de percevoir :
 - les capteurs d'*objet* détectent un seul objet. Cette détection est globale et effectuée sans tenir compte des problèmes de visibilité ;
 - les capteurs de *localisation* détectent ce qui se trouve à un point précis de l'espace ;
 - les capteurs de *proximité* détectent des objets d'un type donné de façon locale, typiquement dans un cône de vision ;
 - les capteurs de *terrain* détectent la praticabilité du terrain sur lequel évolue le piéton ;
 - les capteurs de *champ de vision* déterminent si un humanoïde est visible depuis un ensemble d'autres humanoïdes. La valeur de sortie du capteur est proportionnelle au nombre d'agents qui peuvent voir le propriétaire du capteur et inversement proportionnelle à la distance de ces agents.
- les nœuds de *contrôle* servent à implanter les boucles de retour d'état de l'automatique en générant des boucles qui visent à minimiser l'erreur perçue en fonction des valeurs désirées de la sortie. Deux nœuds de contrôle ont été implantés : les nœuds *attracteurs* servent à attirer l'humanoïde vers quelque chose qui a été perçu et les nœuds *répulsifs* servent au contraire à éviter un obstacle.
- les nœuds *moteur* sont les nœuds qui vont avoir un effet direct sur l'environnement en faisant effectivement bouger les humanoïdes de la simulation.

Cette approche en tant que telle ne permet pas un niveau d'abstraction important, elle est utilisée pour des actions de bas-niveau comme la simulation de la locomotion humaine. Nous verrons plus loin que combinée à une approche de plus haut niveau, les applications deviennent plus intéressantes.

15.5.4 RÉSUMÉ

L'approche stimulus-réponse à base de réseaux est une première brique de l'animation comportementale et notamment de l'approche «Perception-Décision-Action» telle que nous la décrirons plus loin. La différence entre les approches que nous avons étudiées porte plus sur l'algorithme d'identification des paramètres que sur la nature réelle du réseau utilisé. Les mouvements générés peuvent être très différents, mais le niveau d'abstraction reste très bas. La partie décisionnelle est très limitée, tant dans sa portée que dans son adaptativité. De plus, une modification d'une réaction à un stimulus quelconque de l'environnement nécessite souvent un recalcul complet des paramètres du réseau.

15.6 APPROCHE À BASE DE RÈGLES

Les approches dites à base de règles tentent de répondre au problème du niveau d'abstraction de la description du comportement : la structure reste la même, les données perçues sont transformées en actions. Simplement, il existe un ensemble de règles explicites qui spécifie l'action en fonction de la perception.

De fait, les différents comportements possibles d'un objet peuvent être représentés par un arbre de décision. Cette approche est celle qui se rapproche le plus des systèmes multi-agents [Giroux, 1993, Bouron, 1993] de type cognitif.

15.6.1 NUÉES, TROUPEAUX ET BANCS

Historiquement, C. Reynolds [Reynolds, 1987] a été un des premiers à introduire des règles de comportement pour les objets à animer. Il s'est intéressé au vol des nuées d'oiseaux ; il définit à cette fin un vol géométrique qui est un mouvement le long d'une trajectoire, c'est-à-dire une transformation géométrique spatiale et temporelle d'une courbe 3D (la trajectoire de référence). À chaque objet sont associés un repère local et deux opérateurs de changement de repère. Deux forces contradictoires sont utilisées : l'une sert à rester dans la nuée en suivant la trajectoire de référence, l'autre sert à éviter les collisions. Trois règles de comportement sont définies : *éviter les collisions*, *harmoniser sa vitesse à celle de ses voisins* et *rester proche de ses voisins*. La figure 15.6 présente un exemple de nuée d'oiseaux dont le comportement respecte les règles précédentes.

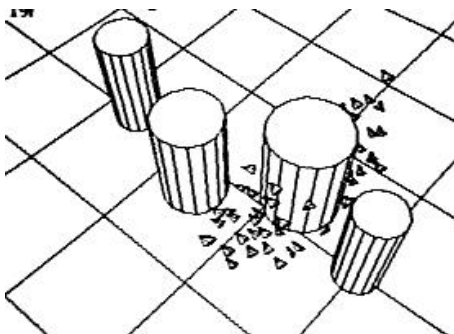


Figure 15.6 : Règles de comportement : une nuée d'oiseaux évitant un obstacle

15.6.2 APPROCHE «ÉCOLOGIQUE»

T. Widyanto [Widyanto e.a., 1991] propose une approche fondée sur la psychologie *écologique*¹ issue des travaux de J. Gibson. Dans son livre [Gibson, 1986], J. Gibson s'intéresse à la perception en tant que moteur de décision du comportement. En effet, les objets qui sont perçus sont *destinés* à une ou plusieurs activités : ainsi, un couteau va être utilisé pour découper, une pomme va être mangée... les informations concernant l'environnement d'une entité comportementale sont disponibles à tout moment et la richesse des informations traitées ne dépend que des capacités de perception d'une entité. Le recueil d'informations est effectué en deux étapes : capture (par lancer de rayon) et évaluation. La partie décisionnelle s'effectue en deux étapes :

- l'orientation du regard au prochain pas de temps est déterminée par une pondération des différents centres d'intérêt du regard ;
- chaque entité possède ses propres buts (à travers un modèle comportemental unique). À chaque pas de temps, les buts sont évalués afin de choisir celui qui sera réalisé à l'aide d'un certain nombre d'actions (ce sont des hiérarchies d'unités fonctionnelles élémentaires). Toutes les unités sont validées et ordonnées par ordre de priorité et la plus importante est effectuée.

Le travail de Widyanto et Marriot se concentre essentiellement sur la perception et incorpore des travaux écologiques. L'intérêt de cette approche réside donc plus dans la sémantique apportée aux données perçues qu'au moteur de décision.

15.6.3 IN VITRAM

Le but de ce système [Mouli e.a., 1992, Mouli, 1994] est de permettre la production de films d'animation, à partir de la mise en scène de personnages autonomes doués d'un comportement propre, le résultat étant la génération de la trajectoire des différents personnages. In Vitram repose sur la notion d'objets et de personnages : les objets contiennent leurs propres fonctionnalités et donc leur propre comportement de réaction aux stimuli de l'environnement. R. Mouli distingue trois niveaux de comportement :

- le niveau *acomportemental* : il est essentiellement utilisé pour décrire les objets du décor qui ne font que subir les forces qui leur sont appliquées ;

¹au sens biologique du terme : étude des êtres vivants au sein des milieux où ils vivent, ainsi que de leur interaction avec ces milieux.

- le niveau réflexe : les personnages de ce niveau vont établir une liaison simple entre la perception et l'action. C'est typiquement le cas d'un interrupteur ;
- le niveau *intelligent* : il est utilisé pour les personnages ayant un comportement complexe. La liaison entre capteurs et effecteurs est réalisée grâce à un ensemble de règles et un moteur d'inférences.

La simulation se fait de manière discrète, en tenant compte des différentes fréquences internes. Pour chaque pas de temps, le cycle «Perception-Décision-Action» est décomposé pour tous les personnages. Cela permet de présenter à un instant donné un état du monde unique et donc à chaque personnage de raisonner sur les mêmes données. La perception se fait à l'aide de capteurs qui simulent les capacités de perception des personnages : ils sont typés dans la perception (vue, ouïe...) comme dans l'espace (sphérique, conique...) et paramétrés (portée, seuil de sensibilité...)

La boucle de simulation d'In VitroAm s'écrit de la façon suivante :

- **Acquisition/Décision**

1. scrutation de l'état du monde ;
2. mémorisation des informations captées ;
3. génération des faits nécessaires à la réflexion (uniquement pour le niveau intelligent) ;
4. raisonnement sur les actions à réaliser (uniquement pour le niveau intelligent).

- **Exécution**

1. exécution des actions automatiques ;
2. exécution des actions déduites par le raisonnement (uniquement pour le niveau intelligent) ;
3. déplacement du personnage.

Dans [Mouli, 1994], R. Mouli décrit l'utilisation d'In VitroAm pour l'animation de fourmis dans une sorte de jeu de la vie : chaque fourmi évolue sur le sol à la recherche de nourriture. Les fourmis sont conçues à partir des personnages intelligents décrits plus haut et munies d'un certain nombre de capteurs. Il existe des versions plus évoluées en ce qui concerne la complexité des comportements des fourmis et de leurs interactions.

15.6.4 DISCUSSION

Les systèmes à base de règles apparaissent plus simples à manipuler que les systèmes stimulus-réponse que nous avons vu précédemment. Malheureusement, la spécification du comportement d'un système à animer en un petit nombre de règles est une tâche difficile à mener à bien. De plus, il existe un double problème de gestion du temps [Rainjonneau, 1992] :

- les algorithmes des moteurs d'inférences ne sont pas compatibles avec le temps réel et donc avec l'animation réellement interactive ou la réalité virtuelle ;
- la spécification à base de règles ne rend pas compte des fréquences de fonctionnement différentes [Lord e.a. , 1994] des activités menées en parallèle dans le cerveau humain.

15.7 APPROCHE OPTIMALE

Cette troisième approche ne s'applique pas à l'animation interactive : l'environnement est entièrement statique et prédéfini. Tout l'espace des solutions pour le mouvement des objets est accessible et le problème de ce type d'animation réside dans le choix d'un critère d'optimalité de l'animation obtenue et dans l'algorithme d'obtention de cet optimum. Cette approche peut être qualifiée de comportementale en ce sens que les critères utilisés peuvent être de nature comportementale (satisfaction de buts propres aux personnages de l'animation), mais n'est pas compatible avec la notion de temps réel. C'est pourquoi nous ne la détaillerons pas ici. On peut néanmoins trouver un exemple dans [Ridsdale e.a. , 1990].

15.8 APPROCHE AUTOMATE

Dans cette section, nous allons nous intéresser à un certain nombre de systèmes d'animation où la partie décisionnelle des entités comportementales est représentée par un ou plusieurs automates. Dans un premier temps, nous décrirons le système utilisé dans le cadre du simulateur de conduite IDS². Les auteurs de ce système reconnaissent que l'utilisation d'un seul gros automate pose des problèmes. C'est pourquoi ils ont proposé le formalisme HCSM qui introduit parallélisme et hiérarchie. Nous verrons aussi d'autres systèmes fondés sur les automates parallèles hiérarchisés.

15.8.1 AUTOMATE D'ÉTATS FINIS

L'Iowa Driving Simulator est un simulateur de conduite temps réel où l'on cherche à proposer à l'utilisateur un environnement virtuel réaliste tant statiquement que dynamiquement : il faut donc inclure dans cet environnement des véhicules virtuels dont le comportement se rapproche le plus possible de la réalité.

Le contrôle de ces véhicules [Booth e.a., 1993] s'effectue grâce à un automate d'états finis. À chaque état de l'automate (figure 15.7) sont associés des attributs qui permettent d'obtenir une commande qui s'adapte à la situation. Les conducteurs sont dotés d'un certain nombre de paramètres (tels que temps de réaction, agressivité...). La sortie du module de conduite est donc un ensemble de commandes pour les véhicules virtuels de la simulation.

Le scénario distribue statistiquement tous ces paramètres et fournit la répartition des véhicules sur le réseau routier.

Les auteurs du conducteur virtuel d'IDS [Booth e.a., 1993] reconnaissent eux-mêmes un certain nombre de limitations dues à l'utilisation d'un automate unique :

- la moindre petite modification du comportement peut nécessiter une refonte globale de l'automate ;
- l'introduction d'une interruption inconditionnelle (exception) en réponse à un environnement requiert l'introduction de transitions de tous les états de l'automate vers un état de gestion de l'exception ;

²Iowa Driving Simulator

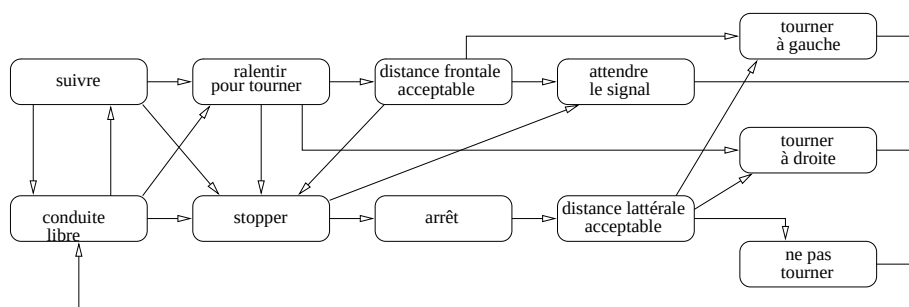


Figure 15.7 : Iowa Driving Simulator : extrait de l'automate de contrôle du conducteur virtuel

- lorsque certains groupes d'états représentent une activité cohérente (dépassement, franchissement d'un croisement...), il n'y a pas de mécanisme qui permette d'absorber le comportement lors de cette activité : la modularité est inexistante ;
- enfin, il est extrêmement difficile d'exprimer la concurrence des comportements : lors de la conduite, plusieurs tâches, éventuellement conflictuelles sont effectuées en même temps. Il appartient au conducteur de prendre en compte toutes les «sorties» possibles de ces tâches et d'en déduire la pression qu'il va exercer sur les pédales ou la manière dont il va tourner le volant.

15.8.2 PILES D'AUTOMATES

15.8.2.1 Virtual Life

Le terme Virtual Life représente un mélange des termes réalité virtuelle et vie artificielle : D. Thalmann [Magenat-Thalmann e.a. , 1984, Thalmann, 1998] cherche à créer une vie artificielle dans un environnement de réalité virtuelle. Les acteurs synthétiques sont doués d'un comportement autonome reprenant une structure «perception-décision-action». La partie décisionnelle est représentée par une séquence d'actions : chaque action est décomposée en un automate séquentiel. Les actions sont empilées par l'utilisateur au début de la simulation (ce qui fournit une sorte de scénario). Une fois qu'une action est complétée, l'automate qui la décrit est dépilé. À noter qu'il est possible qu'un automate ait lui-même besoin de sous-automates pour exécuter sa tâche : dans ce cas, il empile ses sous-automates, suspend sa propre exécution et ne la reprend que lorsqu'il revient au sommet de la pile. Le contrôle de l'humanoïde peut aussi être passé à un système de règles de production [Noser, 1997] : cela permet de définir de nouveaux comportements à partir de quelques règles simples. H. Noser et D. Thalmann [Noser e.a. , 1997] appliquent ce modèle de définition du comportement à la simulation d'un match de tennis entre humains virtuels (cf. figure 15.8³). Ils définissent un certain nombre d'automates régissant le comportement des joueurs comme de l'arbitre. Ainsi, l'automate `play_tennis` d'un joueur doit contrôler les systèmes auditifs et visuels en même temps que son état interne. Si l'automate détecte par exemple, l'événement sonore `faute` émis par l'arbitre, il suspend son exécution en s'empilant

³image due à H. Noser (LIG, Lausanne)

et lance l'automate **marche** pour aller se replacer correctement sur le terrain avant de servir.

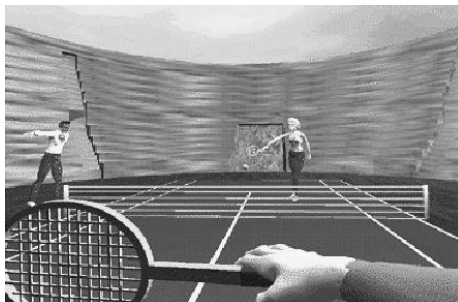


Figure 15.8: *Virtual Life* : jeu de tennis virtuel

Les automates permettent donc de contrôler le comportement du joueur, ainsi que celui de l'arbitre. Le modèle de perception comporte une partie auditive originale, qui est d'ailleurs reprise dans les différents projets du LIG (comportements sociaux et comportements de foule). Par contre, la concurrence et la gestion du temps n'apparaissent pas explicitement.

15.8.2.2 SCA et Pat-Nets

Dans le cadre du projet Jack, N. Badler [Badler e.a., 1997, Webber e.a., 1995] propose un modèle composé à la fois d'une partie stimulus-réponse et d'une partie automates parallèles hiérarchisés. L'architecture globale de contrôle du système est présentée figure 15.9.

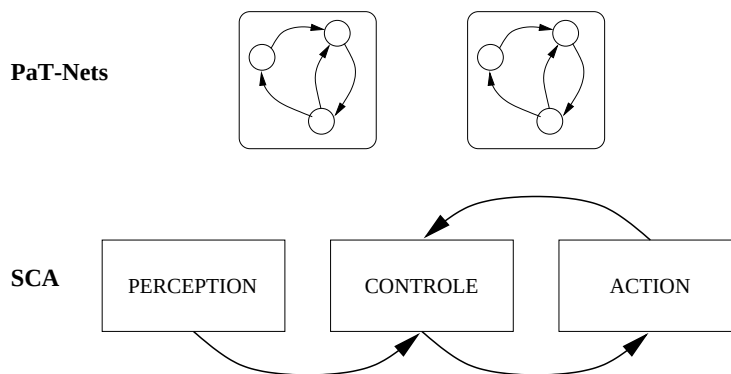


Figure 15.9: *Architecture de contrôle comportemental de Jack*

La partie stimulus-réponse est implantée sous la forme de SCA [Granieri e.a., 1995] : les SCA⁴ sont des boucles de contrôle «Perception-Contrôle-Action» qui servent à produire un comportement réactif localement adaptatif. En général, les comportements

⁴les boucles SCA sont décrites plus longuement en 15.5.3

considérés à ce niveau sont des comportements simples comme la locomotion ou le changement de direction du regard.

Les Pat-Nets [Webber e.a. , 1995, Reich, 1997] sont des automates parallèles hiérarchisés avec des possibilités de passage de messages et de synchronisation (par sémaphores). Des processus sont associés aux états, ils peuvent invoquer d'autres PaT-Nets, des comportements exécutables ou bien des actions de haut-niveau (planification, raisonnement). L'invocation se produit à l'entrée dans un état. Les transitions entre états se font lorsqu'une condition de transition (qui peut être locale à l'humanoïde ou globale) est vérifiée. Les conflits entre transitions franchissables sont résolus en donnant une priorité à chacune d'entre elles. Toutes les primitives classiques sur les processus sont implantées dans les PaT-Nets. L'exécution est effectuée à travers un système d'exploitation écrit en Lisp.

En liant les SCA (flots de données à rétro-action) et les PaT-Nets (contrôleurs à état), il semble possible d'atteindre un niveau de flexibilité assez intéressant en utilisant uniquement le formalisme le plus approprié lors de la phase de spécification du comportement. Le système Jack [http2] témoigne de l'utilisation conjointe de ces deux formalismes dans des cadres applicatifs assez différents qui vont du jeu de cache-cache aux univers virtuels [Webber e.a. , 1995, Badler, 1998]. Toutefois, l'interface entre les différents niveaux de contrôle (SCA et PaT-Nets notamment) n'apparaît pas clairement. De plus, il semble que tous les systèmes comportementaux fonctionnent à la même fréquence et que la résolution de conflits entre comportements ne soit soluble que par l'assignation d'une priorité à chaque transition dans un PaT-Net.

15.8.3 CONCURRENCE ET HIÉRARCHIE : HPTS

15.8.3.1 Fondements

Le modèle HPTS [Moreau e.a. , 1998] se base sur les affirmations de Newell quant aux propriétés d'un *bon* modèle de comportement [Newell, 1990]. Ces propriétés sont les suivantes :

- un comportement flexible et adaptatif⁵, c'est-à-dire que l'organisme simulé doit s'adapter à son environnement et à l'évolution de celui-ci afin d'effectuer ce qu'il est supposé faire, sans mettre en danger son existence ;
- des interactions temps réel avec l'environnement. Le temps réel est obligatoire lorsqu'on se place dans le cadre de la simulation interactive et donc de la réalité virtuelle ;
- un environnement riche et complexe : une entité simulée doit pouvoir percevoir son environnement, doit avoir une base de connaissance sur cet environnement (c'est-à-dire qu'elle possède un moyen de se représenter son environnement), et enfin elle doit avoir plusieurs degrés de liberté d'actions sur cet environnement ;
- représentation symbolique et abstraction de la connaissance : une entité doit posséder des informations symboliques sur son environnement ;
- possibilité d'apprentissage : une entité doit être capable de tenir compte de son expérience passée ;

⁵Le mot adaptatif utilisé par Newell doit plutôt être pris au sens *adapté* qu'au sens *adaptatif* couramment utilisé en robotique

- autonomie dans un environnement social : la «vie» d'une entité se déroule dans un environnement où il y a d'autres entités. Il doit alors y avoir une vie sociale, c'est-à-dire des possibilités de coopération, de conflit...
- conscience de soi, perception de soi : en plus des informations qu'elle possède sur son environnement, une entité doit pouvoir raisonner sur elle-même, et modifier ses buts. C'est ce que Th. Bouron [Bouron, 1993] appelle la réflexivité ;
- être réalisable par un réseau neuronal : toutes les activités précédemment citées doivent pouvoir être réalisées par un réseau de neurones ;
- être constructible par un processus embryologique ;
- émergence par évolution.

