



**HAL**  
open science

# Chirurgie augmentée : un cas de réalité augmentée ; Conception et réalisation centrées sur l'utilisateur

Emmanuel Dubois

► **To cite this version:**

Emmanuel Dubois. Chirurgie augmentée : un cas de réalité augmentée ; Conception et réalisation centrées sur l'utilisateur. Interface homme-machine [cs.HC]. Université Joseph-Fourier - Grenoble I, 2001. Français. NNT : . tel-00004676

**HAL Id: tel-00004676**

**<https://theses.hal.science/tel-00004676>**

Submitted on 16 Feb 2004

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER – GRENOBLE I  
U.F.R. D'INFORMATIQUE ET MATHEMATIQUES APPLIQUEES

**THESE**

pour obtenir le grade de

**DOCTEUR de L'UNIVERSITE JOSEPH FOURIER - Grenoble I**

Spécialité : **Informatique**

présentée et soutenue publiquement par

**Emmanuel DUBOIS**

Le 6 juillet 2001

---

**Chirurgie Augmentée,  
un Cas de Réalité Augmentée;  
Conception et Réalisation Centrées  
sur l'Utilisateur**

---

Directeurs de thèse : Mlle Laurence Nigay, Mme Jocelyne Troccaz

**Composition du jury :**

Rapporteur :	Mr Alain Derycke
Rapporteur :	Mr Bernard Gibaud
Président :	Mr Philippe Cinquin
Examineur :	Mme Eve Coste-Maniere
Examineur :	Mr Philip D. Gray

Thèse préparée au sein des laboratoires :  
CLIPS : Communication Langagière et Interaction Personne-Système  
TIMC : Techniques de l'Imagerie, de la Modélisation et de la Cognition

Fédération IMAG – Université Joseph Fourier – Grenoble I



## Remerciements

Mes travaux de Thèse, présentés dans ce manuscrit, n'auraient pu aboutir sans l'aide et la présence de nombreuses personnes, que je tiens à remercier très sincèrement ici.

Tout d'abord un grand et très chaleureux Merci à mes deux responsables de Thèse : Laurence Nigay de l'équipe IIHM du laboratoire CLIPS, et Jocelyne Troccaz de l'équipe GMCAO du laboratoire TIMC. Merci pour votre aide, vos conseils, vos encouragements, pour m'avoir guidé, remotivé parfois et surtout pour m'avoir fait confiance et fourni une ouverture sur d'autres travaux de recherche en m'invitant à m'investir dans des projets parallèles.

A l'origine de ces travaux se trouve également un nom que j'ai rencontré au bas d'un livre et qui m'a incité à venir en DEA à Grenoble puis à continuer en thèse. Merci donc à Joëlle Coutaz de m'avoir fait connaître et donné goût à l'IHM puis de m'avoir accueilli dans l'équipe sur un sujet qui n'était pas le sien !

Au cours de ma Thèse, je me suis également appuyé sur des collaborations étroites avec le Centre Hospitalier Universitaire (CHU) de Grenoble et le Laboratoire de Psychologie Expérimentale (LPE) de Grenoble. Je tiens donc à remercier tout particulièrement le Dr. Chavanon du service de chirurgie cardiaque du CHU, ainsi que le Pr. Ohlmann et son équipe du LPE.

Je remercie également le Centre National de la Recherche Scientifique (CNRS) et l'institut d'Informatique et Mathématiques Appliquées de Grenoble (IMAG) pour leur soutien, respectivement par le biais d'une Bourse Doctorale pour Ingénieur (BDI), et par l'intermédiaire du "projet IMAG" intitulé "Chirurgie Augmentée".

Je suis également très reconnaissant envers les membres de mon jury de Thèse. Les questions soulevées au cours de la présentation affinent et complètent les perspectives à donner à ce travail. De plus, les échanges établis avec mes rapporteurs avant la soutenance m'ont permis de compléter et d'éclaircir certains points de mon manuscrit.

Enfin pour m'avoir permis de garder un pied dans le monde vraiment réel et non augmenté, si ce n'est par leur amitié et leurs échanges, merci à chaouki, chris, dave, doms, fanf, fred, gaëlle, gaëtan, galou, laurent, lionel, ninie, olivier, phil, raph, yann, les *BRK* et tant d'autres que je m'excuse d'avoir oubliés !