Moreau traduit en termes informatiques les contraintes issues des propriétés précédentes ainsi que de la psychologie comportementale [Moreau, 1998]. Ces contraintes sont reprises dans le tableau 15.1. HPTS est construit sur ces contraintes en y ajoutant la prise en compte explicite du temps.

Contrainte	Motivation
Modularité	mise à jour
Réactivité	simulation interactive
Hierarchie	architecture de contrôle
Concurrence	activités menées simultanément
Préemption	résolution des conflits
Exceptions	traitement des situations exceptionnelles
Gestion du temps	activités cognitives de fréquence différentes
Flots de données	communications entre les différents niveaux

Tableau 15.1: Modèles de comportement

15.8.3.2 Définition

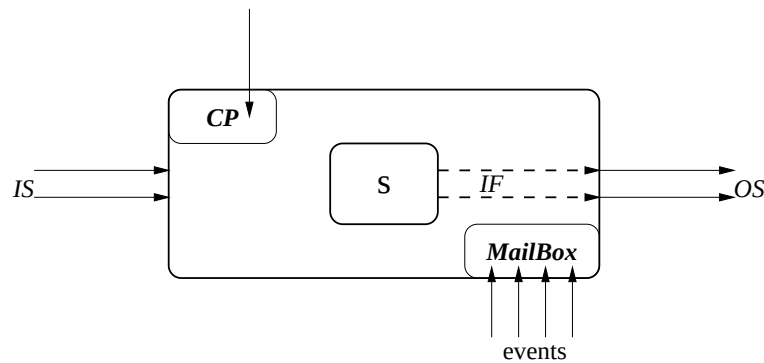


Figure 15.10: HPTS : architecture

Le modèle décisionnel HPTS (Hierarchical Parallel Transition Systems) est un système réactif, constitué d'une hiérarchie d'automates (qui sont les comportements possibles).

Chaque machine à états peut être vue comme une boîte noire (cf. figure 15.10) qui comporte des flots d'entrée et de sortie, et des paramètres de contrôle. On définit donc une machine à états comme étant :

$$s = \langle S, \Gamma, IS, OS, CP, LV, IF, M, \nu \rangle \quad (15.38)$$

où

- S est un ensemble de sous-états ;
- Γ est la fonction d'activité des sous-états ;
- IS est un ensemble de signaux d'entrée ;
- OS est un ensemble de signaux de sortie ;
- CP est un ensemble de paramètres de contrôle ;
- V est un ensemble de variables locales ;
- IF est la fonction d'intégration ;
- M est la boîte aux lettres du système ;
- ν est la fréquence de fonctionnement.

s est soit un état atomique du système, dans ce cas $S = \emptyset$, soit un état composite ($S \neq \emptyset$). À chaque état s est attribué un paramètre d'activité a_s tel que :

$$a_s \in \mathcal{A} = \{idle, active, suspended\} \quad (15.39)$$

La fonction d'activité Γ est donc la fonction qui fait évoluer le paramètre d'activité $a_{s'}$ de tous les sous-états $s' \in S$. Par exemple, après avoir débuté, un état se retrouve actif. La fonction d'activité Γ à valeurs dans $\mathcal{A}$ s'écrit :

$$\forall s' \in S, a_{s'}(t) = \Gamma(a_{s'}(t-1), IS, CP, V, M) \quad (15.40)$$

Il existe une suite logique entre les statuts. La fonction d'activité est contrainte par cette suite, qui peut être elle-même représentée sous la forme d'un automate repris figure 15.11. Le carré situé dans l'état *active* signifie que la fonction d'intégration (décrite plus loin) est exécutée. Les transitions entre états se font à chaque pas de temps de simulation. Il n'est donc pas possible de passer d'*idle* à *suspended* en un seul pas de temps.

La boîte aux lettres M est un système de messagerie qui permet de recevoir des messages (provenant notamment des sous-états ou de l'état père). Un message est ici un événement accompagné de données. On distingue un certain nombre d'événements spécifiques qui commandent la fonction d'activité Γ : ce sont les événements qui apparaissent sur la figure 15.11 et qui permettent de modifier l'état d'une machine HPTS. À chaque primitive est associé un événement renvoyé par la machine HPTS à qui le message était destiné. Cet événement permet à la machine de signaler son nouveau statut.

On distingue deux types de messages :

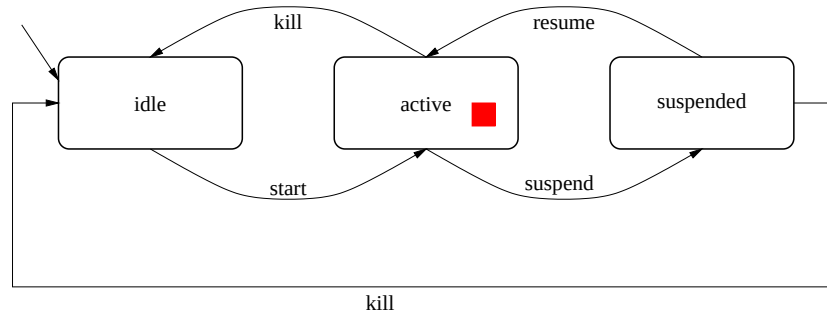


Figure 15.11 : HPTS : automate de fonctionnement

- les messages globaux : leur diffusion est assurée dans un environnement de broadcasting, c'est-à-dire que toutes les machines HPTS y ont accès, et que toutes peuvent les utiliser ;
- les messages locaux : ceux-ci ont un destinataire particulier spécifié lors de l'envoi. Ce mécanisme permet de mettre en œuvre la communication entre machines HPTS, voire entre entités comportementales. On pourrait imaginer une diffusion générale de tous les messages, où le protocole de communication incorporerait le destinataire (par exemple comme dans le protocole IP), mais ce serait moins efficace d'un point de vue transmission de données.

La fonction d'intégration gère la cohérence entre les différents comportements proposés par les sous-états. On distingue trois cas :

- $S = \emptyset$: s est donc un état atomique, IF est donc une fonction qui ne dépend que des entrées IS , des variables locales et des paramètres de contrôle ;
- $\text{card } S = 1$: s a un seul sous-état s_1 actif, dont la sortie est le résultat de la fonction d'intégration de s_1 appliquée sur les sous-états actifs de s_1 . Typiquement, la fonction IF de s est alors une fonction identité. On peut imaginer que IF soit aussi un filtre ;
- $\text{card } S > 1$: c'est le cas le plus complexe : IF est alors l'intégration des différents comportements proposés par les sous-états actifs s' de S . Nous étudierons ce problème plus tard. De manière formelle, IF s'écrit simplement :

$$OS(s) = IF(OS(s'), \forall s' \in S), V, IS, PC) \quad (15.41)$$

Un exemple relativement simple de fonction d'intégration, dans le cadre de la simulation de conduite, est présenté figure 15.12. Les trois sous-états de l'état *supervisor* proposent chacun une action paramétrée. Aux actions sont associées des priorités, comme le montre la relation d'ordre définie dans la boîte *action*. Si deux états proposent la même action, une loi conservative est appliquée sur ses paramètres : en cas de demandes de freinage simultanées, la distance d'arrêt la plus courte est choisie.

15.8.3.3 Aspects temporels

Le n-uplet qui définit une machine HPTS incorpore un paramètre ν , défini comme la fréquence d'exécution. En pratique, seul le cas où la fréquence d'exécution d'un automate s est un multiple de la fréquence de son automate père a été envisagé. Tout

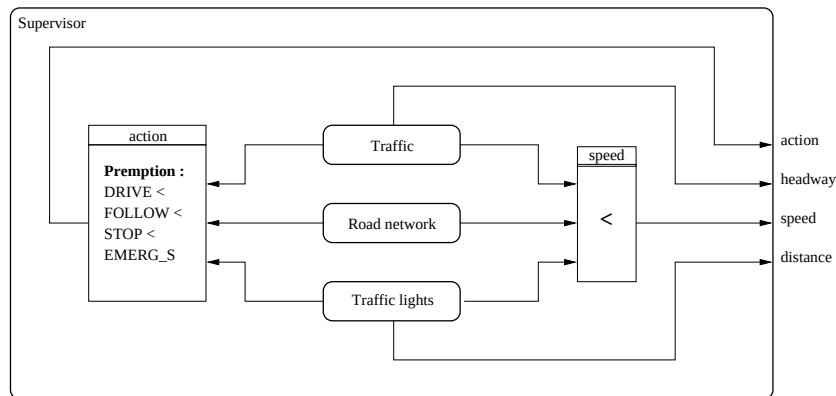


Figure 15.12: HPTS : exemple de fonction d'intégration

en évitant les problèmes de déphasage, cela pose néanmoins un certain nombre de problèmes, que nous allons aborder dans les paragraphes suivants.

Sorties et variables locales

Si on se place dans le cas d'un module de simulation comportementale qui doit produire des données pour un contrôleur de mouvement à une fréquence de référence ν_0 , les sorties OS de la machine HPTS de plus haut-niveau s_0 doivent être produites elles-aussi à une fréquence ν_0 . Cela implique que les fonctions d'intégration des automates placés au niveau le plus-bas de la hiérarchie doivent être exécutées aussi à une fréquence nominale ν_0 , puisque ce sont elles qui ont le temps de réponse le plus faible (cf. [Moreau, 1998] première partie, sur la théorie du contrôle). Maintenant, si les fonctions d'intégration des sous-automates produisent leurs données à une fréquence ν_0 , il devient nécessaire d'intégrer ces différents comportements proposés à cette même fréquence. On en déduit donc que la fréquence de calcul de IF pour s et tous ses sous-automates est égale à la fréquence de calcul des sorties OS .

Les variables locales pourraient être calculées de deux manières : soit en même temps que la fonction d'activité, soit en même temps que la fonction d'intégration. Pour permettre des calculs temporaires sur les comportements et des calculs liés au temps (donc à la fréquence d'activité), les deux possibilités sont offertes.

Événements

Par opposition aux entrées, les événements sont sporadiques. Néanmoins, pour gérer les différences fréquentielles, ils doivent avoir une certaine durée de vie, c'est-à-dire jusqu'à leur consommation effective par l'automate auquel il sont destinés. C'est la fonction de la boîte aux lettres M . Celle-ci est donc remplie de façon asynchrone, mais consultée et vidée à chaque exécution de la fonction d'activité, seule consommatrice de ces événements.

Fonctions d'activité

Nous avons montré que la gestion des différentes fréquences de fonctionnement de la hiérarchie n'était pas possible au niveau des sorties, elle se fait donc au niveau des fonctions d'activité. La fonction d'activité de l'automate de plus haut-niveau s est exécutée à sa fréquence ν , et les sous-automates sont exécutés à des fréquences ν_i multiples de ν . Grâce au mécanisme de boîte aux lettres, cela permet à un automate de haut-niveau de contrôler à l'aide d'événements des fonctions de plus bas-niveau.

Entrées

Les entrées doivent être disponibles à chaque fois qu'un calcul le nécessite, c'est-à-dire :

- lors du calcul de la fonction d'intégration, soit à la fréquence ν_0 ;
- lors du calcul de la fonction d'activité de s et de chacun de ses sous-automates, soit à $\bar{\nu} = \text{ppcm}(\nu, \text{ppcm}_{i \in S}(\nu_i))$;
- lors de l'évaluation des variables locales, c'est-à-dire à $\text{ppcm}(\nu, \bar{\nu})$.

On remarque qu'un choix judicieux de ν et ν_0 simplifie considérablement les calculs.

HPTS a été utilisé avec succès pour différentes applications ; on peut citer le projet européen DIATS dont le but était l'amélioration de la sécurité et de la fluidité du trafic sur l'infrastructure autoroutière existante grâce à de nouveaux systèmes de signalisation (panneaux à messages variables, limitations de vitesse variables, contrôle du flux d'accès sur les autoroutes saturées) qu'il fallait simuler avec plusieurs milliers de véhicules. Récemment s'est ouvert à la Cité des Sciences et de l'Industrie à la Villette une application de musée virtuelle, où les modèles de comportement des entités autonomes sont décrits avec HPTS ou avec son extension pour la description de scénarios présentée dans la section suivante.

15.9 AUTONOMIE SOUS CONTRÔLE : SCÉNARIOS D'ANIMATION

Un des problèmes cruciaux qui n'a pas encore été abordé dans ces pages est le problème du contrôle des entités autonomes. Les termes peuvent paraître contradictoires, mais il s'agit en quelque sorte de scénariser l'action de l'environnement virtuel. Même si les entités autonomes sont libres, le concepteur de l'environnement virtuel peut avoir besoin de contrôler certaines actions à la fois dans l'espace et dans le temps. Par exemple, dans un simulateur de conduite dont le but est de recréer des conditions accidentogènes, il est important de pouvoir contrôler un véhicule afin qu'il arrive à un instant précis dans une intersection. D'après Kearney [Kearney e.a., 1999], le composant scénario d'un simulateur de conduite porte la responsabilité de l'orchestration des activités des agents semi-autonomes qui peuplent un environnement virtuel. Dans la plupart des expériences menées en réalité virtuelle, on cherche à pouvoir prédire les comportements des agents. Cette fonction est rendue possible par la direction des objets. Pour faciliter leur coordination, les objets autonomes sont construits avec des interfaces qui leur permettent de recevoir des instructions et de modifier leur comportement. Dans cette partie, nous allons nous intéresser aux travaux permettant la description de scénarios d'animation ou de simulation.

Les premiers travaux dans le domaine sont l'œuvre de Reynolds [Reynolds, 1982] qui définit le langage ASAS (Actor/Scriptor Animation System). Le concepteur écrit un

script qui définit l'environnement statique, gère le temps et l'enchaînement des tâches pour les acteurs. Dans le même ordre d'idées, Ridsdale et Calvert [Ridsdale e.a. , 1990] proposent en 1990 un système d'animation dite théâtrale utilisant un système expert pour répartir les tâches entre les acteurs. L'inconvénient majeur de ce type de système est bien sûr le manque d'interactivité, les scénarios sont spécifiés et compilés à l'avance. Ils ne répondent pas bien non plus au problème des entités semi-autonomes dont on ne peut pas prédire à l'avance le comportement. Dans ce cas, on cherche plutôt à coordonner les actions. Différents types de coordination peuvent être implantés :

- *orientée but* : on définit des buts à atteindre pendant l'animation, le problème le plus important consistant à planifier les actions à effectuer par les acteurs ;
- *par des règles* : on observe les acteurs évoluant dans une scène, certaines actions sont exécutées dans des configurations spécifiques ;
- *ambiante* : il y a un personnage principal (généralement l'utilisateur du système), tous les actions des autres acteurs sont déterminés en fonction de lui celui-ci, dans le but de tester ses réactions ;
- *séquence d'actions* : on détermine un ensemble d'actions et leurs dépendances temporelles. Un algorithme de planification s'occupe du reste.

Dans tous les cas, le script du scénario n'est pas exhaustif et les acteurs gardent une part d'autonomie. Un script spécifiera seulement les actions d'un petit nombre d'entités à chaque fois tandis que toutes les autres évolueront de façon autonome.

L'équipe du simulateur IDS (Iowa Driving Simulator) a proposé un modèle de description de comportement HCSM similaire à HPTS décrit précédemment. Ces deux modèles ont été étendus conjointement pour la description de scénarios de simulation [Kearney e.a., 1999, Donikian e.a., 1999]. L'université d'Iowa, à partir de travaux plus anciens [Ahmad e.a., 1994, Cremer e.a. , 1995], met l'accent sur le parallèle entre le script de cinéma et celui de la simulation en créant des chefs d'orchestre, sorte de metteurs en scène de simulation. Devillers [Donikian e.a., 1999] propose une extension de HPTS associée à une surcouche du langage de description pour proposer un langage de description de scénarios de simulation prenant en compte les contraintes temporelles. La sémantique du langage est exprimée en HPTS par l'intermédiaire de logique des intervalles définie par Allen [Allen, 1981].

15.10 CAS PARTICULIER : LE MOTION CAPTURE

De même qu'on peut construire des modèles 3D à partir d'objets qu'on va numériser, on peut effectuer des acquisitions de mouvement à partir de mouvement réels et rejouer ces mouvements. A priori quoi de plus naturel que cette technique baptisée *motion capture* (qu'on retrouve en anglais sous l'abréviation MoCap) ?

Les systèmes de Motion Capture comprennent généralement un ensemble de marqueurs qui permettent de repérer les positions de points particuliers d'un élément mobile, un ensemble de capteurs qui détectent ces positions, un logiciel de calibrage et de traitement des données obtenus. Un exemple commercialisé par la société Vicon est présenté sur les photos de la figure 15.13 : la photo de gauche montre une jeune femme dont le corps est couvert de marqueurs plastique dont la signature infrarouge est spécifique. La photo de droite montre une caméra infrarouge sensible à la signature de ces marqueurs. Un simple traitement par seuillage de l'image infrarouge permet de détecter les positions des marqueurs. Le système se compose de plusieurs caméras (jusqu'à

24 pour le système présenté). Une fois les caméras calibrées, il est possible de retrouver la position dans l'espace 3D des marqueurs.



Figure 15.13 : Le système d'acquisition de mouvement commercialisé par la société Vicon, copyright photos Vicon

Le processus d'acquisition comprend donc les étapes suivantes :

- préparation du volume d'acquisition : les caméras doivent être disposées de façon à ce qu'un maximum de caméras perçoivent le plus de marqueurs possibles lors des mouvements. Les acquisitions doivent avoir lieu dans des endroits relativement mats pour éviter les reflets ;
- calibrage du système : il s'agit de repérer les positions et orientations des caméras dans la scène, ainsi que leurs caractéristiques respectives. Cette phase est très importante puisque la précision obtenue sera le maximum de précision atteignable par les acquisitions à faire ;
- positionnement des marqueurs : les marqueurs ne doivent pas être très éloignés les uns des autres pour ne pas les confondre. Il faut les fixer sur la source de mouvement. Lors d'une acquisition de mouvement humain (le plus fréquent), il faut réussir à fixer les marqueurs sans qu'ils se déplacent, mais aussi sans que la personne qui les porte soit gênée dans ses mouvements. En dernier lieu, pour pouvoir rejouer ces mouvements en utilisant un minimum de données, il est intéressant de positionner les marqueurs sur les points d'articulation ; ainsi on peut diminuer le nombre de paramètres ;
- repérage des marqueurs : chaque marqueur est repéré sur chaque image, évitant ainsi au logiciel un calcul d'appariement. Cette étape peut être fastidieuse si le nombre de caméras est important ;
- acquisition de la scène : le sujet effectue le mouvement, tandis que les caméras enregistrent la scène ;
- traitement des images : les images sont seuillées, les coordonnées des points correspondant aux marqueurs sont extraites. Chaque point est associé à un marqueur en utilisant des algorithmes de suivi. Cette étape est sujette à erreur, puisqu'en raison des mouvements un marqueur peut apparaître ou disparaître d'une image ;
- reconstruction 3D : chaque marqueur est associé à un certain nombre d'images, la position des caméras étant connue, on peut recalculer aisément la position 3D du

marqueur. Le système est généralement sur-dimensionné (deux images suffisent à calculer la position d'un marqueur), on utilise des algorithmes de moindres carrés ;

- conversion des positions 3D en paramètres du système : les systèmes de motion capture actuels permettent de traiter des squelettes et donc de retrouver automatiquement les coordonnées articulaires du sujet.

Les coordonnées articulaires obtenues sont ensuite importées dans les grands logiciels d'animation (3Ds MAX par exemple) pour y être utilisées. Le Motion Capture a des applications en biomécanique (il existe par exemple un laboratoire pour l'analyse du geste sportif dont le but est l'amélioration de l'efficacité des gestes des sportifs de haut-niveau), dans les jeux vidéo ou le cinéma (il donne ainsi des mouvements réalistes aux personnages synthétiques), mais il a aussi donné naissance à une toute nouvelle génération de héros et d'héroïnes virtuels, dont Cléo, conçue par Médialab filiale de Canal+ a été un précurseur : Cléo était une présentatrice entièrement synthétique dont les mouvements étaient régis par une vraie actrice bardée de capteurs. L'intérêt du motion capture ici est que l'émission est réalisée en direct.