Un grand Merci également à mes parents et ma famille sans qui je ne pourrais écrire ces remerciements aujourd'hui.

Le dernier mot, tout simplement personnel, est pour Nath pour sa présence à mes côtés, sa patience et sa compréhension, malgré les errances de mon humeur, les nombreux coups de bourre, et finalement mes préférences accordées aux grenouilles et oiseaux nocturnes durant les nuits et matinées d'avril ...



# **TABLES DES MATIERES**

---



<b>Introduction</b>	<b>11</b>
<b>1 Motivations</b>	<b>11</b>
<b>2 Objectifs</b>	<b>12</b>
<b>3 Structure du mémoire</b>	<b>13</b>
<b>Partie A : Les Deux Facettes de l'Espace Problème</b>	<b>17</b>
<b>Chapitre I : Les Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO)</b>	<b>21</b>
<b>1 Introduction aux GMCAO</b>	<b>21</b>
1.1 Objectif des GMCAO	21
1.2 Etapes de fonctionnement des systèmes de GMCAO	22
<b>2 Exemples de systèmes de GMCAO</b>	<b>25</b>
2.1 Systèmes de GMCAO et formes de rendu	26
2.1.1 Rendu visuel	27
2.1.1.1 Des informations de stratégie	27
2.1.1.2 Des informations anatomiques	28
2.1.1.3 La combinaison d'informations anatomiques et de stratégie	32
2.1.1.4 Des informations fonctionnelles	34
2.1.2 Rendu haptique	35
2.1.2.1 Des informations de stratégie	35
2.1.2.2 Des informations anatomiques	37
2.1.2.3 La combinaison d'informations anatomiques et de stratégie	38
2.1.3 Rendu sonore	38
2.2 Systèmes de GMCAO et sources d'informations	39
2.2.1 "Laisser Faire"	40
2.2.2 "Faire"	41
2.2.3 "Faire Faire"	41
2.2.4 "Faire Ensemble"	42
2.3 Synthèse des systèmes de GMCAO	43
<b>3 Approches et limites de la littérature</b>	<b>44</b>
3.1 Classification "Passifs, Semi-Actifs, Actifs"	44
3.2 Avantages et inconvénients	45
<b>4 Bilan de l'analyse des systèmes de GMCAO</b>	<b>47</b>
<b>5 Références</b>	<b>49</b>
<b>Chapitre II : La Réalité Augmentée (RA) : Fusion du Monde Réel et du Monde Virtuel</b>	<b>55</b>
<b>1 Introduction à la RA</b>	<b>55</b>
1.1 Principes fondamentaux	55
1.2 Positionnement de la RA par rapport à la RV	57
<b>2 Exemples de systèmes de réalité augmentée existants</b>	<b>58</b>
2.1 RA : Intégration du monde réel dans le monde virtuel	58
2.1.1 Intégration pour les sorties du système (du système vers l'utilisateur)	59
2.1.2 Intégration pour les entrées du système (de l'utilisateur vers le système)	59
2.1.2.1 Le système Illuminating Light	59
2.1.2.2 Le Tangible Geospace	60
2.2 RA : Intégration du monde virtuel dans le monde réel	61
2.2.1 Systèmes de RA mono-utilisateur	61
2.2.1.1 Le Musée Augmenté	61