Cléo fut la première animatrice entièrement virtuelle, depuis d'autres personnages sont venus au monde, à commencer par une vedette des jeux vidéo, la fameuse Lara Croft. Bien qu'elle n'existe pas, on trouve sur le Web des sites entiers à sa dévotion, elle a même fait de la publicité pour une marque de voitures. Mieux encore, pour lui faire tourner un film, on n'a pas hésité à lui trouver un clone réel ! Le cas des jeux de Lara Croft permet de mettre en évidence les limitations du motion capture :

- Lara Croft est par définition unique, ses mensurations aussi. De ce fait, les mouvements de Lara Croft ne peuvent pas s'adapter au corps de tout un chacun : ils ne sont pas appropriés, ni d'un point de vue esthétique, ni d'un point de vue biomécanique. La question ouverte est la suivante : comment à partir d'un mouvement acquis peut-on généraliser ce mouvement à un ensemble d'individus de mensurations différentes ? c'est à cette question qu'essaie de répondre le projet MOCA⁶ ;
- si on a acquis deux mouvements utiles, par exemple la marche et la course pour Lara Croft, comment peut-on passer de l'un à l'autre ? on peut étendre la question à comment enchaîner deux fois le même mouvement ? la réponse passe actuellement par l'adoption d'une posture neutre (position de repos). F. Multon [Multon, 1998] propose une interpolation entre deux mouvements ;
- indépendamment des paramètres physiques comme la taille et le poids, il faudrait aussi pouvoir identifier dans des séquences animées l'influence de la fatigue par exemple. A moins d'acquérir des quantités énormes de données, les animations actuelles manquent cruellement de diversité, notamment les simulations des jeux de football, où tous les joueurs auront le même gabarit, et effectueront le même superbe geste sans se préoccuper ni de leur état de forme, ni même de la position exacte du ballon.

En résumé, le motion capture est une technique ambitieuse, qui permet de répondre à un certain nombre de besoins de l'animation, mais qui pose des problèmes spécifiques.

15.11 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Ahmad e.a., 1994] O. Ahmad, J. Cremer, J. Kearney, P. Willemsen, e.a. . Hierarchical, concurrent state machines for behavior modeling and scenario control.

⁶projet ESPRIT IV de la Communauté européenne, dont les partenaires sont le LIG de l'EPFL, Art & Magic, Actisystem, BITS et Duran. Pour plus d'informations, voir [http3].

- In *Conference on AI, Planning, and Simulation in High Autonomy Systems* (1994), Gainesville, Florida, USA.
- [Allen, 1981] J. F. Allen. An interval based representation of temporal knowledge. In *Proceedings of the seventh International Joint Conference on Artificial Intelligence* (1981), pages 221–226.
- [Arnaldi, 1988] B. Arnaldi (1988). *Conception du noyau d'un système d'animations de scènes tridimensionnelles intégrant les lois de la mécanique*. Thèse de doctorat, Université de Rennes I.
- [Arnaldi, 1994] B. Arnaldi (1994). *Modèles physiques pour l'animation*. Habilitation à diriger des recherches, Université de Rennes I, Rennes.
- [Arnaldi e.a. , 1992] B. Arnaldi e.a. . Vehicle simulation versus vehicle animation. In *3rd Eurographics Workshop on Animation and Simulation* (1992), Cambridge, Royaume-Uni.
- [Badler, 1998] N. Badler. Virtual humans, Computer Animation'98. In *The Virtual Humans World : Tutorial Notes* (1998), Philadelphie, USA.
- [Badler e.a., 1997] N. I. Badler, B. D. Reich, e.a. . Towards personalities for animated agents with reactive and planning behaviors. *Lecture Notes in Artificial Intelligence, Creating Personalities for synthetic actors*, (1195) :43–57 (1997).
- [Bamberger, 1981] Y. Bamberger. *Mécanique de l'ingénieur 1 : systèmes de corps rigides*, volume 1. Hermann, 293, rue Lecourbe 75015 Paris (1981).
- [Blaise, 1991] F. Blaise (1991). *Simulation du parallélisme dans la croissance des plantes et application*. Thèse de doctorat, Université Louis Pasteur, Strasbourg.
- [Booth e.a., 1993] M. Booth, J. Cremer, e.a. . Scenario control for real-time driving simulation. In *Fourth Eurographics Workshop on Animation and Simulation* (1993), pages 103–119, Barcelone, Espagne.
- [Boulic e.a., 1990] R. Boulic, N. Magnenat-Thalmann, e.a. . A global human walking model with real-time kinematic personification. *Visual Computer*, 6 (6) :344–358 (1990).
- [Boulic e.a. , 1996] R. Boulic e.a. (1996). Hierarchical kinematic behaviors for complex articulated figures. In D. Thalmann e.a. , editors, *Advanced Interactive Animation*, chapter 3. Prentice-Hall Europe.
- [Bouron, 1993] T. Bouron (1993). *Structures de communication et d'organisation pour la coopération dans un univers multi-agents*. Thèse de doctorat, Université de Paris VI, Paris.
- [Cozot, 1996a] R. Cozot (1996a). *Environnement de simulation des systèmes physiques*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Cozot, 1996b] R. Cozot. From multibody systems modelling to distributed real-time simulation. In ACM, editor, *American Simulation Symposium* (1996b), New Orleans, USA.
- [Cozot e.a. , 1997] R. Cozot e.a. . Fast simulation models for an effective aircraft training. In *Simtect'97 Advancing Simulation Technology* (1997), Canberra, Australie.
- [Cremer e.a. , 1995] J. Cremer e.a. . Hcsm : a framework for behavior and scenario control in virtual environments. *ACM Transactions on Modeling and Computer Simulation*, 5(3) :242–267 (1995).
- [Donikian, 1994] S. Donikian. Les modèles comportementaux pour la génération du mouvements d'objets dans une scène. *Revue Internationale de CFAO et d'Infographie*, 9(6) :847–871 (1994).

- [Donikian e.a., 1999] S. Donikian, F. Devillers, e.a. . The kernel of a scenario language for animation and simulation. In N. Magnenat-Thalmann e.a. , editors, *Eurographics Workshop for Computer Animation and Simulation* (1999), Milan, Italie. Eurographics.
- [Dubreuil e.a. , 1997] N. Dubreuil e.a. . Facial animation. In *Computer Animation'97* (1997), pages 98–109, Genève, Suisse.
- [Escher e.a., 1998] M. Escher, I. Pandzic, e.a. . Facial deformations for mpeg-4. In *Computer Animation'98* (1998), pages 56–62, Philadelphie, USA. IEEE Computer Society.
- [Essa e.a., 1997] I. Essa, S. Bassu, T. Darrell, e.a. . Modeling, tracking and interactive animation of faces and heads using input from video. In *Computer Animation'97* (1997), pages 68–79, Genève, Suisse.
- [Fua e.a., 1998] P. Fua, A. Gruen, R. Plaenkers, N. D'apuzzo, e.a. . Human body modeling and motion analysis from video sequences. In *International Symposium on Real Time Imaging and Dynamic Analysis* (1998), Hakodate, Japon.
- [Fua e.a. , 1997] P. Fua e.a. . Fitting sophisticated facial animation models to image data. In *Optical 3-D Measurement Techniques Conference* (1997), Zurich, Suisse.
- [Gibson, 1986] J. J. Gibson. *The ecological approach to visual perception*. Lawrence Erlbaum Associates, Hillsdale, USA (1986).
- [Giroux, 1993] S. Giroux (1993). *Agents et systèmes, une nécessaire unité*. Thèse de doctorat, Faculté des arts et des sciences, Université de Montréal, Montréal, Canada.
- [Granieri e.a., 1995] J. P. Granieri, W. Beckett, B. D. Reich, J. Crabtree, e.a. . Behavioral control for real-time simulated agents. In *Symposium on interactive 3D Graphics* (1995), pages 173–180, Monterey, USA. ACM Computer Press.
- [Hégron e.a. , 1992] G. Hégron e.a. . *Computer Animation : Motion and Deformation Control*. Eurographics'92 Tutorial Notes, Eurographics Technical Series, Cambridge, Grande-Bretagne (1992).
- [Hodgins, 1996] J. Hodgins. Three-dimensional human running. In *IEEE Conference on Robotics and Automation ICRA'96* (1996).
- [Jaeger e.a. , 1992] M. Jaeger e.a. . Basic concepts of computer simulation of plant growth. *J. Biosci*, 17 (3) :275–291 (1992).
- [Kearney e.a., 1999] J. Kearney, P. Willemsen, S. Donikian, e.a. . Scenario languages for driving simulation. In *Driving Simulation Conference, DSC'99* (1999), Paris, France.
- [Lee e.a., 1989] P. Lee, C. Phillips, E. Otani, e.a. . The jack interactive human model. In *First Annual Symposium on Mechanical Design in a Concurrent Engineering*, volume 1 (1989), pages 179–198, Iowa City, USA.
- [Lee e.a., 1995] Y. Lee, D. Terzopoulos, e.a. . Realistic modeling for facial animation. In R. Cook, editor, *SIGGRAPH'95* (1995), pages 55–62, Los Angeles, USA. ACM Computer Press.
- [Lindemayer, 1968] A. Lindemayer. Mathematical models for cellular interaction in development. *Journal of Theoretical Biology*, 18 :280–315 (1968).
- [Lord e.a. , 1994] R. G. Lord e.a. . Moving from cognition to action : A control theory perspective. *Applied Psychology : an international review*, 43 (3) :335–398 (1994).
- [Magnenat-Thalmann e.a. , 1984] N. Magnenat-Thalmann e.a. , editors. *Artificial Life and Virtual Reality*. John Wiley, Chichester (1984).

- [Moreau, 1998] G. Moreau (1998). *Modélisation du comportement pour la simulation interactive : application au trafic routier multimodal*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Moreau e.a. , 1998] G. Moreau e.a. . From psychological and real-time interaction requirements to behavioural simulation. In B. Arnaldi e.a. , editors, *Eurographics Workshop on Computer Animation and Simulation* (1998), Lisbonne, Portugal.
- [Mouli, 1994] R. Mouli (1994). *Un modèle d'animation comportementale basé sur le concept de personnage*. Thèse de doctorat, Université Paul Sabatier, Toulouse.
- [Mouli e.a., 1992] R. Mouli, Y. Duthen, e.a. . Un modèle de simulation comportementale orienté objet pour l'animation intelligente. In *IMRV92 Interface des Mondes Réels et Virtuels* (1992), pages 111–121, Montpellier.
- [Multon, 1996] F. Multon. Animation d'humanoïdes synthétiques par un modèle biomécanique. In *4èmes Journées de l'Association Française Graphique* (1996), Dijon.
- [Multon, 1998] F. Multon (1998). *Contrôle du mouvement des humanoïdes de synthèse*. Thèse de doctorat, Université de Rennes I, Rennes.
- [Multon e.a., 1999] F. Multon, L. France, M.-P. Cani-Gascuel, e.a. . Computer animation of human walking : a survey. *The Journal of Visualization and Computer Animation*, 10(1) :39–54 (1999).
- [Newell, 1990] A. Newell. *Unified Theories of Cognition*. Harvard University Press (1990).
- [Ngo e.a. , 1993] J. T. Ngo e.a. . Spacetime constraints revisited. In J. T. Kajiya, editor, *SIGGRAPH'93* (1993), pages 343–350, Anaheim, USA. ACM Computer Press.
- [Noser, 1997] H. Noser (1997). *A behavioral animation system based on L-systems and Synthetic Sensors for Actors*. Thèse de doctorat, École Polytechnique Fédérale de Lausanne, Lausanne, Suisse.
- [Noser e.a. , 1997] H. Noser e.a. . Sensor based synthetic actors in a tennis game simulation. In *Computer Graphics International* (1997), pages 189–198, Hasselt, Belgique.
- [Ostermann, 1998] J. Ostermann. Animation of synthetic faces in mpeg-4. In *Computer Animation'98* (1998), pages 49–55, Philadelphie, USA. IEEE Computer Society.
- [Phillips e.a. , 1988] C. Phillips e.a. . Jack : a toolkit for manipulating articulated figures. In *ACM/SIGGRAPH Symposium on user interface software* (1988), Banff, Canada.
- [Picinbono, 2001] G. Picinbono (2001). *Modèles géométriques et physiques pour la simulation d'interventions chirurgicales*. Thèse de doctorat, Université de Nice Sophia-Antipolis.
- [Prusinkiewicz e.a., 1993] P. Prusinkiewicz, M. S. Hammel, e.a. . Animation of plant development. In J. T. Kajiya, editor, *SIGGRAPH'93* (1993), pages 351–360, Anaheim, USA. ACM Computer Press.
- [Prusinkiewicz e.a., 1994] P. Prusinkiewicz, M. James, e.a. . Synthetic topiary. In *SIGGRAPH'94, Annual Conference Series* (1994), pages 351–358, Orlando, USA. ACM SIGGRAPH.
- [Rainjonneau, 1992] S. Rainjonneau (1992). *Un modèle orienté-objet pour la simulation comportementale*. Thèse de doctorat, Université Paul Sabatier, Toulouse.
- [Reich, 1997] B. D. Reich (1997). *An architecture for behavioral locomotion*. Thèse de doctorat, Université de Pennsylvannie, Philadelphie, USA.

- [Reynolds, 1982] C. Reynolds. Computer animation with scripts and actors. *Computer Graphics (Proceedings of SIGGRAPH'82)*, 16(3) :289–296 (1982).
- [Reynolds, 1987] C. W. Reynolds. Flocks, herds, and schools : A distributed behavioral model. *Computer Graphics (Proceedings of SIGGRAPH'87)*, 21(4) :25–34 (1987), Anaheim, USA.
- [Ridsdale e.a. , 1990] G. Ridsdale e.a. . Animating microworlds from scripts and relational constraints. In N. Magnenat-Thalmann e.a. , editors, *Computer Animation'90* (1990), pages 107–118, Genève, Suisse. Springer Verlag.
- [Thalmann, 1998] D. Thalmann. Interactive perceptive virtual humans. In *Computer Animation'98* (1998), Philadelphia, USA.
- [van de Panne e.a. , 1993] M. van de Panne e.a. . Sensor-actuator networks. In J. T. Kajiya, editor, *Computer Graphics (SIGGRAPH'93 Proceedings)* (1993), pages 335–342, Anaheim, USA. ACM Computer Press.
- [Watt, 2000] A. Watt. *3D Computer Graphics*. Addison-Wesley (2000).
- [Webber e.a. , 1995] B. Webber e.a. . Animation through reactions, transition nets and plans. In *International Workshop on Human Interface Technology* (1995), Aizu, Japon.
- [Widyanto e.a., 1991] T. Widyanto, A. Marriot, e.a. . Applying a virtual perception model to behavioral animation system. In *Eurographics Workshop on Animation and Simulation* (1991), pages 89–98, Vienne, Autriche.
- [Wilhelms e.a. , 1990] J. Wilhelms e.a. . a notion for interactive behavioral animation control. *IEEE Computer Graphics and Applications*, 10 (3) :14–22 (1990).
- [Zeltzer, 1991] D. Zeltzer (1991). Task-level graphical simulation : abstraction, representation and control. In N. I. Badler, B. A. Barsky, e.a. , editors, *Making them move : Mechanics, control and animation of articulated figures*, pages 3–33. Morgan Kaufmann.
- [Zhao e.a., 1996] X. Zhao, D. Tolani, B. Ting, e.a. . Simulating human movements using optimal control. In *Computer Animation and Simulation'96* (1996), pages 109–120.

Adresses des sites Web :

- [http1] <http://www.cirad.fr/produits/amap/amap.html>
- [http2] <http://www.cis.upenn.edu/~hms/jack.html>
- [http3] <http://ligwww.epfl.ch/~thalmann/MOCA.html>

Septième partie

Aspects généraux de la réalité virtuelle

16 LES EFFETS SUR L'HOMME DES ENVIRONNEMENTS

Jean-Paul Papin

Il a été dit à plusieurs reprises, dans les chapitres précédents, qu'une immersion totalement virtuelle était impossible. L'espace d'immersion comprend toujours des environnements réels et virtuels. Chacun de ces environnements peut entraîner des effets néfastes. Le but de ce chapitre est de montrer quelques effets indésirables directs (sur chaque organe) et généraux (sur l'ensemble de l'organisme), liés à certaines valeurs des environnements naturels et artificiels, en particulier virtuel.

16.1 LES EFFETS SUR L'HOMME DES ENVIRONNEMENTS NATURELS

16.1.1 LA LUMIÈRE

16.1.1.1 L'éblouissement

Toute source brillante placée dans le champ visuel crée dans son voisinage une chute des sensibilités de l'œil. Cette source brillante agit comme un voile brillant s'interposant entre l'œil et la tâche visuelle. La brillance du voile s'exprime par la formule $KE/\theta n$ où E est l'éclairement produit sur la pupille, θ l'angle de la ligne du regard avec la direction de la source, K et n des constantes. L'éblouissement est donc d'autant moindre que la source est éloignée de l'axe du regard. C'est pourquoi il y a intérêt à mettre les sources lumineuses le plus haut possible afin qu'elles aient moins de chance de se trouver dans le champ des sujets. De même si la source est issue de tubes l'éblouissement sera à sa plus petite valeur lorsque les tubes seront disposés dans la même direction que le regard. Les sources vives (flash de phare, laser, arc électrique) sont éblouissantes et la seule protection efficace reste l'interposition d'un écran protecteur entre la source et l'œil.

16.1.1.2 Les brûlures

Des sources comme le laser peuvent brûler la rétine. La lumière par son effet calorifique peut entraîner des brûlures au niveau de la peau.

16.1.2 LE BRUIT

Le bruit est généralement défini comme un son parasite ou n'ayant aucun contenu informationnel. Cette définition n'est pas satisfaisante, il est préférable de dire qu'il y a bruit quand il y a gêne. On distingue, les bruits stables, fluctuants, intermittents et impulsifs, ces derniers, même avec des énergies faibles peuvent provoquer des lésions cochléaires importantes conduisant à des surdités.

16.1.2.1 Effets auditifs du bruit

- l'effet de masque. On dit qu'il y a phénomène de masque quand un son ou une bande sonore cessent d'être perçus si l'on fait entendre simultanément un autre son ou une autre bande sonore ou un bruit complexe. Dans le cadre de l'intelligibilité de la parole, le masque de la voix par le bruit peut altérer la compréhension ;
- la fatigue auditive se caractérise par une élévation temporaire du seuil de perception, on parle de TT (Temporary Threshold Shift). La plage la plus sensible se situe entre 4000 et 6000 Hz pour des bruits supérieurs à 80 dB (Figure 16.1). Cependant le temps d'exposition joue un rôle important ainsi que la forme du signal sonore. En particulier s'il s'agit d'un bruit impulsif caractérisé par une attaque très brusque avec une montée au niveau maximal se faisant en un temps inférieur au temps de réaction des muscles de l'oreille moyenne (coup de feu), le temps d'exposition est réduit à sa plus simple expression. Une fatigue, voire un traumatisme définitif, peut s'installer. Dans l'état actuel de nos connaissances il est difficile d'établir avec certitudes les liens qui existent entre la fatigue auditive et la surdité. La sensibilité interindividuelle est fluctuante, le temps de récupération aussi (en général cette récupération se fait en 20 minutes).

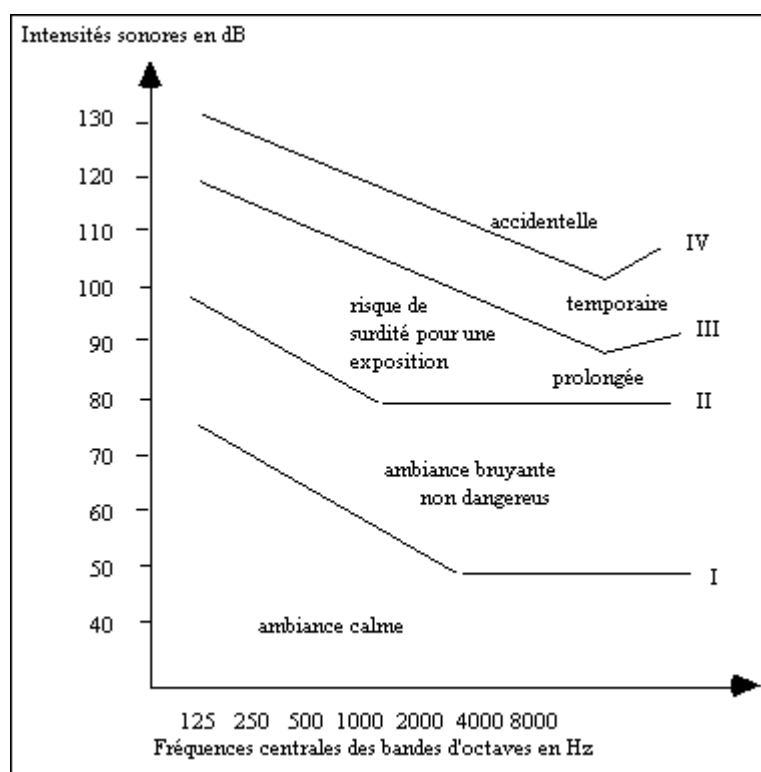


Figure 16.1 : Risques de traumatisme sonore (D'après Wisner 1967)

Mais même si le bruit n'est pas dangereux, il peut être source de gêne subjective et surtout objective en s'accompagnant d'une diminution des performances de travail.

16.1.2.2 Effets psychologiques du bruit

Ils sont de deux ordres affectifs et intellectuels. Sur le plan psychomoteur et intellectuel un bruit bref, intense et inattendu perturbe la performance. Cependant dans des tâches monotones de surveillance un bruit peut relever le niveau de vigilance et donc être favorable. Un bruit continu non significatif entraîne une baisse de performance sur des tâches complexes mettant en jeu la mémoire immédiate et la prise de décision. Sur le plan affectif c'est une sensation désagréable qui entraîne une gêne parfois insupportable pouvant se traduire par une perte de sommeil ou une augmentation de l'agressivité. La gêne ressentie est fonction en particulier de l'âge (sédatif pour les nourrissons de 3 ou 4 mois, très agressifs pour les sujets âgés), du lieu (90 dB sont plus gênant au domicile qu'à l'usine) et de la personnalité (les névrotiques sont plus sensibles).

16.1.2.3 Effets généraux du bruit sur l'organisme

Le bruit surtout s'il est intense peut aussi entraîner des effets secondaires nocifs du type perte de l'équilibre et de manière indirecte des nausées, des vertiges. Pour des ambiances bruyantes continues, il a été signalé, par exemple, des pertes d'appétit, des insomnies, des amaigrissements. Cependant il est difficile de faire la part réel du bruit et de l'état de stress lié à certaines situations de vie ou de travail.

16.1.2.4 Protection contre le bruit

La protection peut se faire :

- par réduction des bruits à la source ;
- par isolation acoustique de l'habitable ou du local ;
- par le port de protecteurs d'oreilles, moyen, qui s'il protège, diminue d'autant les possibilités de communication verbales ;
- enfin, Il existe des normes particulières à respecter dans les locaux en fonction du type d'activité. Par exemple :
 - salle de concert : 21 à 30 dBA
 - chambre, hôpitaux : 34 à 47 dBA
 - bureaux : 38 à 47 dBA
 - pools de secrétaires : 47 à 56 dBA
 - cuisines : 52 à 61 dBA
 - ateliers : 56 à 66 dBA
 - pour les locaux ou les conversations : 66 à 80 dBA

16.1.3 LA CHALEUR

Les effets de la chaleur se traduisent soit par une sensation de douleur qui apparaît lorsque la température de la peau atteint 45° C, soit par une sensation d'inconfort, voire de malaise par stockage.

La douleur d'origine thermique apparaît lorsque la température de la peau atteint 45 °C. La tolérance dépendra du temps mis par la peau pour atteindre le seuil de douleur. Le diagramme de Webb 1949 [Ministère de la défense, 1981] est plus sévère que d'autres valeurs rapportées, comme celles concernant une boulangère qui en 1760 a supporté 142 °C pendant 12 minutes ou celles de Blockey et Taylor en 1949 [Ministère de la défense, 1981] montrant que l'on peut rester 30 minutes à 115 °C. La température de la peau varie dans ces conditions de 38 °C à 42 °C.

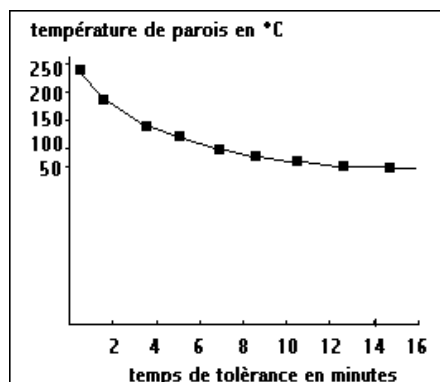


Figure 16.2 : Temps de tolérance à la chaleur radiante de parois (d'après Webb 1963)

Le stockage s'effectue lorsqu'un sujet est placé dans une ambiance thermique assez sévère pour dépasser les limites de thermorégulation. La sensation d'inconfort à la chaleur apparaît lorsque l'organisme stocke 375 Kilojoules pour un homme de 70 Kg et de 1,75 m. Ceci correspond à une augmentation de la température centrale de 1,5 °C. La zone dangereuse se situe vers 600 Kilojoules et s'accompagne d'une élévation de 2,5 °C. A cette température de 39,5 °C, la fréquence cardiaque est de 180 battements par minute. Dans ces conditions, 50 % des sujets ne sont plus efficaces car incommodés totalement ou partiellement par la chaleur. Le risque mortel 100 % est obtenu pour un stockage de 1000 Kilojoules entraînant des températures profondes de l'organisme de 41 °C. Pour obtenir de tels états de multiples facteurs interviennent, certes la température de l'air ou la radiante, l'humidité (Figure 16.3) et la vitesse du vent, mais aussi l'âge, le sexe, l'activité, l'adaptation, l'entraînement, les vêtements et enfin la durée d'exposition. La survie dépend aussi de la durée qui s'écoule entre l'apparition des symptômes et les soins. C'est ainsi qu'un coup de chaleur avec déshydratation entraîne la mort dans 70 % des cas en deux heures.

Pour l'O.M.S il ne faut jamais dépasser 38 °C de température centrale corporelle pendant 8 H par jour. Ces chiffres expliquent que les bons indices de surveillance de l'homme pour la contrainte thermique sont la température profonde, la fréquence cardiaque et la perte pondérale. Cette dernière rend compte de la perte hydrique. Ces paramètres servent à évaluer l'astreinte. Les autres mesures d'ordre climatiques évaluent la contrainte. Il faut signaler que les différences individuelles concernant la tolérance sont importantes et que les effets néfastes peuvent apparaître plus ou moins rapidement chez des sujets soumis à une ambiance chaude.

Sur le plan des performances psychomotrices ou intellectuelles la chaleur va faire varier celles-ci en agissant sur le niveau d'activation du système nerveux central. Les variations de température ont des effets différents en fonction du niveau d'activation

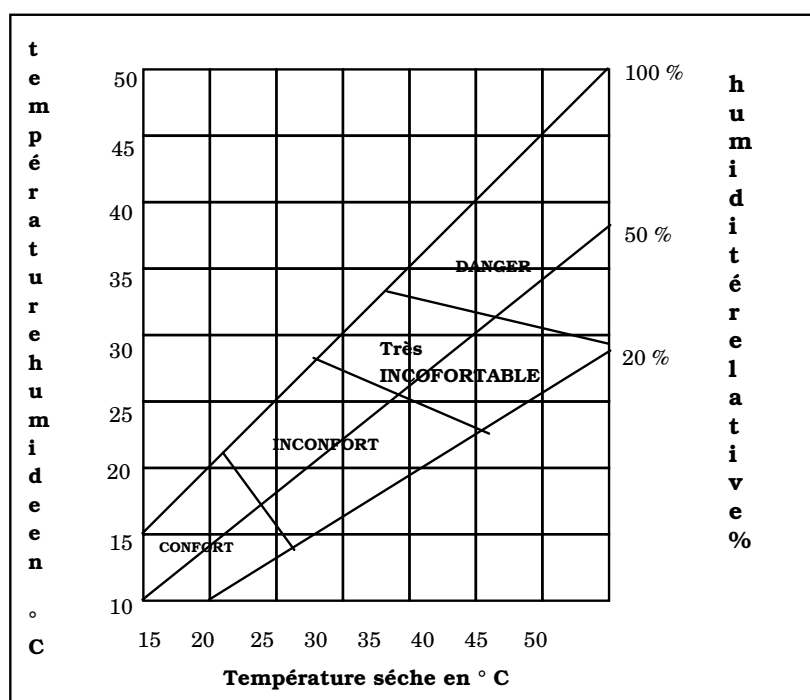


Figure 16.3 : Zones de confort, d'inconfort et de danger (D'après Nevins et coll. 1966) [Ministère de la défense, 1981]

antérieur. Si ce niveau correspondait à l'optimal pour une tâche donnée quel que soit le sens de variation de la température, la performance va être détériorée. Si le niveau d'activation était trop bas, une chaleur modérée le détériore encore plus mais une chaleur intense peut pendant un certain temps l'améliorer. L'effet est inverse si ce niveau était trop élevé. Ceci est vrai pour les tâches de réactivité ou de rapidité. Pour les tâches de précision, la chaleur entraîne toujours une détérioration dès le premier quart d'heure. Pour le travail physique, la chaleur, surtout si elle est humide, entraîne une baisse de production pouvant atteindre, comme le montre Wyndham [Ministère de la défense, 1981], 90 % lorsque la température dans l'ambiance humide d'une mine passe de 27 °C à 36 °C.

16.1.4 LE FROID

A côté des effets du froid sur la peau qui se traduisent par des gelures, comparables aux brûlures par le feu, il y a des effets généraux du froid sur l'être humain qui se traduisent par un refroidissement du corps. Un refroidissement de la température centrale à 24°C donne une symptomatologie de mort apparente et malgré la réanimation conduit presque toujours à la mort. Il existe une relation entre la baisse de la température corporelle et les symptômes qui peuvent apparaître (Figure 16.4).

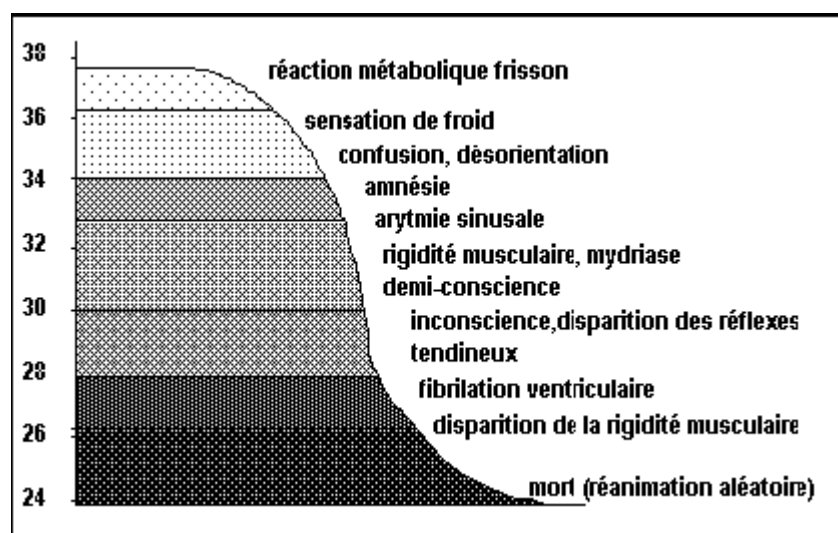


Figure 16.4 : Baisse de la température centrale et symptômes (d'après Goden 1973) [Ministère de la défense, 1981]

Il faut cependant distinguer les effets du froid dans l'air avec du vent et dans l'eau. Dans l'air le normogramme en fonction des valeurs du W.CI. montre que la zone 1 peut être considérée comme confortable avec des précautions normales port de vêtements hermétiques. Dans la zone 2 il y a sensation de très froid, le port de gant devient nécessaire, et ces sensations de froids sont encore plus fortes par temps couvert. Dans la zone 3 cela devient très pénible sauf par temps ensoleillé. Dans la zone 4 il y a des risques de gelures en particulier s'il n'y a pas d'activité physique. Enfin dans la zone 4 sans gants très chauds, par exemple, les gelures se produisent en moins d'une minute

et la survie est en jeu assez rapidement (Figure 16.5).

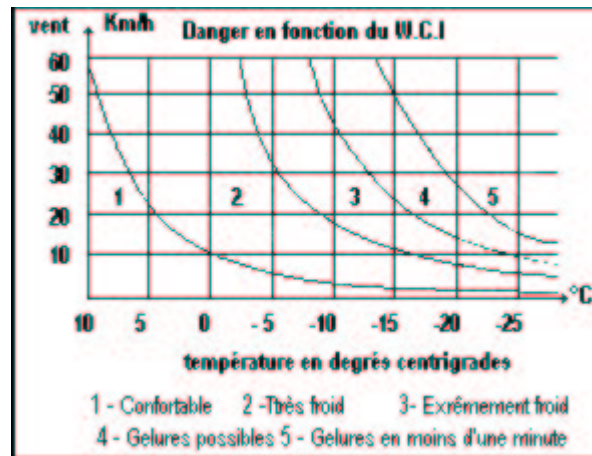


Figure 16.5 : Danger en fonction des valeurs du W.C.I (d'après Rivolier 1981)
[Scherrer, 1981]

16.1.5 LES ACCÉLÉRATIONS

16.1.5.1 Les accélérations de longues durées

Il faut considérer les positives et les négatives, les longitudinales et les transversales ; Les Gz + et - se rencontrent essentiellement en aéronautique au cours des montées et des descentes, les autres lors de déplacements et en freinages.

- pour les Gz +, outre la sensation de poids accru entraînant une difficulté pour bouger les membres (très nette à partir de 2,5 G), il apparaît à 3 G des troubles de la vision conduisant à partir de 4 ou 5 G au voile noir puis à la perte de conscience dès 5 ou 6 G. Ces chiffres sont à prendre avec prudence car il existe de grandes variations individuelles expliquant le fait que l'on pratique des tests en centrifugeuses pour la sélection des pilotes de chasse ou des cosmonautes. Ces tests permettent de plus, comme on le pratique pour les effets liés à l'hypoxie, de faire prendre, aux sujets, conscience des prodromes (signes annonciateurs) précédant les effets dangereux ;
- pour les Gz - à partir de - 3 G les effets deviennent intolérables en quelques secondes. Il y a sensation d'oppression, de gêne respiratoire et voile rouge. Ce dernier semble lié à l'engorgement du sang dans les paupières et non à une hémorragie ;
- pour les transversales les effets sévères apparaissent à partir de 5 G et sont essentiellement respiratoires, à 10 G la vision périphérique est perdue, à 15 G le voile noir apparaît.

Les positives se rencontrent au cours des catapultages, les négatives au cours des apontages ou des crashes (Figure 16.6).

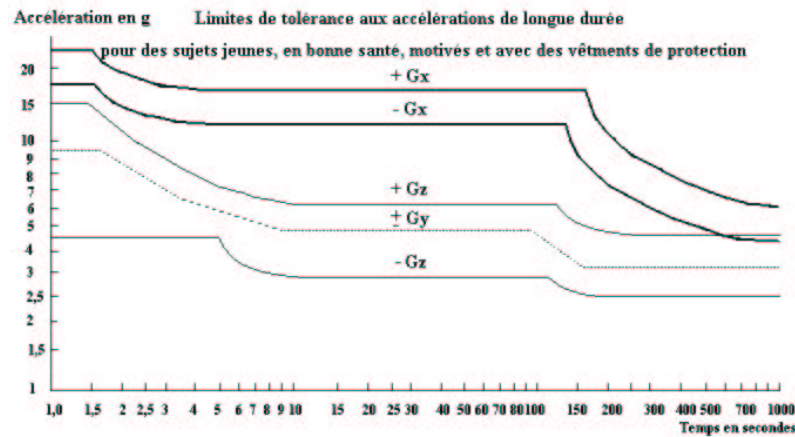


Figure 16.6 : Tolérance aux accélérations de longue durée (D'après Fraser 1966)

La protection de ces effets n'est efficacement possible que pour les $G_z +$ grâce aux combinaisons anti-g qui compriment les vaisseaux sanguins du bas du corps. Ceci permet une irrigation de la tête et donc du cerveau.

16.1.5.2 Les accélérations brèves

Elles se traduisent par des troubles mécaniques à type de fractures d'os ou d'éclatement d'organes. Leurs effets dépendent des possibilités d'amortissement et de résonance du corps. En effet, ce dernier peut être assimilé à un ensemble de masses élémentaires réunies entre elles par des ressorts et des amortisseurs.

On peut admettre qu'en dessous de 0,01 s tout est amorti. Entre 0,01 et 1 s, il y a transmission avec des phénomènes d'amplification possible si les vitesses d'installations sont dans les zones de périodes naturelles du corps (phénomène d'overshoot). Au-delà de 1 s, le corps se comporte comme un ensemble rigide (Figure 16.7 et Figure 16.8).

16.1.6 LES VIBRATIONS DES SOLIDES ET LEURS EFFETS GÉNÉRAUX SUR L'ORGANISME HUMAIN

On distingue trois types de vibrations :

- de très basses fréquences : 0 à 2 Hz ;
- de basses fréquences : 2 à 20 Hz ;
- de hautes fréquences : 20 à 1000 Hz ;

Les effets vont être liés à la résonance possible d'un organe qui a sa fréquence propre. L'organisme, composé de matière, se comporte comme un système de masses suspen-

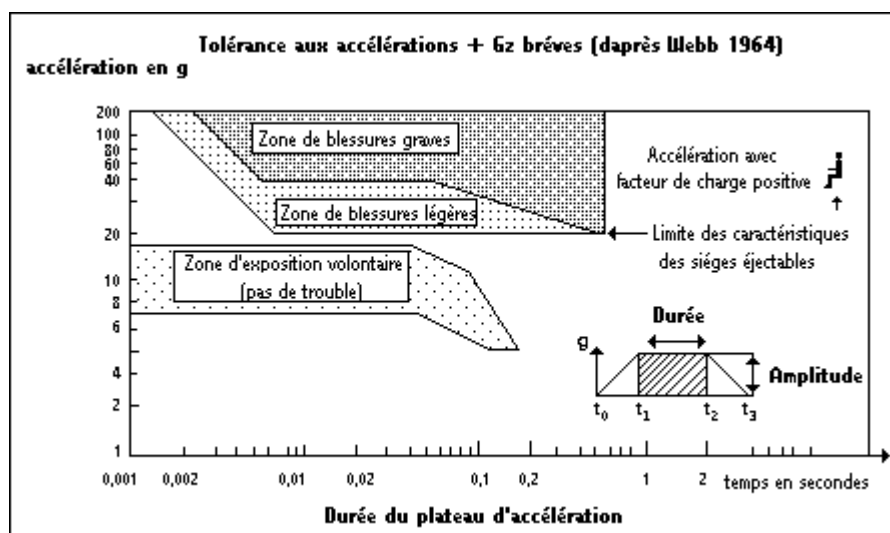


Figure 16.7: Tolérance aux accélérations + Gz de durée brève [Ministère de la défense, 1981]

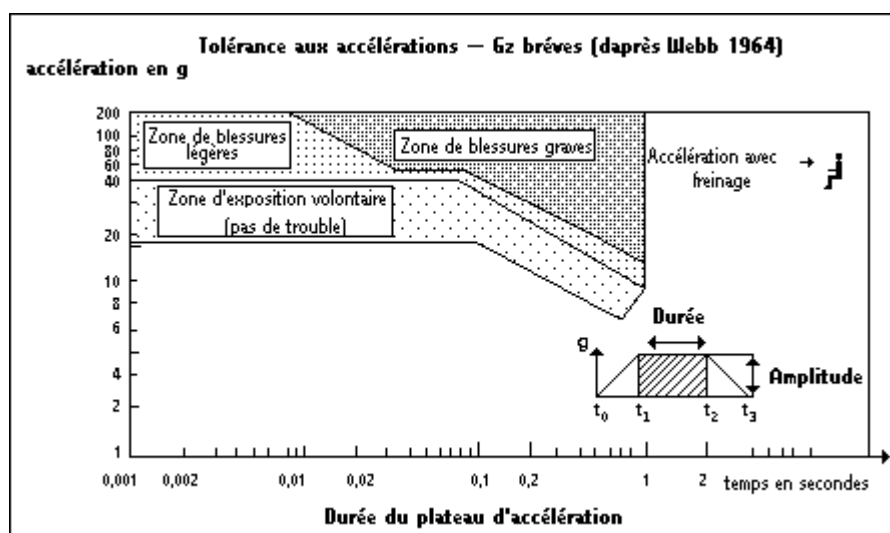


Figure 16.8: Tolérance aux accélérations - Gz de durée brève [Ministère de la défense, 1981]

dues avec la fréquence appliquée (Figure 16.9) :

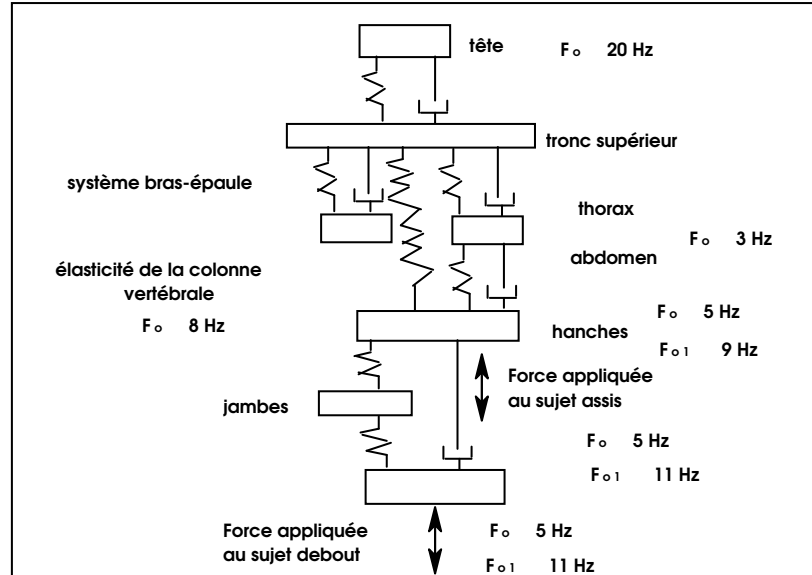


Figure 16.9: Modèle analogique de l'homme debout ou assis soumis à des vibrations verticales. F_o = fréquence naturelle ou propre (D'après COERMANN 1963) [Ministère de la défense, 1981]

16.1.6.1 Les très basses fréquences

Elles sont synonymes de cinétose ou mal du mouvement. Elles sont fréquentes dans tous les moyens de transports et dans les simulateurs. Il s'agit d'un désordre neurovégétatif avec pâleur, sueurs froides, nausées et vomissements. Tout rentre dans l'ordre en quelques minutes et au maximum en quelques heures après l'arrêt des causes ayant déclenchées le malaise. Les cinétoses liées à l'emploi de la réalité virtuelle seront traitées dans le paragraphe effets des immersions.