2.2.1.2	Le système ARC : "Augmented Reality for Construction"	62
2.2.1.3	Le Paper-Based Audio Notebook	63
2.2.1.4	Le système Haptic interfaced telemanipulation	64
2.2.1.5	Le système Chameleon	64
2.2.1.6	Le Digital Desk	64
2.2.2	Systèmes de RA multi-utilisateurs	65
2.2.2.1	La Gesture Cam	66
2.2.2.2	Le système DOLPHIN	66
2.2.2.3	Les médiaspace	66
2.2.2.4	Les systèmes de vidéo-conférence	67
2.2.2.5	Le jeu de Mah-Jong et le système CAVE	67
<b>3</b>	<b>Approches et limites de la littérature</b>	<b>68</b>
3.1	Deux approches concurrentes	68
3.1.1	Approche du domaine de la synthèse d'images	68
3.1.2	Approches de la communauté IHM	69
3.2	Avantages et inconvénients	71
3.2.1	Diversité des aspects pris en compte	71
3.2.2	Contradictions des conclusions	72
3.2.3	Limitations propres à chaque approche	73
<b>4</b>	<b>Bilan des systèmes de RA</b>	<b>75</b>
<b>5</b>	<b>Références</b>	<b>77</b>
 <b>Chapitre III : Notre Taxonomie des Systèmes Mixtes</b>		<b>83</b>
<b>1</b>	<b>Démarche d'analyse</b>	<b>83</b>
<b>2</b>	<b>Deux caractéristiques intrinsèques aux tâches</b>	<b>83</b>
2.1	Objet de la tâche	83
2.2	Type d'augmentation	86
<b>3</b>	<b>Illustration de notre taxonomie</b>	<b>87</b>
<b>4</b>	<b>Bilan de notre taxonomie</b>	<b>92</b>
<b>5</b>	<b>Références</b>	<b>96</b>
 <b>Partie B: L'espace Solution : Conception et Réalisation</b>		<b>97</b>
 <b>Chapitre IV : ASUR : Support à la Caractérisation des Systèmes Mixtes</b>		<b>101</b>
<b>1</b>	<b>Composants</b>	<b>102</b>
1.1	Utilisateur et Système informatique	102
1.2	Objets réels	102
1.3	Lien entre le monde réel et monde virtuel	102
1.4	Règles d'utilisation et illustration	103
<b>2</b>	<b>Relations</b>	<b>104</b>
<b>3</b>	<b>Caractéristiques de l'interaction Utilisateur-Système Mixte</b>	<b>106</b>
3.1	Caractéristiques des composants	106
3.2	Caractéristiques des relations	108
<b>4</b>	<b>Exemples de descriptions ASUR</b>	<b>113</b>
4.1	Visée Pédiculaire, ACRobot et Craniotomie Assistée par Ordinateur	113
4.1.1	Composants et Relations ASUR	113
4.1.2	Caractéristiques des composants	114
4.1.3	Caractéristiques des relations	115
4.1.4	Apports de la description ASUR	116
4.2	KARMA et le Musée Augmenté	117

4.2.1	Composants et relations ASUR	117
4.2.2	Caractéristiques des composants	118
4.2.3	Caractéristiques des relations	118
4.2.4	Apports de la description ASUR	119
4.3	Tangible Geospace et Digital Desk	119
4.3.1	Composants et relations ASUR	119
4.3.2	Caractéristiques des composants	120
4.3.3	Caractéristiques des relations	121
4.3.4	Apports de la description ASUR	121
<b>5</b>	<b>Intérêts et limitations</b>	<b>122</b>
<b>6</b>	<b>Références</b>	<b>125</b>

### **Chapitre V : Eléments de Conception et de Réalisation de Systèmes Mixtes** **129**

<b>1</b>	<b>Etapas de conception et de réalisation d'un système interactif</b>	<b>129</b>
<b>2</b>	<b>Propriétés ergonomiques de la dualité de l'interaction dans un système mixte</b>	<b>132</b>
<b>3</b>	<b>Démarche d'analyse prédictive de la compatibilité et de la continuité</b>	<b>137</b>
<b>4</b>	<b>Illustration de la démarche</b>	<b>139</b>
4.1	Analyse de KARMA	140
4.1.1	Mise en œuvre de la démarche	140
4.1.2	Conclusion et discussion	141
4.2	Analyse de la Visée Pédiculaire	142
4.2.1	Mise en œuvre de la démarche	142
4.2.2	Conclusion et discussion	143
<b>5</b>	<b>Éléments de réalisation logicielle</b>	<b>145</b>
5.1	Modèle d'architecture de référence	145
5.2	Adaptation du modèle aux systèmes mixtes	146
5.3	Flux d'informations dans le modèle adapté	149
<b>6</b>	<b>Conclusion</b>	<b>150</b>
<b>7</b>	<b>Références</b>	<b>152</b>

### **Chapitre VI : D'un Problème Clinique à la Conception d'un Système : CASPER** **157**

<b>1</b>	<b>Problème clinique</b>	<b>157</b>
<b>2</b>	<b>Conception et développement de la solution informatisée</b>	<b>160</b>
2.1	CASPER : Principe de fonctionnement et matériels utilisés	160
2.1.1	Principe de fonctionnement	160
2.1.2	Matériels utilisés	161
2.1.3	Etapas d'utilisation	163
2.2	Conception de l'interface homme-machine	166
2.3	Bilan	168
<b>3</b>	<b>Positionnement de CASPER dans notre cadre taxinomique</b>	<b>168</b>
<b>4</b>	<b>Analyse de l'interaction dans CASPER</b>	<b>169</b>
4.1	Description ASUR de CASPER	169
4.2	Caractérisation de l'interaction	170
4.2.1	Caractérisation des composants de l'interaction	171
4.2.2	Caractérisation des relations de l'interaction	171
4.3	Analyse ergonomique de l'interaction	173

<b>5 Conclusion</b>	<b>176</b>
<b>6 Références</b>	<b>178</b>

### **Chapitre VII : Nouvelle Version de CASPER :**

#### **Conception, Réalisation et Evaluation Expérimentale \_\_\_\_\_ 181**

<b>1 Re-conception</b>	<b>181</b>
1.1 Incompatibilité et discontinuité perceptuelles	181
1.2 Facteurs d'incompatibilité et de discontinuité cognitives	184
<b>2 Analyse ergonomique prédictive de CASPER-V2</b>	<b>187</b>
<b>3 Réalisation logicielle</b>	<b>191</b>
3.1 Architecture logicielle	191
3.1.1 Pilier informatique	191
3.1.2 Contrôleur de Dialogue	192
3.1.3 Pilier réel	194
3.2 Représentation 3D	196
3.2.1 Représentation dans un référentiel de présentation trajectoire	197
3.2.2 Représentation dans un référentiel de présentation aiguille	198
3.2.3 Représentation dans un référentiel de présentation utilisateur	200
3.3 Calibrage du dispositif d'affichage	204
<b>4 Evaluation Expérimentale</b>	<b>207</b>
4.1 Cadre expérimental	208
4.1.1 Utilisateurs	208
4.1.2 Tâche	208
4.1.3 Dispositif technique	209
4.1.4 Mesures	209
4.2 Méthodologie expérimentale	210
4.2.1 Facteurs ou variables indépendantes	210
4.2.2 Mesures ou variables dépendantes	211
4.2.3 Hypothèses concernant le signal	212
4.3 Analyse statistique du signal global	214
4.3.1 Effets principaux et effets d'interaction	214
4.3.2 Identification des effets	216
4.3.3 Analyse des effets principaux	218
4.3.3.1 Le support	218
4.3.3.2 Le type de présentation	219
4.3.3.3 Le référentiel de présentation	220
4.3.3.4 Les mouvements selon les trois axes	220
4.3.4 Analyse des interactions	221
4.3.4.1 L'interaction entre représentation et axes	221
4.3.4.2 L'interaction entre représentation, axes et référentiel de présentation	222
4.3.4.3 L'interaction entre représentation et référentiel de présentation	224
4.3.5 Bilan de l'analyse du signal global	224
4.4 Analyse statistique de chaque composante du signal	225
4.5 Bilan de l'évaluation expérimentale	226
<b>5 Conclusion</b>	<b>228</b>
<b>6 Références</b>	<b>231</b>

<b>Chapitre VIII : Conclusion</b>	<b>235</b>
<b>1 Résumé de notre contribution</b>	<b>235</b>
<b>2 Perspectives</b>	<b>238</b>
2.1 Extensions à court terme	238
2.1.1 Réalisation logicielle	238
2.1.2 Evaluation	239
2.2 Prolongements	239
2.2.1 Identification du niveau d'abstraction des composants	240
2.2.2 Approche logicielle et composants réutilisables	242
2.2.3 Ouverture à d'autres domaines	242
 <b>Annexe A : Spécifications Technique du Casque SONY LDI-D100/BE</b>	 <b>249</b>
 <b>Annexe B : Instructions Fournies aux Sujets de l'Expérimentation Menée sur CASPER-V2</b>	 <b>253</b>
 <b>Table des Figures</b>	 <b>261</b>
<b>Bibliographie Générale</b>	<b>267</b>



# INTRODUCTION

" Entre  
Ce que je pense,  
Ce que je veux dire,  
Ce que je crois dire,  
Ce que je dis,  
Ce que vous avez envie d'entendre,  
Ce que vous croyez entendre,  
Ce que vous entendez,  
Ce que vous avez envie de comprendre,  
Ce que vous comprenez,  
Il y a dix possibilités qu'on ait des  
difficultés à communiquer.