16.1.6.2 Les basses fréquences

Elles entraînent l'installation d'une fatigue précoce conduisant à une diminution des performances psychomotrices et à des douleurs au niveau de la colonne vertébrale. A long terme elles produisent une pathologie du rachis.

L'acuité visuelle est la plus diminuée pour 5 Hz. La ventilation pulmonaire est augmentée pour des fréquences comprises entre 4 et 12 Hz. Pour les mêmes fréquences, il peut y avoir des perturbations de l'activité cardiaque avec extrasystoles et tachycardie.

16.1.6.3 Les vibrations de hautes fréquences

Les fréquences de 20 à 40 Hz de grandes amplitudes (cas du marteau piqueur), se traduisent par une augmentation du tremblement du corps et surtout de la main qui entraînent, à long terme, des lésions ostéo-articulaires à type de décalcification, surtout au niveau du poignet et de la main. Les fréquences de 40 à 300 Hz, caractéristiques des moteurs, conduisent à des troubles vasomoteurs type le syndrome de RAYNAUD. Les fréquences supérieures à 300 Hz, dont l'amplitude est généralement de quelques millimètres, sont obtenues avec des polisseuses et s'accompagnent de troubles trophiques et sensitifs de la main. Pour se protéger, le temps d'exposition doit être limité.

16.2 LES EFFETS SUR L'HOMME DES ENVIRONNEMENTS VIRTUELS

Les effets se rencontrent sur le plan visuel, musculo-squelettique, neurovégétatif et psychologique. Une place particulière doit être faite, en s'appuyant sur les connaissances du mal des simulateurs et celui des transports, aux malaises rencontrés au cours des immersions [Regan, 1998]. Dans ce domaine seront décrits les facteurs qui peuvent jouer un rôle dans la sévérité et la gravité de la symptomatologie rencontrée. De même on donnera des exemples d'outils permettant de mettre en évidence cette pathologie.

16.2.1 LES EFFETS VISUELS

Les effets secondaires visuels peuvent avoir de sérieuses conséquences sur les actions à effectuer en environnement virtuel. Il est possible en effet de rencontrer :

- une hétérophorie (écart de position en vision binoculaire lié à des variations de l'équilibre des muscles moteurs des globes oculaires). Dans la vision normale (orthophorie), il n'y a pas de déviation des axes visuels. Dans l'hétérophorie, soit ils divergent (exophorie), soit ils convergent (ésophorie), soit l'une des lignes tend à se porter sur un plan plus élevé que l'autre (hyperphorie);
- une diplopie (vision double);
- une altération du champ visuel;
- une vision brouillée;
- une fatigue visuelle.

La survenue de maux de tête est probablement largement reliée à ces troubles de la vision. L'intensité des troubles visuels dépend en partie de la nature des systèmes optiques de présentation des informations :

- binoculaire (où chaque œil regarde un écran différent, montrant des images différentes permettant la vision stéréoscopique);
- bioculaire (où chaque œil regarde un écran différent mais qui montrent la même scène.);
- monoculaire (où les deux yeux regardent le même écran).

Il peut être admis [Mon-Williams, 1993] que les problèmes visuels rencontrés sont largement dus à la dissociation entre l'accommodation et la convergence induite par l'utilisation surtout des systèmes binoculaires. En vision normale, les yeux convergent et

accommodent pour une même distance. La vision stéréoscopique sur un écran ou dans un visiocasque a pour conséquence le phénomène de dissociation entre l'accommodation et la convergence. Il est rapporté, [Mon-Williams, 1993], que 50 % d'un échantillon de sujets immergés dans un environnement virtuel pendant 10 minutes présentent des troubles de la vision subjectifs et vérifiés par des tests ophtalmologiques conventionnels après l'immersion. En particulier l'hétérophorie que les auteurs interprètent comme la base de l'effort nécessaire à l'accommodation/ convergence. Des recherches similaires [Rushton, 1994] ont été réalisées avec des casques bioculaires plus récents et trouvé que des troubles apparaissent après un port de 30 minutes. Ces résultats entraînent les auteurs à confirmer que l'effort intense qui est demandé en terme d'accommodation/convergence avec les systèmes binoculaires est responsable de la courte durée supportable. Cet effort est moins important avec les systèmes bioculaires car c'est la même image qui est présentée aux deux yeux. Une analyse expérimentale porte sur la dissociation accommodation et convergence et les troubles qui en découlent [Howarth et Bradbry, 1994]. Pour cela les auteurs ont utilisé des systèmes visuels à base de prismes et dans deux expériences ils ont mis en évidence des hétérophories. De plus il semble que les troubles sont aggravés quand en plus il y a des variables techniques associées aux systèmes visuels de tête comme la faible illumination, le faible contraste et une distance de travail exceptionnellement faible. Ces travaux se sont poursuivis [Howarth et Costello, 1996] et montrent des changements liés au handicap du port des systèmes de poursuite de tête au cours de période d'une heure. De même [Miyake, 1994], il a été montré l'existence de changements intervenant au niveau visuel après le port pendant vingt minutes des différents systèmes de présentations visuelles de réalité virtuelle. D'autres troubles visuels liés aux visiocasques sont probablement dus à des problèmes d'alignement des images ou des lentilles. Par exemple des troubles apparaissent si les lentilles sont montées sans être parfaitement adaptées à un individu donné. Beaucoup des premières générations de visiocasque étaient à base d'écrans à cristaux liquides, séparés d'une distance fixe. Dans ce cas [Regan, 1993], il apparaît que si l'utilisateur a une distance inter pupillaire plus petite que la distance entre les deux écrans les problèmes visuels sont alors très fréquents. D'autre part si la distance inter pupillaire est plus grande que la distance entre les deux écrans les problèmes visuels semblent moins importants. Cela serait lié à l'effort plus important fourni par les yeux pour diverger que pour converger. Cependant les systèmes récents apportent des possibilités de régler la distance entre les lentilles. Le confort visuel est plus grand mais malgré tout la durée acceptable de port reste toujours courte. C'est pourquoi, il est recommandé de se servir dans la mesure du possible de lunettes à cristaux semi-transparentes et de bien réfléchir aux applications nécessitant une immersion totale et une vision stéréoscopique. Il faut aussi se souvenir que des facteurs comme la dépendance ou l'indépendance au champ visuel jouent un rôle. C'est ainsi que les dépendants du champ sont plus gênés par le port d'un casque de qualité médiocre que les sujets indépendants du champ. Cela peut s'expliquer en partie par le fait que les indépendants du champ ayant une prise d'information globale d'une scène ont moins besoins que les autres d'un bon relief pour appréhender la scène.

16.2.2 LES EFFETS MUSCULO-SQUELETTIQUES

Ils sont essentiellement dus au poids des visiocasques qui peuvent impliquer de grandes fatigues au niveau du cou. En fait cela pose problème que si le port est de longue durée. De même la conception de dispositifs par exemple pour l'ensemble des spectateurs d'une salle est un véritable défi si l'on veut avoir des possibilités de réglages pour toutes les tailles de tête. Beaucoup de modèles courants contiennent dans une armature

solide des systèmes de serrage qui sont souvent très inconfortables car il y a des points de compressions tout autour de la tête qui réduisent la circulation sanguine et donnent des sensations de chaleur désagréable. D'autres problèmes sont souvent causés par une limitation des mouvements et des efforts musculaires intenses.

16.2.3 CINÉTOSES

Dans cette partie il s'agit d'aborder le problème du mal de l'immersion qui est comparable à celui des transports ou des simulateurs. Dans tous les cas même si les situations sont différentes (l'existence de mouvements de basses fréquences dans le mal de mer et aucun mouvement dans certains simulateurs) les mêmes symptômes sont relevés. Ces symptômes sont plus ou moins tous présents, suivant les sujets et pour un même sujet suivant les différentes immersions. Dans les expériences d'immersion [Regan, 1994], il est noté l'existence de nausée, de sensation d'estomac noué, de vertiges, de désorientation, de somnolence, de transpiration et de maux de tête. Autant de symptômes qui sont retrouvés dans le mal des transports ou celui des simulateurs. Sur le plan théorique il est possible de s'appuyer sur la théorie du conflit sensoriel [Reason, 1974] qui peut s'appliquer aussi bien au mal de mer, qu'à celui de l'espace ou celui des simulateurs. D'autre part, une thèse [Draper, 1998] étudie en détail le problème de l'adaptation physiologique de l'homme aux interfaces de réalité virtuelle et montre l'incidence du réflexe vestibulo-oculaire sur le mal des simulateurs. Dans tous ces cas, la théorie du conflit sensoriel montre que les malaises ressentis se caractérisent par la transmission conflictuelle de signaux à partir du système visuel, du système vestibulaire et du système de perception de position extra-vestibulaire. L'immersion en environnement virtuel peut être vue comme induisant un conflit sensoriel entre les systèmes visuel et vestibulaire. En effet dans ce type d'environnement le mouvement qui apparaît sur l'écran peut être commandé par un dispositif manuel. Dans ces conditions le système visuel suggère des mouvements du corps, mais le système vestibulaire évoque une position statique du corps. De plus les retards de présentation des systèmes de réalité virtuelle induisent un conflit pour l'utilisateur quand le temps entre une action (un mouvement de tête par exemple) et le résultat escompté (changement de la scène visuelle) devient discernable et qu'ainsi il y a conflit entre l'attente et le réel. La théorie écologique soutient que les malaises sont liés à la stabilité posturale. Dans le même sens [Riccio, 1991] les malaises résulteraient du fait qu'il n'y a pas eu un apprentissage d'une stratégie pour maintenir une posture verticale au niveau des activités psychomotrices de la tête du torse, des hanches, des genoux, des chevilles car l'anticipation est difficile en simulateur. Cette idée d'impossibilité d'anticipation des activités et des événements [Vaden, 1996] est une cause d'aggravation des malaises.

Les utilisateurs ont noté une diminution significative de symptômes en rapport avec la répétition des immersions. Cela a été mis en évidence expérimentalement [Regan, 1994a], il y a une diminution spectaculaire des symptômes au cours de quatre immersions de 20 minutes. Un peu plus de la moitié des sujets qui avait présenté des effets secondaires au cours de la première immersion n'en avait plus à la fin de la quatrième ou notait que les nausées devenaient très faibles. Il peut s'agir d'une simple accoutumance à un environnement donné, mais cette adaptation est-elle tenace et intéresse-t-elle d'autres environnements virtuels ? Ces questions méritent d'être approfondies.

Des méthodes pour réduire les effets secondaires particulièrement les nausées ont été étudiées. Il a, par exemple, été utilisé avec succès des drogues anti mal de mer. Dans une expérimentation avec un groupe placebo et un groupe médicamenté [Regan, 1994b].

L'administration de 300 micros grammes de « Hyoscine hydrobromide » 40 minutes avant 20 minutes d'immersion en réalité virtuelle entraîne une décroissance significative des malaises ressentis. Il est apparu que 25 % des sujets du groupe placebo faisaient l'expérience de nausée comparés à 5 % des sujets du groupe prenant de l'hyoscine. De plus l'hyoscine avait non seulement un effet sur les nausées, mais aussi sur les autres symptômes comme la sensation d'estomac noué et les maux de tête qui sont communément observés durant les immersions.

Enfin, plusieurs études ont analysé ce que pouvait apporter la possibilité de pouvoir naturellement se déplacer dans les environnements virtuels. Certaines de ces recherches ont été motivées par le désir de réduire les conflits sensoriels et par-là les effets secondaires tandis que d'autres étaient destinées en priorité à provoquer plus de flexibilité en terme de mouvements pour les utilisateurs de la réalité virtuelle. En fait les résultats finaux sont similaires. S'il est possible d'exciter le système vestibulaire durant l'immersion, la symptomatologie du mal du mouvement est alors réduite. Pour cela, dans les recherches, [Mon-Williams, 1993, Regan, 1994c], des bicyclettes sont utilisées pour naviguer à travers les mondes virtuels. Les développements les plus récents dans ce domaine sont réalisés aux Etats-Unis à la « Naval Postgraduate School » de Monterey en Californie [Delaney, 1996]. C'est ainsi que pour entraîner les soldats, un tapis roulant omnidirectionnel a été développé. Le but de ce projet est d'améliorer la mobilité des individus dans les environnements synthétiques. Le tapis roulant permet à l'utilisateur de marcher en rond et de tourner librement dans l'environnement virtuel. Ceci est très important pour les développements concernant la marche des fantassins débarqués dans les simulations distribuées interactives. En se basant sur la théorie du conflit sensoriel il est raisonnable de penser que l'utilisation du tapis roulant réduira les effets secondaires de l'immersion et en particulier les nausées. Il a été trouvé une corrélation inverse entre les cinétoses et le sentiment de présence [Witmer, 1994]. Ce qui permet de faire la liaison avec les effets psychologiques de l'immersion.

16.2.4 EFFETS PSYCHOLOGIQUES

Une revue intéressante de la presse [Wartella, 1993] met en évidence que la possibilité d'effets psychologiques ou comportementaux liés à l'immersion est abordée de manière similaire dans les médias depuis quelques années comme cela c'est réalisé dans le passé lors de l'introduction de la télévision. Néanmoins, il y a des risques garantis concernant l'utilisation de la réalité virtuelle. Par exemple les environnements virtuels peuvent entraîner des phénomènes de dépendance pour certaines personnes comme c'est le cas rapporté très récemment de gens devenu dépendant du grattage. Il est possible de penser aussi que les systèmes de réalité virtuelle, devenant peut-être un jour partie intégrante de l'amusement à domicile, peuvent jouer un rôle sur les enfants. Cependant, il semble [Sherman, 1992] qu'une éventuelle dépendance serait liée aux différences individuelles à la vulnérabilité.

La réalité virtuelle peut être vue comme une fuite du réel par certaines personnes. Il a été discuté que pour certaines personnes ayant des maladies psychologiques, il était difficile en réalité virtuelle de faire la distinction entre le réel et le fantasmatique. Johnston [Johnston, 1995] se pose le problème de savoir quelles peuvent être les conséquences à long terme de l'utilisation de la réalité virtuelle par exemple pour l'adulte avec des expériences d'immersion pendant l'enfance. L'immersion peut soulever des problèmes éthiques [Stone, 1993], en particulier, si la création de certains mondes entraîne un retentissement sur certaines personnes.

Sur le plan privé, il y a des risques d'isolement ou de toute puissance, voire de pratiques magiques (agir à travers un simulacre pour agir sur le réel). Cela demande que la mise en place de cette technique s'accompagne d'études sur la signification que cela représente pour l'utilisateur. Il faut analyser le vécu des sujets plongés dans cet autre monde. Il y a des risques de perte du sens du réel.

En terme d'études scientifiques un des problèmes abordé est celui du niveau de présence. Il est possible de citer les questionnaires passés pendant et après immersion [Regan, 1995, Slater, 1993, Henry, 1993]. De même il faut citer les travaux concernant les téléopérations [Schloerb, 1994, Schloerb, 1995] où le sentiment de présence se caractérise par celui d'être en même temps en deux endroits et de pouvoir agir à distance comme si on y était réellement.

Trois sous-classes ont été identifiées dans le questionnaire de présence (Presence Questionnaire ou PQ) [Witmer, 1994]. La première (Involved/Control) implication et direction rend compte de l'expérience d'être bien impliqué et de bien diriger les événements de l'environnement virtuel. La deuxième naturelle (natural) indique si les sensations expérimentées sont naturelles et aussi consistantes dans le monde virtuel que le monde réel. La troisième concerne la perturbation de l'interface dans la part qu'elle peut avoir en terme d'interférence et de distraction face à l'interaction avec l'environnement virtuel.

D'une manière plus expérimentale les études réalisées dans une thèse [Morineau, 1996] méritent une attention particulière. La première étude concerne l'encodage et la restitution en fonction d'un environnement réel ou virtuel. Il s'est référé aux travaux de Farnsworth 1924 et Smith et Guthrie 1921 qui montrent « l'effet de dépendance contextuelle ». On réalise mieux une tâche dans un endroit où elle a été apprise. Cela est vrai pour la restitution de mémorisation, par exemple de listes de mots. Morineau a créé deux situations : un bureau réel et un bureau virtuel assez ressemblant. Il a utilisé 20 sujets de 21 à 26 ans (3niveau CAP, 6 niveau bac, 11 supérieur au niveau bac). Le matériel comprenait une Silicon graphics, une carte graphique reality Engine2 avec une fréquence de rafraîchissement 80 Hertz, un taux d'affichage de 30 image par seconde, un casque Eyegen3 de champ visuel 40° en horizontal, un traqueur Polhemus et un joystick 3D. La représentation de la main virtuelle est à plat et figée. De plus, il y a possibilité de voler dans l'espace virtuel en apesanteur, il n'y a pas de résistance pour les objets et on peut passer à travers ceux-ci. Il a utilisé 4 listes de 25 mots tous à forte valeur d'imagerie et fort degré de concrétude (cheval, bandit). Les sujets étaient naïfs par rapport à la réalité virtuelle. Chaque sujet fait une visite de l'environnement virtuel et peut manipuler les objets durant cinq minutes. Le plan expérimental comprend 4 conditions d'encodage et de restitution ; Réel-Réel, Réel-Virtuel, Virtuel-Réel, Virtuel-Virtuel. L'ordre de passation est contrebalancé. On peut rappeler les mots dans n'importe quel ordre. Les mots sont présentés de manière auditive. Les résultats montrent un effet principal : le rappel est plus faible quand le mot est appris en virtuel. Le tableau suivant résume les résultats obtenus.

Il compare ses résultats avec ceux de Godden et Baddeley 1975 [Morineau, 1996] qui avaient trouvé que le rappel en plongée était moins bon qu'à l'air. On peut noter que quelle que soit la condition d'encodage les items centraux sont moins bien restitués. Ses autres études portent sur la comparaison des résultats obtenus par des adultes immergés et ceux d'enfants : des tests de mimes gestuels, de résolutions de problèmes d'inclusion ou d'estimation de chemin parcouru. D'une règle générale au cours d'une première immersion les adultes adoptent des comportements d'enfants de 8 ans (répé-

Encodage	restitution réel	restitution virtuel	Total
Réel	9,6 (2,2)	9,65 (2,5)	9,63
Virtuel	8,7 (2,7)	7,95 (2,5)	8,33
Total	9,15	8,8	8,98

Tableau 16.1: Moyenne de rappel de 25 mots et entre parenthèses écart-type

tition itérative d'une consigne comme taper avec marteau, identification à un pistolet quand on demande de mimer de tirer avec un pistolet sur une bouteille). En ce qui concerne les problèmes d'inclusion tous les adultes commettent des erreurs en immersion. Ces comportements infantiles diminuent au fur et à mesure des immersions. Tout se passe comme si les adultes utilisaient les schèmes archaïques pour découvrir le nouveau monde comme ceux qu'ils avaient utilisé pour découvrir le monde réel.

16.2.5 FACTEURS INFLUENTS

Les facteurs intervenant dans le mal de la simulation peuvent être classés en 3 catégories [Kolasinsky, 1995] : les facteurs individuels, les facteurs liés au simulateur, les facteurs liés à la tâche. C'est en s'appuyant sur cette classification que vont être abordés ces facteurs.

16.2.5.1 Différence individuelle dans la susceptibilité

Il y a quelques données qui suggèrent que la nature et la sévérité des effets secondaires sont en relation avec les facteurs individuels. Par exemple il y a une relation entre les antécédents de mal des transports et la susceptibilité aux effets secondaires de la réalité virtuelle. D'autres facteurs peuvent jouer un rôle important comme le sexe, l'âge la qualité de la vision, le niveau de concentration, le style de perception et la stabilité posturale. Quelques études suggèrent que les femmes sont plus susceptibles que les hommes au mal du mouvement. L'âge apparaît être significatif [Reason, 1974] qui rapporte que le mal des transports est important entre 2 et 12 ans et décroît rapidement pour être presque inexistant après 50 ans. Les données sont plus contradictoires en ce qui concerne le rôle du style perceptif ou de la stabilité posturale.