Mais essayons quand même ..."

*Encyclopédie du Savoir Relatif et Absolu*  
EDMOND WELLS



# Introduction

Nos travaux de recherche ont trait à la conception et à la réalisation de systèmes interactifs exploitant des éléments du monde réel et visant à l'association harmonieuse de ce monde réel et du monde informatique ou virtuel : les systèmes reposant sur le paradigme de la Réalité Augmentée. Nos travaux sont pluridisciplinaires et s'inscrivent dans deux domaines de recherche, l'Interaction Homme-Machine (IHM) et les Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO).

Menée en collaboration entre les équipes de deux laboratoires de l'institut IMAG, l'équipe IIHM (Ingénierie des Interfaces Homme-Machine) du laboratoire CLIPS (Communication Langagière et Interaction Personne-Système) et l'équipe GMCAO (Gestes médico-Chirurgicaux Assistés par Ordinateur) du laboratoire TIMC (Techniques de l'Imagerie, de la Modélisation et de la Cognition), cette thèse repose sur une confrontation et un enrichissement mutuel entre expériences complémentaires.

Les travaux actuels dans le domaine de l'IHM visent à sortir l'ordinateur de son cadre d'utilisation habituel, l'écran, le clavier et la souris posés sur un bureau, pour laisser une part de plus en plus grande à l'interaction avec l'environnement physique de l'utilisateur, c'est-à-dire le monde réel : on parle alors de Réalité Augmentée (RA). Parallèlement, le domaine des systèmes de GMCAO est en pleine expansion et les interventions médico-chirurgicales assistées par ordinateur sont de plus en plus fréquentes dans des spécialités médicales de plus en plus variées. Ces deux constatations se rejoignent, puisqu'un système de GMCAO repose par définition sur le paradigme de RA, et ouvrent la voie au travail que nous présentons dans ce mémoire.

## 1 Motivations

La RA est un paradigme d'interaction qui est né de la volonté de fusionner les capacités de traitement informatique et l'environnement physique. Plusieurs principes de base constituent les fondements de la RA. Elle permet notamment de sortir les capacités de l'ordinateur de son boîtier. L'objectif est de parvenir à rompre la frontière entre l'ordinateur et le monde réel, qui empêche l'entrelacement simple et transparent des informations issues du monde réel avec celles issues du monde informatique. Il s'agit ainsi d'accroître les capacités de l'utilisateur à percevoir les informations, à exécuter des tâches, ou encore à communiquer, le tout sans coupure avec son environnement habituel.

Face aux avancées des technologies de l'interaction, la RA constitue un sujet d'étude d'importance dans le domaine de l'Interaction Homme-Machine (IHM). Le dynamisme de cet axe d'étude se traduit par des avancées brutales. Prenons pour preuve une terminologie mal assise : interfaces tangibles, réalité virtuelle, mixte ou augmentée, environnement et objet augmentés, vidéo augmentée, bit/atomes, etc. Ainsi les facettes de la RA sont multiples. La RA est étudiée pour des applications mono-utilisateur ainsi que collaboratives. Ce paradigme est également exploité pour l'informatique mobile ainsi que l'informatique ubiquitaire. Orthogonalement aux types d'applications étudiés, nous constatons aussi une grande variété dans les domaines d'application de la RA, comme l'architecture ou encore l'archéologie.