16.2.5.2 Facteurs liés à la nature de l'environnement virtuel et à la technique

Il est rapporté que des facteurs comme l'étendue de la scène visuelle, sa complexité, la gamme de couleur utilisée, la présence ou l'absence de mouvement, l'étendue des mouvements de tête nécessaire, la présence d'indicateurs auditifs, tactiles ou olfactifs joue un rôle dans l'apparition des effets secondaires à l'immersion. Par exemple, il est rapporté [Regan, 1994] que la présence de mouvements dans un environnement virtuel augmente les étourdissements et la désorientation. L'impact des autres variables, au cours des investigations, est cependant difficile à envisager sans prendre en compte leur interaction avec les caractéristiques techniques des systèmes de réalité virtuelle.

Par exemple il est difficile d'envisager le rôle de la complexité de la scène sans prendre en compte les effets de latence. D'autre part, il semble que plus les sujets immergés ont le sentiment que la simulation est réelle (concept de présence), moins ils présentent d'effets secondaires. Cependant deux caractères physiques méritent un peu plus de développement car ils jouent un rôle important dans l'apparition de symptômes. Il s'agit des délais temporels et du champ visuel.

Les systèmes retardateurs

Les délais sont fortement impliqués dans les effets secondaires de l'immersion dans les environnements virtuels. Il y a trois sources de délais :

- la vitesse de rafraîchissement des présentations (généralement autour de 60 Hz) qui est importante pour avoir une image de bonne qualité ;
- le taux de mise à jour des scènes qui est la fréquence à laquelle la scène visuelle peut changer. Quand la complexité croît, le taux de mise à jour décroît ;
- le temps nécessaire entre l'action de l'utilisateur et la perception par l'utilisateur de la conséquence de son action.

Quand le temps entre la construction d'un mouvement et ce mouvement dans l'environnement virtuel devient perceptible les effets secondaires deviennent probables. Des retards de 100 ms ont été signalés comme cause des malaises [Piantanida, 1993].

Champ de vision

Bien qu'actuellement les effets secondaires liées à des variations du champ visuel n'aient pas fait l'objet d'études systématiques, il semble que lorsqu'il y a un large champ visuel la sensation contraignante de présence entraîne une augmentation de ces effets. D'autre part les champs visuels réduits demandent de plus grands mouvements de tête et ainsi augmentent la désorientation et les nausées.

16.2.5.3 Les facteurs liés à la tâche

Il semble que des facteurs comme la durée, le degré de surveillance, la vitesse d'exécution soient des facteurs qui jouent un rôle en se superposant à beaucoup d'autres facteurs.

16.2.6 OUTILS D'ANALYSE DES ENVIRONNEMENTS VIRTUELS ET DES EFFETS DE L'IMMERSION

16.2.6.1 Outils d'analyse des effets de l'immersion

La revue de la littérature [Regan, 1994] montre que la plupart des auteurs utilisent des échelles et des outils différents pour évaluer les effets secondaires de l'immersion dans les environnements virtuels. D'autre part, il s'agit le plus souvent d'approches subjectives et les risques de biais expérimentaux ne sont pas nuls. Il est cependant possible de

retenir parmi ces outils le questionnaire du mal des simulateurs (Simulator Sickness Questionnaire) développé par Kennedy, Lane, Berbaum et Lilienthal (1993). Ce questionnaire est une auto évaluation concernant 16 symptômes (inconfort général, apathie, sueurs, nausées, difficulté de concentration, fatigue, maux de tête, fatigue visuelle, étourdissements, vertiges, pâleur, somnolence, estomac lourd, salivation, déséquilibre postural ou ataxie) qui sont classés par le sujet selon une échelle de sévérité en quatre points (0 = absent, 1 = léger, 2 = modéré, 3 = sévère). Ces données sont regroupées en trois sous-classes (nausée, inconfort oculaire, désorientation) et le total donne un score de sévérité.

L'Université de Nottingham a mis au point des outils plus objectifs pour mesurer ce type de malaise comme des analyses cardiovasculaire et neuroendocrinienne.

16.2.6.2 Outils d'évaluation des environnements virtuels

Un outil standard est utilisé aux Etats-Unis la VEPAB (Virtual environment performance assessment battery). Sa conception et les premiers essais sont décrits dans un article [Lampton, 1994]. Cet outil comprend

- tâches visuelles
 - acuité (lire des lettres présentées à 20 pieds de distance) ;
 - test de couleurs d'Ishihara ;
 - reconnaissance d'objet (silhouette humaine à 40 pieds de distance) ;
 - estimation de la taille (silhouette humaine à 40 pieds de distance) ;
 - estimation de la distance (dire si la silhouette humaine apparaît à 30, 20, 10, 5 ou 2,5 pieds) ;
 - détection d'une cible (balle rouge) apparaissant sur les murs, le plancher ou le plafond. Le sujet étant au centre de la pièce et regardant droit devant lui.
- tâches de locomotion à travers des couloirs des portes et des pièces, le déplacement se fait par exemple à l'aide d'un joystick et simule la progression d'un humain ayant une stature d'homme correspondant au 50^{ème} percentile ;
 - la première épreuve consiste à se mouvoir en suivant un couloir qui décrit un cercle et revient à son point de départ ;
 - la deuxième épreuve consiste à se mouvoir en suivant le même couloir que précédemment mais en marche arrière ;
 - la troisième épreuve consiste à se mouvoir dans un couloir qui décrit 10 virages alternativement à droite et à gauche ;
 - la quatrième épreuve consiste à suivre un itinéraire en 8 ;
 - la cinquième épreuve consiste à passer dans des salles par un ensemble de portes ;
 - la sixième épreuve consiste (tâche de vol) à se mouvoir dans l'espace, en particulier à passer par des fenêtres de 7 sur 3 pieds pour aller d'une pièce à une autre.
- tâches de poursuite. Quatre cas :
 - la tête se déplace, la cible est fixe ;
 - la tête se déplace, la cible se déplace ;
 - utilisation d'une commande avec une cible fixe ;

- utilisation d'une commande avec une cible mobile.
- tâches de manipulations d'objet à l'aide d'un curseur ;
 - saisir un barreau lisse (comme un manche à balai) et le déplacer horizontalement jusqu'à une cible en moins de trente secondes ;
 - saisir un bouton (comme une commande de radio) et le tourner d'une valeur donnée en moins de trente secondes ;
 - saisir une balle et la replacer dans une boîte ;
 - tâches de temps de réaction simple et avec choix.

Cet outil maintenant ancien, testait l'immersion visuelle, la manipulation se faisant soit avec un joystick, soit avec une souris 3D. Il a permis de montrer que le joystick permettait d'avoir des réponses plus rapides. Les résultats obtenus aux différents tests sont très dépendants de la qualité des casques qui ont été utilisés.

16.3 RÉFÉRENCES BIBLIOGRAPHIQUES

- [Delaney, 1996] B.Delaney, «*Where do you want to go today ? At last : An omni-directional treadmill*». Cyberedge Journal, Sep/Oct (1996).
- [Draper, 1998] M.H. Draper, «*The adaptive Effects of Virtual Interfaces : Vestibulo-Ocular Reflex and Simulator Sickness*». Doctoral Thesis at the University of Washington. (1998).
- [Howarth, 1994] PA Howarth & J.Bradbury, «An investigation into whether dissociating accommodation and convergence causes physiological stress». Report No. VIRART/94/111, Virtual Reality Applications Research Team, University of Nottingham. (1994).
- [Howarth, 1996] Howarth P.A. and Costello P.J, «*Studies into the Visual Effects of Immersion in Virtual Environments*». Visual Ergonomics Research Group (VISERG) in the Department of Human Sciences, Loughborough University, England. September 24 (1996).
- [Johnston, 1995] R. Johnston, «*Losing everything to VR*», Simulator sickness. Virtual Reality Special Report, May/June (1995).
- [Kennedy, 1993] R S. Kennedy, N.E. Lane, K.S. Berbaum & M.G.Lilienthal, «*Simulator sickness questionnaire (SSQ) : An enhanced method for quantifying simulator sickness*». The International Journal of Aviation Psychology, 3 (3), 203-220. (1993).
- [Kolasinsky, 1995] E.M Kolasinsky, «*Simulator sickness in Virtual environments*». US ARI Technical Report 1027. Alexandria, VA : U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA 295 861). (1995).
- [Lampton, 1994] D.R Lampton, B.W Knerr, S.L. Goldberg, J.P. Bliss, J.M. Moshell & B.S Blau, «*The virtual environment performance assessment battery (VEPAB) : Development and evaluation*». Presence, 3 (2), 145-157. (1994).
- [M.Défense, 1981] Ministère de la défense Manuel de médecine du travail et d'ergonomie appliquées aux armées, 2030 DEF/DCSSA/.TEC, Paris (1981).
- [Miyake, 1994] S. Miyake, J.Akatsu, M. Kumashiro, T. Murakami, H. Hirao, and M. Takayama, «*Effects of Virtual Reality on Human*». Proceedings of The 3rd Pan-Pacific Conference on Occupational Ergonomics. November 13-17. Seoul, Korea. (1994).

- [Mon-Will., 1993] M. Mon-Williams, J.P. Wann & S. Rushton, «*Binocular vision in a virtual world - Visual deficits following the wearing of a HMD*». *Ophthalmic and Physiological optics*, 13, 387-391.(1993).
- [Morineau, 1996] T Morineau, «*Adaptation cognitive à un environnement virtuel lors de premières immersions*». Thèse soutenue à Angers (1996).
- [Papin, 2000] J-P Papin, «*Facteurs humains et réalité virtuelle*». Cours à l'Ecole Navale de Brest (2000)
- [Piantanida, 1993] T Piantanida , D.K Bowman & J. Golle J, «*Human perceptual issues and virtual reality*». *Virtual Reality Systems*, 1 (1), 43-52. (1993).
- [Reason, 1974] J T Reason, «*Man in Motion*». Weidenfeld & Nicholson, London. (1974).
- [Regan, 1998] E C Regan, «*Effects of participating in immersive virtual reality : A review of health and safety issues*», Defence Research Agency Centre for Human Sciences Farnborough, UK Report to NATO group RSG 28. (1998).
- [Regan, 1993] E.C Regan & K.R Price, «*Some side-effects of immersion virtual reality : An investigation into relationship between inter-pupillary distance and ocular problèmes*». APRE Report 93R023. (1993).
- [Regan, 1994] E.C Regan & K.R Price, «*The frequency of occurrence and severity of side-effects of immersion virtual reality*». *Aviation Space, and Environmental Medicine*, 65, 527-530. (1994)
- [Regan,1994a] E.C. Regan & A.D. Ramsey, «*Some side-effects of immersion virtual reality : The results of four immersions*». APRE Report 94R012. (1994).
- [Regan,1994b] E.C. Regan & A.D. Ramsey, «*Some side-effects of immersion virtual reality : An investigation into the effectiveness of hyoscine hydrobromide in reducing nausea induced during immersion*». APRE Report 94025. (1994).
- [Regan,1994c] E.C. Regan & A.D. Ramsey, «*Some side-effects of immersion virtual reality : An investigation into the effect of two types of movement through a virtual environment on reported nausea*» ; CHS Report DRA/CHS/TR94009/1. (1994).
- [Riccio,1991] G. E. Riccio & T.A. Stoffregen, T. A, «*An ecological theory of motion sickness and postural instability*». *Ecological Psychology*, 3(3), 195-240. (1991).
- [Rushton,1994] S Rushton, M. Mon-Williams & J.P. Wann, «*Binocular vision in a biocular world : New generation head-mounted displays avoid causing visual deficit*». *Displays*, 15 (4), 255-260. (1994).
- [Sherman, 1992] B. Sherman & P. Jukins, «*Glimpses of heaven, Visions of hell : Virtual reality and its implications*». Hoddes & Stoughton, Kent. (1992).
- [Sherrer,1981] J. Sherrer et coll. «*Précis de physiologie du travail, notions d'ergonomie*». Masson, Paris, (1981).
- [Stone,1993] V.E. Stone, «*Social interaction and social development in virtual environments*». *Presence*, 2, 153-161. (1993).
- [Vaden,1996] E. A.Vaden, J.A. Ehrlich, & E.M. Kolasinski, «*Usability evaluation of low-end virtual reality systèmes*». *Proceedings of the human factors and ergonomics society annual meeting*, 1292. (1996).
- [Wartella, 1993] E Wartella & B. Reeves, «*Historical trends in research on children and the media : 1900-1960*», *Journal of Communication*, 35 (2), 118.(1993).
- [Witmer,1994] B. Witmer & M.J. Singer, *Measuring immersion in virtual environments*» (ARI Technical Report 1014). Alexandria, VA : U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA 286 183). (1994).

17 ETHIQUE, DROIT ET RÉALITÉ VIRTUELLE

Jean-Paul Papin

17.1 INTRODUCTION

Certaines recherches sur la réalité virtuelle peuvent entrer dans le cadre des recherches dites biomédicales et sont donc soumises à la loi Huriet (1991) du nom du sénateur qui l'a préparée. Ce même sénateur, en 1997, a écrit un rapport intitulé « images de synthèse et monde virtuel : techniques et enjeux de société » dans lequel il évoque les perspectives juridiques et une éthique concernant l'utilisation des images de synthèse et des nouvelles technologies d'information et de communication. Sur le plan international les problèmes éthiques et juridiques de la télé médecine ont été abordés dans un symposium qui s'est déroulé à Bruxelles en octobre 1997. Plus récemment (mars 2001) un congrès s'est tenu à l'Université de Paris X Nanterre sur le thème « Le virtuel : la présence de l'absent ». Parmi les communications présentées, Leleu, sexologue à Amiens, montre qu'Internet peut devenir, dans certains cas, une véritable drogue avec phénomène de dépendance (cela pouvant être accru avec la présentation d'avatar des sujets communicant entre eux). Cependant d'après les enquêtes de Tisseron, il semble qu'une pathologie n'est pas directement induite par ces modes de communication mais révèle plutôt une pathologie préexistante.

Dans différents domaines, en particulier à des fins de recherches de l'immersion de l'homme dans des environnements virtuels, il apparaît utile dans un ouvrage traitant de la réalité virtuelle d'aborder ces différents aspects juridiques et éthiques.

17.1.1 LOI SUR LA PROTECTION DES PERSONNES SE PRÊTANT À DES RECHERCHES BIOMÉDICALES

17.1.1.1 Introduction

La Loi n° 88-1138 (dite Loi HURIET) du 20 décembre 1988 mise en application le 1/01/1991 sur la protection des personnes se prêtant à des recherches biomédicales précise d'une part, ce qui est convenu d'appeler recherche biomédicale et d'autre part les règles régissant ces recherches. Cette loi précise :

- qu'est considérée comme recherche biomédicale tout essai ou toute expérimentation testant la réaction de l'homme, à un médicament, un produit, un objet, un matériel, une méthode ;
- qu'est considéré comme promoteur toute personne qui initie un essai ou une expérimentation ;
- qu'est considéré comme investigateur, un médecin chercheur désigné nominativement, après enquête, par le ministère de la santé publique, responsable d'un lieu de recherches agréé ;
- que le non-respect de la loi pour le promoteur et ou l'investigateur peut entraîner 2 mois à 1 an de prison et 6000F à 100000F d'amende.

17.1.1.2 *Vocabulaire*

- une recherche (essai ou expérimentation) implique une manipulation des conditions dans lesquelles sont effectuées des mesures sur un sujet. Ces conditions et ces mesures peuvent porter atteinte à l'intégrité humaine. Elle est destinée à acquérir des connaissances ou des lois nouvelles ;
- une étude est une observation sans atteinte de l'intégrité humaine par des mesures sur l'homme dans des conditions auxquelles le sujet peut être confronté de manière naturelle dans sa vie de travail ou privée. Elle est destinée à porter un jugement sur l'interaction homme conditions.

Les recherches sont soumises à la Loi Huriet, mais pas les études en général. Cependant il existe des cas particuliers où certaines études peuvent entraîner une atteinte de l'intégrité humaine, soit par la mesure, soit par les conditions. Ces études doivent être considérées comme devant être soumises à la Loi Huriet. A l'inverse certaines recherches ne portent pas atteinte à l'intégrité humaine.

17.1.1.3 *La notion d'intégrité humaine*

Cette notion est complexe. Il y a atteinte de l'intégrité, lorsqu'on pratique un prélèvement biologique (sang, biopsie), lorsqu'on installe une sonde, mais aussi lorsqu'on diminue les capacités physiologiques ou psychologiques par exemple, restriction du champ visuel). L'immersion dans les environnements virtuels peut s'assimiler au cas où les sujets ne seraient pas de manière naturelle, dans leur vie de travail ou privée, confrontés à de telles conditions. Dans l'état actuel il est donc par définition difficile de réaliser des études. Il n'est possible de faire que des recherches. De plus les atteintes de l'intégrité humaine sont une réalité. C'est pourquoi Il est recommandé à tous les promoteurs potentiels de s'assurer qu'ils respectent la loi. Ceci dans les cas où il est prévu des mesures anatomiques, physiologique, psychologique, biologique ou chimique sur l'homme, mais aussi s'il existe un risque humain. Il est donc proposé que chaque promoteur se pose la question de savoir si dans une recherche, une étude, un essai ou une expérimentation on teste en partie les réactions de l'homme.

17.1.1.4 *La notion de promoteur*

D'après la loi, seul est considéré comme promoteur celui qui est juridiquement et financièrement responsable, en général le directeur de l'établissement. C'est donc seulement lui qui légalement initie un essai, une recherche, une expérimentation. Dans la pratique celui qui prépare le dossier et qui fera ou fera faire le travail doit être au fait de cette loi. C'est pourquoi pour plus de simplicité les recommandations qui s'adressent légalement au promoteur, doivent concerner directement ceux qui veulent réaliser une recherche.

17.1.2 *RECOMMANDATIONS AUX PROMOTEURS ET AUX CHERCHEURS*

Nous proposons [Papin, 1999] que, pour les essais, les études et les recherches où un humain est dans la boucle (c'est le cas de la réalité virtuelle), le promoteur se pose

toujours la question suivante : *ceux-ci se font-ils avec ou sur l'homme ?* Par « avec » nous entendons qu'un sujet participe à l'essai, l'étude ou la recherche mais qu'il n'est pas lui-même sujet d'expérience et par « sur » qu'il est sujet d'expérience.

- dans le cas « d'avec » (par exemple télé opération d'un robot pour évaluer la qualité d'un nouveau système de détection d'obstacles sur des pistes d'essais). Le promoteur (ou le chercheur) devrait s'adresser au médecin du travail qui estime si un risque potentiel physique ou mental existe. Si un risque existe, il faut respecter les règles d'hygiène et de sécurité et médicaliser l'environnement (présence par exemple d'une équipe de secours sur les pistes d'essais) ;
- dans le cas de « sur » ou en cas de doute (par exemple télé opération d'un robot pour évaluer soit la diminution du stress des opérateurs (sûr) soit la facilité d'utilisation des commandes (pas sûr) pendant la mise en œuvre d'un nouveau système de détection d'obstacles sur des pistes d'essais), il est demandé au promoteur, quel que soit le degré de confidentialité de l'affaire, de contacter un médecin investigateur. Le médecin investigateur décidera si c'est une recherche biomédicale testant la réaction de l'homme à un médicament, un produit, un objet, un matériel, une méthode. Il est préférable que ce médecin puisse statuer et faire une réponse écrite dans les huit jours. Dans les cas où il n'y a pas lieu d'appliquer la loi, la réponse précisera que c'est pour le protocole présenté à l'exclusion de toute modification à posteriori. Si le médecin investigateur estime que c'est du ressort de la loi Huriet, il va faire une réponse demandant d'appliquer les mesures prévues. Cela demande des actions de l'investigateur et du promoteur. L'investigateur doit demander l'avis au comité consultatif de protection des personnes dans la recherche biomédicale (article R.2029 du décret n° 90-872 du 27 septembre 1990). La demande se fait par lettre recommandée avec accusé de réception. Elle doit présenter des renseignements sur la nature de la recherche et des renseignements sur la garantie des personnes. La réponse demande un délai de cinq semaines. Ce délai peut être prolongé d'un mois si des renseignements complémentaires sont demandés. Les renseignements à fournir concernent d'une part, l'identité du promoteur et du fabricant du produit, le titre et l'objectif de la recherche avec ou sans bénéfice individuel direct, les informations concernant le produit, l'identité des investigateurs avec titres, expériences et fonctions, une synthèse du dernier état des connaissances faisant la preuve de la nouveauté de la recherche et de son intérêt scientifique, les éléments du protocole expérimental et de la méthodologie clinique, la description des lieux et des personnes agréées avec copie des autorisations accordées. D'autre part des renseignements sur la garantie des personnes. En particulier, il est nécessaire de préciser que des informations seront communiquées aux personnes les informant sur les objectifs de la recherche, les bénéfices attendus, le droit pour les personnes de retirer leur consentement. Il faudra fournir un justificatif des modalités de recueil du consentement, de l'attestation d'assurance, du montant des indemnités dues aux personnes et préciser la durée d'exclusion. Parallèlement le promoteur doit déclarer son intention au ministre chargé de la santé (article R.2032 du décret n° 90-872 du 27 septembre 1990) conformément au formulaire enregistré au C.E.R.F.A sous le n° 65-0042.