Parmi les domaines reposant sur le paradigme de la RA, celui des systèmes de GMCAO est en pleine expansion, car il répond à une demande médicale très forte. La médecine moderne cherche en effet à limiter les conséquences post-opératoires des interventions médico-chirurgicales et à garantir la qualité de ces actions. Les gestes médicaux deviennent de moins en moins "invasifs". Ceci implique que les capacités sensorielles du chirurgien soient prolongées par des capteurs artificiels et que la dextérité humaine soit complétée par des systèmes de guidage précis et sûrs. Les systèmes de GMCAO répondent à cet objectif, en proposant des moyens d'acquisition et de traitement des signaux pertinents, puis des moyens de guidage ou de robotisation du geste. Cependant, la conception des systèmes de GMCAO est principalement guidée par les technologies disponibles : les expériences de conception sont ainsi difficilement capitalisables et les composants logiciels rarement réutilisables. De plus, les systèmes, dont la conception est influencée par les technologies et non centrée sur l'utilisateur, imposent souvent au clinicien d'adapter le protocole chirurgical à des outils parfois conçus dans un contexte étranger aux contraintes spécifiquement médicales. On voit ainsi se multiplier des systèmes dont l'introduction dans les salles d'opération s'avère difficile.

Nous le constatons en GMCAO, chaque domaine d'application adapte la RA à ses besoins et développe des principes de conception ad hoc. La RA constitue donc un domaine évolutif tiraillé entre la pratique créative dirigée par le domaine d'application d'une part, et l'approche conceptuelle et mesurée de la théorie d'autre part. Pratique et théorie sont les deux approches complémentaires pour explorer des voies nouvelles dont on ne cerne pas encore les limites. Ainsi, les pratiques du domaine des systèmes de GMCAO et les théories et modèles du domaine de l'IHM sont les deux piliers essentiels de notre travail et nos objectifs se situent dans ce contexte pluridisciplinaire.

## 2 Objectifs

L'objectif de notre travail est d'établir des éléments utiles à la conception et à la réalisation de systèmes interactifs combinant des entités du monde réel et du monde virtuel. Parmi ces systèmes se trouvent les systèmes de GMCAO, dont la conception souffre du manque de prise en compte de l'utilisateur, mais aussi plus généralement les systèmes de RA, dont la définition même n'a pas trouvé de consensus.

Dans ce but, nous visons trois objectifs intermédiaires :

- Comprendre et structurer l'espace des possibilités des systèmes combinant des entités du monde réel et du monde virtuel. Le cadre de classification et de conception visé doit être centré sur l'utilisateur. Il doit permettre d'identifier et d'organiser les éléments utiles à la conception. Bien que s'appuyant sur les pratiques du domaine des systèmes de GMCAO, la généralité de ce cadre conceptuel est primordiale pour notre étude. En effet, un cadre de définition général et consensuel est nécessaire pour comparer les systèmes existants, fédérer les approches existantes et capitaliser les expériences de conception, en vue de l'établissement de méthodes et principes généraux de conception. Nous estimons que ce premier objectif est une contribution importante au domaine de la RA.

- Fournir des éléments utiles à la conception. S'appuyant sur le cadre de classification, le modèle solution visé doit être un support à l'étude de l'utilisabilité du système conçu. Pour cela nous visons l'intégration des recommandations de l'ergonomie dans un modèle conceptuel de l'interaction qui se doit d'être indépendant d'une technologie particulière.
- Fournir des éléments utiles à la réalisation logicielle. La solution générique visée, utile à l'ingénierie des logiciels, prouvera le caractère opérationnel de notre cadre de conception pour la réalisation logicielle. Pour cela, il est nécessaire d'étudier l'adaptation des modèles d'architecture logicielle afin de prendre en considération les spécificités des systèmes étudiés.

### 3 Structure du mémoire

L'organisation du mémoire reflète notre démarche d'analyse, de l'espace des possibilités aux solutions de conception et de réalisation. La première partie composée de trois chapitres décrit l'existant (chapitres I et II) et son organisation selon notre espace taxinomique (chapitre III).

Les possibilités étant cernées, la deuxième partie du mémoire est composée de quatre chapitres, dédiés aux éléments utiles à la conception et à la réalisation (chapitres IV et V) ainsi qu'à leurs applications à notre système *CASPER* (chapitres VI et VII).

La première partie de notre mémoire dresse un bilan des systèmes interactifs combinant des entités du monde réel et du monde virtuel.