17.1.3 AUTORISATION DES LIEUX DE RECHERCHES BIOMÉDICALES SANS BÉNÉFICE INDIVIDUEL DIRECT

Pour obtenir une autorisation des lieux de recherches, ceux-ci doivent comporter :

- une superficie, une distribution et un aménagement compatibles avec un bon déroulement des activités de recherches prévues ;
- des moyens en équipements et en personnels permettant d'assurer une surveillance clinique constante et des soins d'urgence, ainsi que la possibilité d'effectuer, si nécessaire un transfert immédiat dans un service de soins approprié ;
- des moyens d'archivage et de protection des données et des informations concernant les recherches et les personnes.

17.2 ETHIQUE ET DROIT

17.2.1 PERSPECTIVES JURIDIQUES ET ÉTHIQUES SELON HURIET

Le sénateur Huriet [Huriet, 1997] insiste sur les notions de droit positif et de déontologie en faisant remarquer que la tentation réglementaire doit être évitée dans les domaines où les évolutions techniques sont très rapides. Il cite, en particulier, les échecs des réglementations concernant les jeux vidéo (décret n° 96-360 du 23 avril 1996 relatif aux mises en garde concernant les jeux vidéo) et préconise l'adaptation des règles existantes. Par exemple il suggère qu'une réflexion internationale soit menée en ce qui concerne le principe de propriété intellectuelle sur Internet. De même il rappelle qu'un article du code pénal (article 226-8) réprime « la publication des montages réalisés avec l'image d'une personne sans son consentement ». Sur le plan déontologique, il préconise que les professionnels mettent au point des codes en matière de jeux vidéo, de publicité virtuelle, mais aussi qu'il y ait un apprentissage à ces nouvelles techniques.

En ce qui concerne l'éthique le rapport Huriet préconise qu'« il ne paraît pas souhaitable, cependant, de traiter à part le virtuel. Il faut éviter la multiplication d'instances de conseil ou de régulation spécialisées d'autant que l'image, le virtuel, l'information forment un ensemble indissociable. Le virtuel fait partie de l'image, l'image de l'information, l'information des nouvelles technologies d'information et de communication ». Pour lui, les organismes déjà existants, comme par exemple la commission nationale de l'informatique et des libertés ou l'autorité de régulation des télécommunications, peuvent mener à bien cette réflexion éthique.

17.2.2 ETHIQUE ET DROIT EN TÉLÉ MÉDECINE

Ghent [Raes, 1997] a proposé une réflexion sur les problèmes éthiques et juridiques que pose la pratique de la télé chirurgie et du télé diagnostique au cours d'un symposium international de télé médecine à Bruxelles en octobre 1997. En introduction il rappelle la position de Merell de 1995 [Raes, 1997] : « *la télé médecine doit être regardée comme un exercice de la médecine de soins et non comme de l'électronique* ». La télé médecine couvre tous les aspects de transmission d'information à distance dans tous les domaines médicaux (radiologie, psychiatrie, dermatologie, etc.). Cependant elle peut être divisée en deux grandes catégories. La contribution de la télé médecine pour délivrer juste un système de soins plus économique, plus accessible et plus efficace. La contribution de la télé médecine pour accroître la qualité des soins prodigués à des patients particuliers. La qualité porte sur la technique mais aussi sur le vécu de la relation entre patient et praticien. Un risque existe celui de la déshumanisation de la relation. Une des facettes de ces pratiques et celle de la responsabilité entre les

différents acteurs, celui qui conseille, celui qui pratique, celui qui transmet, celui qui conçoit les systèmes. Une autre facette est celle du secret médical au cours des transmissions. En conclusion la télé médecine ne doit pas envisager que la communication entre praticiens mais doit prendre en compte la communication à distance entre patient et praticien et trouver comment ne pas perdre la relation de face à face permettant la confiance et la confiance qui existe dans la relation classique patient praticien.

17.3 RÉFÉRENCES BIBLIOGRAPHIQUES

[Huriet, 1988]	Loi n° 88-1138 (dite Loi HURIET) du 20 décembre 1988.
[Huriet, 1997]	C. Huriet, « <i>Images de synthèse et monde virtuel : techniques et enjeux de société</i> ». Rapport n° 169 du Sénat et n° 526 de l'Assemblée Nationale (11 décembre 1997).
[Leleu, 2001]	P. Leleu, « <i>Internet comme intercorps : Stéphanie et le prince charmant virtuel</i> ». Communication au congrès Le virtuel : présence de l'absent. 3 mars 2001 Université de Paris X Nanterre, (2001).
[Tisseron, 2001]	S. Tisseron, « <i>L'Identité à l'épreuve du virtuel</i> ». Communication au congrès Le virtuel : présence de l'absent. 3 mars 2001 Université de Paris X Nanterre, (2001).
[Papin, 1999]	J-P. Papin, « <i>La loi Huriet</i> ». Présentation au laboratoire de Robotique de Paris, (1999).
[Raes, 1997]	K. Raes, « <i>Ethical aspects of telesurgery and telediagnosics</i> ». International symposium on telemedicine, Bruxelles, (1997).

GLOSSAIRE

Aides Logicielles Comportementales (ALC)

Une Aide Logicielle Comportementale programmée, associée à une PCV et à son interface comportementale, apporte une aide mentale pour l'accomplissement de la tâche par l'utilisateur.

Bureau immersif

Voir à visiobureau

Cabine de simulation de mouvement

Une cabine de simulation de mouvement est une interface à simulation de mouvement pour plusieurs utilisateurs.

Capteur de localisation

Voir à traqueur

Capture de mouvement

Voir à motion capture

Casque immersif

Voir à visiocasque

Cave

Nom anglais d'une marque américaine de visiocube ; terme utilisé couramment pour parler des visiocubes.

Exosquelette

Un exosquelette est un système mécanique composé de barres articulées portées par les différentes parties du corps de l'utilisateur. Les exosquelettes sont presque tous des interfaces sensori-motrices qui ont pour objectif de fournir des retours d'effort sur le corps, tout en mesurant secondairement les mouvements relatifs de toutes les parties du corps entre elles. Quelques rares exosquelettes sont des interfaces motrices qui ont pour objectif de mesurer seulement les mouvements relatifs de toutes les parties du corps entre elles, dispositifs faisant partie de la technique de motion capture.

Fenêtre immersive

Une fenêtre immersive est une interface sensorielle visuelle qui permet à l'observateur de regarder par une fenêtre dans toutes les directions l'environnement virtuel, grâce à un écran plat, localisable et orientable, suspendu à un portique.

Gant de données (ou gant numérique)

Le gant de données est une interface motrice qui mesure en temps réel les mouvements relatifs des doigts de la main. Associé presque toujours avec un traqueur, le gant de données mesure alors les mouvements absolus de la main et des doigts par rapport au repère de référence du traqueur.

Gant numérique

Voir à gant de données.

Immersion et interaction (I²)

Dans un dispositif de réalité virtuelle, l'utilisateur est immergé dans un monde virtuel et peut interagir avec ce monde artificiel. Dans un premier niveau **d'immersion et d'interaction sensori-motrices**, le sujet est connecté à l'ordinateur par ses sens et ses réponses motrices via des interfaces matérielles. Ce premier niveau d'I² sensori-motrices concerne donc les aspects informatiques temps réel pour l'ordinateur, les aspects physiques pour les interfaces matérielles et les aspects psychophysiques pour les caractéristiques sensori-motrices du sujet.

Dans un deuxième niveau **d'immersion et d'interaction mentales**, le sujet s'immerge mentalement dans le monde virtuel, le niveau inférieur d'I² sensori-motrices devant lui être mentalement invisible (transparent). Ce deuxième niveau d'I² mentales concerne les processus mentaux exploités par le sujet et la modélisation comportemental du monde virtuel.

Le troisième niveau **d'immersion et d'interaction fonctionnelles** concerne les objectifs de l'application de réalité virtuelle. Ces objectifs s'attachent à réaliser une immersion et une interaction du sujet pour des fonctionnalités données et non pour une simple immersion mentale de l'homme dans ce monde virtuel. Ces fonctionnalités sont décomposables en Primitives Comportementales Virtuelles (PCV).

Interface à retour d'effort à réaction interne et celles à réaction externe

Il y a deux types d'interfaces à retour d'effort : le mécanisme créant des efforts s'appuie soit sur une autre partie du corps de l'utilisateur, soit sur un socle fixe. La première solution est **une interface à retour d'effort à réaction interne**. La deuxième solution est **une interface à retour d'effort à réaction externe**. Elles font partie des interfaces sensori-motrices, car toutes mesurent en même temps les mouvements du(des) point(s) où sont appliqués les efforts.

Interface à retour tactile

Les interfaces à retour tactile sont des interfaces sensorielles qui interagissent par des actions mécaniques superficielles avec les récepteurs tactiles de la peau.

Interface à retour thermique

Les interfaces à retour thermique sont des interfaces sensorielles qui stimulent les récepteurs thermiques de la peau.

Interface à sensibilité cutanée

La sensibilité superficielle cutanée englobe la sensibilité mécanique (le sens du «toucher») et la sensibilité thermique. Les interfaces à sensibilité cutanée sont des interfaces sensorielles qui sont donc composées des interfaces à retour tactile et des interfaces à retour thermique.

Interface à simulation de mouvement

Une interface à simulation de mouvement est une interface sensorielle qui a pour objectif de modifier l'orientation dans l'espace du corps d'un ou plusieurs utilisateurs et de leur faire subir des accélérations en rotation et(ou) en translation.

Interfaçage comportemental

Dispositif matériel et processus mentaux associés qui permettent à un sujet de s'im-

merger et d'interagir dans un environnement virtuel.

Interface comportementale

Dispositif matériel qui exploite un comportement humain, naturel et qui n'implique pas une période d'apprentissage importante et complexe pour pouvoir être utilisé.

Les interfaces comportementales sont composées de trois catégories d'interfaces :

Interface motrice

Dispositif matériel conçu pour transmettre des réponses motrices de l'homme vers l'ordinateur.

Interface sensorielle

Dispositif matériel conçu pour transmettre des stimuli sensoriels de l'ordinateur vers l'homme.

Interface sensori-motrice

Dispositif matériel conçu pour transmettre des réponses motrices et en réaction des stimuli sensoriels sont renvoyés par l'ordinateur (actuellement, ce ne sont que les interfaces à retour d'effort qui ont ces deux fonctions associées).

Interface olfactive

Une interface olfactive est une interface sensorielle qui est un diffuseur programmable d'odeurs préalablement stockées en petites quantités.

Motion capture (ou capture de mouvement)

Les systèmes de localisation de tout le corps sont utilisés pour mesurer, en temps réel ou en temps différé, les mouvements relatifs de toutes les parties du corps entre elles. Indirectement, ils peuvent aussi déterminer la configuration du squelette à tout instant. Dans le langage habituel, pour désigner ce type de dispositif, on emploie le terme anglais «**motion capture**», plutôt que «**capture de mouvement**». Employé pour la réalité virtuelle, c'est une interface motrice qui fait partie des interfaces spécifiques de localisation corporelle.

Présence du sujet

Le terme «présence» recouvre la notion d'immersion mentale du sujet dans le monde virtuel, où le sujet est soit passif (spectateur), soit actif (acteur). Ce dernier cas, la présence active, correspond à la finalité de la réalité virtuelle. Le deuxième niveau d'immersion et d'interaction mentales correspond donc à la présence active du sujet.

Primitives Comportementales Virtuelles (PCV)

Quand le sujet se trouve en environnement virtuel, il a une activité générale qui se décompose en activités élémentaires, (sensori-motrices et mentales), que nous appelons PCV. Ces dernières peuvent être regroupées en quatre catégories en environnement virtuel :

- observer le monde virtuel ;
- se déplacer dans le monde virtuel ;
- agir sur le monde virtuel ;
- communiquer avec autrui.

Salle immersive

Voir à visiosalle

Siège de simulation de mouvement

Un siège de simulation de mouvement est une interface à simulation de mouvement pour un utilisateur.

Souris 3D

Une souris 3D est une interface manuelle motrice permettant à l'utilisateur **que de commander** des mouvements d'une entité, en général trois translations perpendiculaires et trois rotations perpendiculaires, dans l'espace de l'environnement virtuel.

Traqueur ou capteur de localisation

Un traqueur est une interface motrice permettant une **localisation en temps réel** de tout objet pour le positionnement et le suivi de sa trajectoire. La localisation concerne les translations et(ou) les rotations de l'objet par rapport à un repère fixe de référence.

Vidéolunettes

Une vidéolunette est une interface légèrement différente d'un visiocasque : c'est une interface sensorielle portable de visualisation dont l'objectif est de fournir une *vision monoscopique ou stéréoscopique* à l'aide d'un ou de deux petits écrans et d'optiques proches des yeux, avec un assez *grand champ visuel* et avec une **immersion partielle**, car les écrans et les optiques n'obturent pas toute la vue de l'observateur qui perçoit aussi le monde réel extérieur : la vidéolunette permet de voir les images virtuelles tout en n'étant pas totalement coupé du monde extérieur. **L'immersion du regard** avec la tête traquée est facultative.

Visiobureau ou bureau immersif

Un visiobureau est une interface sensorielle visuelle qui permet d'afficher sur une grande surface, de la grandeur d'un bureau ou d'une planche à dessin, des images stéréoscopiques permettant ainsi à un ou deux observateurs de visualiser des objets tridimensionnels. Ces bureaux immersifs sont composés en général d'un ou de deux écrans, les images étant affichées par l'intermédiaire d'un ou de deux rétroprojecteurs et éventuellement via des miroirs. Le point de vue n'est correct que pour un observateur (rarement deux), ayant la tête traquée.

Visiocasque ou casque immersif

Un visiocasque (traduction de «Head Mounted Display») est une interface sensorielle portable de visualisation dont l'objectif est de fournir une **vision stéréoscopique** à l'aide de petits écrans et d'optiques proches des yeux, avec un assez **grand champ visuel** et avec **une immersion du regard**. Pour cette dernière fonction, un traqueur doit être positionné sur le visiocasque, permettant d'afficher les points de vue correspondant à la direction de la tête de l'observateur, donc pas dans la direction des yeux (*du regard*) car actuellement aucun visiocasque ne traque le mouvement des yeux).

Visiocasque semi-transparent (ou visuel de casque)

Un visiocasque semi-transparent (en anglais See-through HMD) a les mêmes fonctionnalités qu'un visiocasque. Cette interface sensorielle permet en plus de mixer des images de synthèse avec la vision réelle de l'observateur. Des écrans semi-transparents permettent de superposer en temps réel des images de synthèse à la vision du monde réel.

Visiocube

Un visiocube est une interface sensorielle de visualisation composée principalement d'une station graphique et d'un ensemble de vidéoprojecteurs (trois au minimum) permettant l'affichage d'une image *sur plusieurs grands écrans à taille humaine, composant trois à six faces d'un cube, l'observateur étant à l'intérieur du cube (ou d'un parallélogramme)*. Il permet ainsi *l'immersion du regard dans plusieurs directions* (dans toutes pour un visiocube à six faces) pour un observateur ayant sa tête, traquée pour un

parfait affichage de son point de vue en temps réel.

Visiosalle ou salle immersive

Une visiosalle est une interface sensorielle de visualisation composée principalement d'une station graphique et d'un ensemble de vidéoprojecteurs (trois au minimum) permettant l'affichage d'une image à haute résolution *sur grand écran à taille humaine et de forme plate ou semi-cylindrique*. Elle permet ainsi *l'immersion dans un grand champ de vision* pour quelques dizaines de personnes qui regardent simultanément des images de synthèse générées en temps réel.

Visuel de casque

Voir à visiocasque semi-transparent.

TABLE DES ILLUSTRATIONS

1.1	Simulation de la conduite d'une voiture avec visiocasque ou sur écran de simulateur, copyright Photo PSA Peugeot Citroën	4
1.2	Comparaison entre un monde réel et un monde virtuel figé	8
1.3	Comparaison entre un monde réel et un monde virtuel interactif	9
1.4	«Réalité à Compréhension Augmentée», photo Ecole des Mines de Paris	14
1.5	Réalité à Visibilité Augmentée avec représentation des contours visibles et invisibles, photo Ecole des Mines de Paris	15
1.6	Association d'un distributeur virtuel autour d'un mécanisme réel ayant le même distributeur (Simulation de montage du mécanisme), photo Ecole des Mines de Paris	15
1.7	Le robot téléopéré est affiché simultanément en images réelles en relief et en images de synthèse en relief, photo Ecole des Mines de Paris .	16
1.8	Domaines précurseurs de la réalité virtuelle	19
1.9	Les fonctions intrinsèques de la réalité virtuelle	23
1.10	Simulation chirurgicale avec retour d'effort, copyright photo INRIA Sophia Antipolis	26
1.11	La conduite réelle avec les techniques de réalité virtuelle et augmentée, photo J.P. Papin	28
1.12	Présentation virtuelle d'une voiture pour la vente, réalisée par PSA, copyright Photo PSA Peugeot Citroën	29
1.13	Schéma de la «CAO classique»	30
1.14	Schéma de la«CARV»	30
1.15	Etude de la visibilité avec un visiocasque, copyright photo PSA Peugeot Citroën	33
1.16	Simulation de la tenue d'un navire sur la mer, photo Médialab	34
1.17	Revue de projet d'un véhicule, copyright photo PSA Peugeot Citroën . .	35
2.1	Transfert de l'information dans l'organisme	39
2.2	Schéma anthropocentrique de l'homme dans le monde réel	40

2.3	Interfaçage sensoriel	41
2.4	Interfaçage moteur	42
2.5	Schéma anthropocentrique de l'homme dans un monde virtuel	44
2.6	Différence d'architecture entre les systèmes de réalité virtuelle et de téléopération	46
2.7	Schéma anthropocentrique de l'homme dans un monde virtuel	47
2.8	Schéma technocentrique d'Immersion et d'Interaction sensori-motrices .	48
2.9	Schéma triadique de l'interfaçage en réalité virtuelle	49
2.10	Schéma technocentrique de référence en RV	55
2.11	Observation du magasin sur un écran de 4m par 2m, photo Ecole des Mines de Paris	63
2.12	Réalisation de l'écran avec six vidéoprojecteurs LCD (3072x1536 pixels au total), photo Ecole des Mines de Paris	64
2.13	Manipulation d'un produit avec un capteur de localisation (6ddl), inclus dans un cube, photo Ecole des Mines de Paris	65
2.14	Schéma des déplacements du chariot instrumenté	66
2.15	Interfaçage comportemental par chariot instrumenté pour les déplacements dans le magasin, photo Ecole des Mines de Paris	66
2.16	Interfaçage comportemental par un tapis roulant asservi et un «guidon» pour se déplacer, photo Ecole des Mines de Paris	69
2.17	Immersion visuelle sur grand écran avec rétroprojection, photo SNCF .	70
2.18	Interface comportementale : gant de données pour manipuler le téléphone, l'aiguillage, etc., photo SNCF	71
3.1	L'expansion de l'homme	81
3.2	Les différentes ambiances réelles et les organes des sens	82
3.3	Importance relative des consultations visuelles, au cours d'un vol à vue en hélicoptère, en fonction du champ visuel. Chaque point représente 1 % des consultations	86
3.4	Performances humaines et ambiance thermique	91

3.5	Facteurs intervenant dans la réussite d'une tâche	92
4.1	Transfert de l'information dans l'organisme	95
4.2	Projections rétinienne vers le corps genouillé latéral, puis le cortex strié	98
4.3	Structure anatomique de l'œil humain	99
4.4	Accommodation à l'infini et accommodation à courte distance (schéma optique simplifié)	99
4.5	La structure anatomique de la rétine	100
4.6	Définition de la fréquence spatiale : sur l'exemple, la fréquence spatiale est de 4 cpd	104
4.7	Les réponses d'une cellule ganglionnaire à différents stimuli lumineux .	105
4.8	Le phénomène de contraste de la grille d'Hermann-Hering	106
4.9	Modélisation du profil des champs récepteurs d'une cellule ON-OFF par une différence de gaussiennes	106
4.10	Les variations de lumière et les ombres augmentent le « relief » des deux cubes	109
4.11	Voyez-vous 4 roses de même taille plus ou moins éloignées ou voyez-vous 4 roses sur un même plan ?	110
4.12	Voyez-vous une ellipse devant un rectangle ou une ellipse à même distance d'un rectangle découpé sur un côté ?	110
4.13	Variation de la visibilité et de la perception de la texture, photo A. Fuchs	110
4.14	Deux parallélépipèdes et non un ensemble de lignes sur une feuille plane !	111
4.15	Les indices permettant la perception de la profondeur de l'espace tridimensionnel	112
4.16	Fixation sur le doigt ou fixation sur l'arbre	112
4.17	Perception de la différence de profondeur entre le point de fixation F et le point A, à partir de la disparité	113
4.18	Courbe d'adaptation à l'obscurité	116
4.19	Sensibilité au contraste temporel (d'après de Lange 1958)	116
4.20	Schéma pour calculer l'acuité stéréoscopique Δr	118

4.21	La parallaxe horizontale est définie par l'angle DPG	120
4.22	Propagation des sons	123
4.23	Son périodique mais non sinusoïdal	124
4.24	Coupe de l'oreille	126
4.25	Coupe de l'oreille interne	127
4.26	Les voies auditives	127
4.27	Tonotopie cochléaire	128
4.28	Fonctionnement de la cochlée	129
4.29	Schéma des propagations des sons aux oreilles	129
4.30	Voies sensibles tactiles et kinesthésiques	135
4.31	Détails des cellules olfactives et l'organisation générale	137
4.32	La langue avec les différentes zones du goût	140
4.33	Classification des accélérations en fonction des axes du corps	142
4.34	Voies Kinesthésiques (d'après Guyton 1984)	143
4.35	Illusion de mouvement (d'après Roll et al. 1986)	144
5.1	Principe basé sur les mesures de distances entre points	157
5.2	Principe basé sur les mesures de champ	158
5.3	Traqueur de localisation mécanique à leviers articulés	160
5.4	Mesure de position d'un doigt et retour d'effort sur celui-ci	161
5.5	Mécanisme de mesure et de tension d'un fil par moteur-encodeur	161
5.6	Principe d'un inclinomètre	162
5.7	Le traqueur Gyropoint II Pro sans fil de IXMicro Gyropoint, copyright photo IXMicro Gyropoint	164
5.8	Principe d'un accéléromètre	165
5.9	Champ électromagnétique autour d'une bobine	167

5.10	Dispositifs d'un capteur de localisation électromagnétique	167
5.11	Traqueur Ultratrak de Polhemus, copyright photo SimTeam	169
5.12	Traqueur Bird d'Ascension, copyright photo Immersion S.A.	170
5.13	Les antennes SpacePad de Ascension, copyright photo Immersion S.A. .	171
5.14	Très petit récepteur du traqueur miniBird d'Ascension, copyright photo Immersion S.A.	171
5.15	Le traqueur sur carte PC : pcBird d'Ascension	172
5.16	L'InterTrak2 à 3 ddl d'orientation à technologie inertielle d'InterSense, copyright photo SimTeam	173
5.17	Le traqueur 600 Mark2 de Intersense, copyright photo SimTeam	174
5.18	Dispositifs d'un traqueur de localisation à ultrasons	175
5.19	Traqueur acoustique de Logitech qui n'est plus commercialisé, copyright photo Immersion S.A.	175
5.20	Capteur de localisation optique avec récepteurs ponctuels	178
5.21	Traqueur de localisation optique avec récepteurs ponctuels	179
5.22	Le traqueur optique DynaSight pour le suivi de mouvement de paires de lunettes, copyright photo Origin Instruments Corporation	180
5.23	Le traqueur optique à 6 ddl Active Target Adapter pour le suivi de mou- vement de visiocasque, copyright photo Origin Instruments Corporation	181
5.24	Principe de l'extérieur vers l'intérieur pour un capteur de localisation optique	181
5.25	Le dispositif de 3 caméras et deux cibles sur deux doigts du système Coda motion capture pour le suivi de doigts, copyright photos Charnwood Dynamics	182
5.26	Capteur de localisation optique : Système Optotrak	183
5.27	Principe de l'intérieur vers l'extérieur pour un capteur de localisation optique	184
5.28	Schéma résumant symboliquement les différences entre les différents tra- queurs	186
5.29	Repères de référence	187

5.30	Couplage entre l'émetteur et le récepteur	188
6.1	La combinaison de données MotionStar sans fil d'Ascension, copyright photo Immersion S.A.	196
6.2	La capture de mouvement par exosquelette Gypsy de Analogus Corporation, copyright photo Analogus Corporation	198
6.3	Système Optotrak de Northern Digital Inc., copyright photo Northern Digital Inc.	200
6.4	Le dispositif de 3 caméras du système de Coda motion capture, copyright photo Charnwood Dynamics	200
6.5	La motion capture avec le Cyberdôme/CNBDI de XD Productions, copyright photos XD Productions	202
6.6	L'interface de localisation des membres supérieurs de Exos, copyright photo Exos	203
6.7	Exosquelette Upper Body Waldo de Character Shop, copyright photo Character Shop	204
6.8	Exosquelette Dual Arm Waldo de Character Shop, copyright photo Character Shop	205
6.9	Dispositif d'analyse de la posture d'une personne, photo J.P. Papin	206
6.10	Interface de locomotion à patins à roulettes	207
6.11	Interface de locomotion : un tapis roulant asservi et un «guidon» pour se déplacer, photo Ecole des Mines de Paris	208
6.12	Interface de locomotion par tapis 1D de Sarcos, copyright photo Sarcos . . .	209
6.13	Tapis 1D avec écrans panoramiques de Sarcos, copyright photo Sarcos . .	209
6.14	Interface de locomotion par tapis 2D Treadmill, copyright photo Professor Iwata	210
6.15	Schéma des mouvements sur le tapis 2D Treadmill, copyright photos Professor Iwata	211
6.16	Pédalier Stepper de Sarcos, copyright photo Sarcos	211
6.17	Le système V-Actor de SimGraphics, copyright photo SimGraphics	213
6.18	Système de capture du visage Facial Waldo de Character Shop, copyright photo Character Shop	214

6.19	L'oculomètre de SensoMotoric Instruments, copyright photo SensoMotoric Instruments	216
6.20	L'oculomètre de SensoMotoric Instruments, copyright photo SensoMotoric Instruments	216
6.21	Dispositifs électrophysiologiques Biomuse, copyright photo Moto Development Group	218
7.1	Gant de données à fibres optiques	223
7.2	Le gant de données SuperGlove de Nissho Electronics Corporation, copyright photo Nissho Electronics Corporation	223
7.3	Mesure de la déformation d'un doigt par translation d'une fibre optique	224
7.4	Le gant de données optique 5DT DataGlove, copyright photo SimTeam	224
7.5	Gant CyberGlove de Virtual Technologies, copyright photo Immersion S.A.	225
7.6	Gant de données à effet Hall	226
7.7	Le gant de commandes binaires Pinch Glove de Fakespace Labs, copyright photos Fakespace Labs	227
7.8	Souris 3D SpaceBall, copyright photo Immersion S.A.	229
7.9	La souris 3D Space mouse de 3Dconnexion, copyright photo 3Dconnexion	230
8.1	Poignée avec capteurs de pression pour la commande d'une main virtuelle	236
8.2	L'interface électromagnétique à retour d'effort : Haptic Magnetic Levitation device	238
8.3	La manette Magic Wrist d'IBM à six ddl à retour d'effort, copyright photo IBM	238
8.4	Les deux architectures d'une interface à retour d'effort	239
8.5	Le bras maître MA 23 M de Cybernétix à 7 ddl à retour d'effort (position et pince), copyright photo Cybernectix	241
8.6	Le Phantom Premium, version 1.5 de SensAble, copyright photo SimTeam	243
8.7	Le Phantom premium, version 3.0 de SensAble, copyright photo SimTeam	245
8.8	Le Phantom Desktop de SensAble, copyright photo SimTeam	246

8.9	Le Phantom 1.5/6DOF de SensAble, copyright photo SimTeam	247
8.10	Le Virtuouse 6D de Haption, copyright photo Haption	248
8.11	Le Virtuouse 3D du CEA-LIST et de Haption, copyright photo Haption	249
8.12	L'interface à retour d'effort Freedom 6S de MPB Communications Inc., copyright photo MBP Communications Inc.	250
8.13	Schéma de l'interface à retour d'effort à structure parallèle Haptic Master	251
8.14	L'interface à retour d'effort à structure parallèle HapticMaster, copyright photo Professor Iwata	252
8.15	L'interface à retour d'effort Delta haptic Device de ForceDimension, co- pyright photos ForceDimension	253
8.16	L'interface à retour d'effort Impulse Engine de Immersion Corporation, copyright photo Immersion Corporation	255
8.17	Mesure de position d'un doigt et retour d'effort sur celui-ci	257
8.18	Mécanisme de mesure et de tension d'un fil par moteur-encodeur	257
8.19	La manette Robotop de Matra Marconi Space, copyright Matra Marconi Space	258
8.20	L'interface Excalibur à trois ddl à retour d'effort du Biorobotics Labora- tory, copyright photo Biorobotics Laboratory	259
8.21	La souris à deux ddl à retour d'effort Mouse Cat, copyright photo Im- mersion Corporation	260
8.22	L'exosquelette semi-portable Dextrous Arm Master de Sarcos, copyright photo Sarcos	261
8.23	L'exosquelette semi-portable de BDI, copyright photo Boston Dynamics	262
8.24	Le gant à retour d'effort Rutgers Master I	263
8.25	Le gant de données à cinq retours d'effort CyberGrasp de Virtual Tech- nologies, copyright Immersion S.A.	264
8.26	le Sac à dos GraspPack, copyright photo Immersion S.A.	264
8.27	L'interface à retour d'effort CyberForce de Virtual Technologies, copy- right photo Immersion S.A.	265
9.1	L'interface thermique X/10 et une thermode, copyright photo CM Research	273

9.2	Principe d'un retour tactile par une matrice d'aiguilles	275
9.3	Le tactor TDITAC8x9 de la société Transdimension, copyright photo Transdimension	275
9.4	L'interface tactile «Tactor Evaluation System» de Engineering Acoustics, copyright photo Engineering Acoustics	276
9.5	Le gant de données Cyberglove avec retour vibro-tactile Cybertouch de Virtual Technologies, copyright photo Immersion S.A.	277
9.6	L'interface tactile TouchMaster de EXOS, copyright photo Exos	278
9.7	Le système TOPIT pour cockpit virtuel, copyright photo CGSD Corporation	279
9.8	Le gant et la platine mobile de boutons du système TOPIT, copyright photo CGSD Corporation	280
10.1	Triple écran pour une immersion grand angle mono-utilisateur, copyright photo Panoram technologies	285
10.2	Principe de la vision avec l'écran autostéréoscopique à illumination . .	286
10.3	Principe de la vision avec l'écran autostéréoscopique à réseau lenticulaire	286
10.4	Projection stéréoscopique avec lunettes filtrant par différenciation colorimétrique, photo TAN ProjektionsTechnologie GmbH	288
10.5	Principe de la vision stéréoscopique avec lunettes actives	289
10.6	Vision en relief par lunettes actives Stereotek de Nuvision, copyright photo Immersio S.A.	289
10.7	Principe de la vision stéréoscopique avec lunettes passive et multiplexage temporel	290
10.8	Principe de la vision stéréoscopique avec lunettes passives et multiplexage spatial	291
10.9	Système Kappa de restitution stéréoscopique d'images réelles, copyright photo ThetaScan	292
10.10	Window VR de Virtual Research, distribuée par Immersion S.A., copyright photo PSA	292
10.11	Le bureau immersif à deux écran Consul de Barco, copyright photo Barco	294

10.12	Le bureau immersif Holobench de TAN, copyright photo Tan Projektionstechnologie GmbH	294
10.13	Le visiobureau Baron de Barco, copyright photo Barco	295
10.14	Le bureau immersif Workbench de Fakespace Systems, copyright photo Fakespace Systems	296
10.15	La visionstation d'Elumens, copyright photo Immersion S.A.	296
10.16	Les visiondome V5 et V31 d'Elumens, copyright photo Immersion S.A.	297
10.17	Dispositif de visualisation pour l'Immersion Adaptée à un Véhicule, IMAVE (photo Ecole des Mines de Paris)	299
10.18	VRCOVE (Virtual Reality COMmand Visualization Environment), copyright photo VREX	300
10.19	Le Reality Center 3300W à trois canaux de SGI, copyright photo SGI	301
10.20	Salle immersive à écran plat Immersive Workwall de Fakespace Systems, copyright photo Fakespace Systems	302
10.21	Salle immersive avec écran semi-cylindrique et trois vidéoprojecteurs	302
10.22	Illustration du principe du système CAVE	303
10.23	Le SAS ³ , visiocube transportable et multi plate-forme de Clarté à Laval, copyright photo Clarté	304
10.24	La visiosalle reconfigurable Rave de Fakespace Systems, copyright photo Fakespace Systems	304
10.25	Visiocasque AddVisor de Saab Avionics, copyright photo SimTeam	308
10.26	Contraintes technologiques pour la conception d'une interface de vision individuelle	309
10.27	Schéma de principe d'un visiocasque à écrans CRT	311
10.28	Visiocasque à écrans CRT Datavisor 80 de n-vision, copyright photo PSA Peugeot Citroën	311
10.29	Système BOOM HF de Fakespace Labs, copyright photo Fakespace Labs	312
10.30	Système PUSH 640 de Fakespace Labs, copyright photo Fakespace Labs	313
10.31	Schéma de principe d'un visiocasque à écrans LCD	314

10.32	Vidéolunettes à un seul écran pour ordinateur à la ceinture, copyright photo Xybernaut Corporation	318
10.33	Les célèbres Vidéolunettes I-Glasses de Virtual I O, copyright photo ThetaScan	318
10.34	Bureau immersif avec écran semi-transparent de ReachIn Technologies, Copyright photo SimTeam	320
10.35	Positionnement des écrans dans le champ de vision	325
10.36	Vision en relief à partir d'une simple image	327
10.37	Parallaxe positive permettant de visualiser des objets derrière l'écran ; DIO : Distance Intra Oculaire (65 mm en moyenne)	328
10.38	Parallaxe nulle permettant de visualiser des objets au niveau de l'écran .	328
10.39	Parallaxe négative permettant de visualiser des objets devant l'écran . .	329
10.40	Différence entre vision naturelle et restitution stéréoscopique	330
10.41	Mouvements pseudoscopiques quand l'observateur bouge la tête non traquée	331
10.42	Mouvements pseudoscopiques quand l'observateur bouge la tête non traquée	331
10.43	Positionnement des deux caméras fictives ou réelles dont les axes optiques convergent	332
10.44	Les parallaxes verticales sont exagérées sur la figure.	332
10.45	Dispositif de téléopération et de téléprésence stéréoscopique	335
10.46	Zone d'observation possible sans fatigue oculaire en fonction des focales	337
10.47	Fonction de confort oculaire en vision en relief par rapport à la disparité donnée en minute d'angle, Wöpkings, 1995	338
10.48	Images de base pour les différents filtrages (ces images peuvent être vues en relief comme la figure 10.36)	339
10.49	Le calque virtuel : les zones dont la disparité et la fréquence locale sont incompatibles du point de vue de la fonction de confort sont flouées	340
10.50	Le flou adaptatif : là où la disparité et la fréquence locale sont incompatibles, les hautes fréquences ont été supprimées	340
10.51	Schéma montrant la position virtuelle de la planche de bord, vue en relief	342

10.52	En a : champ électrique faible, en b : champ moyen, en c : champ fort	346
10.53	Cheminement de la lumière dans la cellule	346
10.54	R = Retard, Rt = Retard sur tout le long du parcours et L = longueur d'onde	348
10.55	Modèle optique pour l'affichage d'une image d'un visiocasque	349
10.56	Modèle optique d'un visiocasque	350
10.57	Schéma du principe de perception du relief sur une feuille	351
10.58	Coordonnées dans la scène observée	352
10.59	Coordonnées dans l'espace visuel de l'observateur	353
10.60	Coordonnées dans la scène observée	355
10.61	Coordonnées dans l'espace visuel de l'observateur	356
11.1	Gyroscope de Aerotrim, copyright photo Aerotrim	365
11.2	Le Personal Motion Simulator de Flyit Simulation, copyright photo Flyit Simulation	368
11.3	Le simulateur de deltaplane DREAMGlider de Dreamality Technologies, copyright photo Dreamality Technologies	369
11.4	Le micro simulator de Panoram technologies avec 3 écrans, copyright photo Panoram Technologies	370
11.5	Le siège de simulation Cyberchair, copyright photo ViRtogo	370
11.6	Un simulateur de moto	371
11.7	Simulateur d'avion	372
11.8	Simulateur d'engin d'abattage d'arbre Arvester conçu par l'ENSAM et l'Institut de l'Image, copyright photo ENSAM	373
11.9	Simulateur pour les loisirs	373
11.10	Système de déplacement électrique à deux axes de Servos and Simulation, coyright photo Servos and Simulation	374
11.11	Plate-forme à 4 ddl avec des vérins électriques de Moog Inc., copyright photo Moog Inc.	375
11.12	La plate-forme Maxcue de Motionbase, copyright photo Motionbase . . .	375

11.13	Système de mouvement à six axes de Servos and Simulation, copyright photo Servos and Simulation	376
11.14	Cabine de simulation Venturer S2 de AI GROUP (Thomson Entertainment), copyright photo AI Group	377
11.15	La cabine PLC de Motionbase, copyright photo Motionbase	377
11.16	Le Générateur d'Environnement Agressif pour le Travail, GEPAT	379
12.1	L'interface olfactive fixe d'Olfacom sous forme de tour ou de petit diffuseur, copyright photo Olfacom	382
12.2	L'interface olfactive portable Pinoke développée par AromaJet, copyright photo AromaJet	383
12.3	L'interface olfactive d'AromaJet incorporée dans un siège ou dans un visiocasque, copyright photo AromaJet	384
12.4	Le diffuseur d'odeur ScentCartridge et les bornes de diffusion d'odeur de Aerome, copyright photo Aerome	385
13.1	Schéma technocentrique de référence en RV	388
14.1	Etude d'impact sur l'intégration d'un tramway dans un carrefour nantais (image IRISA/Siames).	403
14.2	Solide rigide : un cube.	405
14.3	Solide articulé : un bi-pendule.	405
14.4	Objet déformable : la lampe Luxo se déformant sous son propre poids (image IRISA/Siames).	405
14.5	Maillage régulier en deux dimensions d'un lapin.	407
14.6	Construction d'un quadtree à partir d'un maillage régulier d'un espace à deux dimensions (dans cet exemple, on a modélisé l'espace libre, c'est-à-dire le complémentaire des objets).	408
14.7	Reconstruction du sous sol de la région de Morges (image INRIA/Prisme). .	409
14.8	Exemple d'arbre CSG.	410
14.9	Exemple de pièce modélisée par CSG avec povray	411
14.10	Exemple de pièce modélisée à l'aide d'un maillage triangulaire.	412

14.11	Courbe de Bézier de degré 3, avec ses 4 points de contrôle.	413
14.12	Exemple de construction des triangles à partir de la valuation des sommets des cubes (les sommets valués à 1 sont indiqués par un point noir).	417
14.13	Coupes scanner d'une mâchoire humaine (image Ecole des Mines de Paris/Centre de Robotique)	417
14.14	Reconstruction 3D de la mâchoire à l'aide de l'algorithme des Marching Cubes (image Ecole des Mines de Paris/Centre de Robotique)	418
14.15	Types de configurations de triangles.	423
14.16	Suppression d'un sommet dans un maillage.	423
14.17	Suppression d'une arête dans un maillage, méthode dite du edge collapse.	423
14.18	Représentation de la terre à l'aide de textures.	425
14.19	Représentation des niveaux de détail dans l'arborescence d'un fichier de données 3D	426
15.1	Interpolation entre des images clé	432
15.2	Animation d'un bras de pelleuse	433
15.3	Exemple de plante générée par un L-system	444
15.4	Exemple d'arbre généré par le logiciel AMAP du CIRAD	444
15.5	Les réseaux SAN	447
15.6	Règles de comportement : une nuée d'oiseaux évitant un obstacle	450
15.7	Iowa Driving Simulator : extrait de l'automate de contrôle du conducteur virtuel	453
15.8	Virtual Life : jeu de tennis virtuel	454
15.9	Architecture de contrôle comportemental de Jack	454
15.10	HPTS : architecture	456
15.11	HPTS : automate de fonctionnement	458
15.12	HPTS : exemple de fonction d'intégration	459

15.13	Le système d'acquisition de mouvement commercialisé par la société Vicon, copyright photos Vicon	462
16.1	Risques de traumatisme sonore (D'après Wisner 1967)	472
16.2	Temps de tolérance à la chaleur radiante de parois (d'après Webb 1963)	474
16.3	Zones de confort, d'inconfort et de danger (D'après Nevins et coll. 1966) [Ministère de la défense, 1981]	475
16.4	Baisse de la température centrale et symptômes (d'après Goden 1973) [Ministère de la défense, 1981]	476
16.5	Danger en fonction des valeurs du W.C.I (d'après Rivolier 1981) [Scherrer, 1981]	477
16.6	Tolérance aux accélérations de longue durée (D'après Fraser 1966)	478
16.7	Tolérance aux accélérations + Gz de durée brève [Ministère de la défense, 1981]	479
16.8	Tolérance aux accélérations – Gz de durée brève [Ministère de la défense, 1981]	479
16.9	Modèle analogique de l'homme debout ou assis soumis à des vibrations verticales	480