Dans le chapitre I, nous présentons une revue des systèmes de GMCAO. L'essence même de ces systèmes est d'aider le clinicien à réaliser une stratégie opératoire de sorte que la zone d'action soit la plus réduite possible, afin de limiter les lésions causées au patient : il s'agit de chirurgie minimalement invasive. Nous présentons des systèmes développés dans différentes spécialités médico-chirurgicales. L'étude du cadre de classification des systèmes de GMCAO révèle une approche centrée sur les technologies utilisées et leur niveau de participation dans la réalisation de la tâche. Cette étude nous amène alors à envisager une approche de classification prenant mieux en compte l'utilisateur et ses besoins. A ces fins, nous nous tournons vers le domaine de l'Interaction Homme-Machine (IHM) et en particulier la Réalité Augmentée (RA).

Nous présentons dans le chapitre II des systèmes interactifs combinant les mondes réel et virtuel. Certains d'entre eux sont qualifiés de systèmes de RA. Les systèmes présentés soulignent l'étendue des possibilités. L'étude des approches de caractérisation de ces systèmes met en évidence l'absence de consensus.

Devant l'absence d'un cadre unificateur pour ces systèmes, nous fédérons les approches existantes et proposons au chapitre III un cadre taxinomique pour les systèmes mixtes, terme que nous utilisons pour désigner tout système mêlant entités réelles et virtuelles. Cette taxonomie est destinée à classer et comparer des systèmes existants mais aussi à explorer de nouvelles solutions. Nous l'illustrons par les systèmes mixtes présentés dans les chapitre I et II. En particulier elle distingue deux classes de systèmes : les systèmes de Réalité Augmentée, où l'objet de la tâche est une entité réelle appartenant

à l'environnement physique de l'utilisateur et les systèmes de Virtualité Augmentée, où l'objet de la tâche est une entité virtuelle dont la manipulation ou la perception s'appuie sur des outils du monde réel.

Notre taxonomie des systèmes mixtes conclut la première partie de ce mémoire, visant à présenter l'espace des possibilités des systèmes mixtes combinant entités réelles et virtuelles. La deuxième partie de ce mémoire complète ce cadre en présentant nos éléments utiles à la conception ergonomique et à la réalisation logicielle. L'exposé de nos solutions et leurs applications à notre système *CASPER* comprend quatre chapitres.

Dans le chapitre IV, nous complétons notre cadre taxinomique par notre notation ASUR. Il s'agit d'une notation de description permettant d'identifier les entités composant les systèmes mixtes ainsi que de caractériser l'interaction. Deux types d'interaction sont mis en évidence grâce à la notation : l'interaction de l'utilisateur avec le système (monde virtuel) et l'interaction de l'utilisateur avec l'environnement physique (monde réel).

Reposant sur ASUR, le chapitre V présente nos éléments utiles à la conception ergonomique et à la conception logicielle de systèmes mixtes. Pour cela, nous utilisons ASUR et ses caractéristiques pour asseoir la définition et l'étude de nouvelles propriétés ergonomiques spécifiques à ces systèmes. Sur ces bases, nous présentons une démarche d'analyse prédictive ergonomique. Nous traitons ensuite la conception logicielle en exposant notre modèle d'architecture logicielle, une adaptation du modèle d'architecture logicielle PAC-Amodeus pour intégrer les spécificités des systèmes mixtes. Nous concluons ce chapitre en établissant des liens entre notre modèle d'architecture logicielle et nos éléments utiles à la conception ergonomique.

Nous consacrons le chapitre VI à une application de chirurgie cardiaque assistée par ordinateur : *CASPER*. La présentation clinique et technique de *CASPER* est suivie de son positionnement dans notre taxonomie. Nous menons ensuite une analyse ergonomique de l'interaction Clinicien-*CASPER*, basée sur une description du système avec ASUR et sur notre démarche d'analyse ergonomique présentée au chapitre V.

Dans le chapitre VII, nous présentons des solutions de conception permettant de résoudre certains problèmes ergonomiques identifiés lors de l'analyse ergonomique de *CASPER*. Après avoir évalué à nouveau nos solutions en appliquant notre démarche d'analyse ergonomique, nous exposons la réalisation logicielle selon notre modèle d'architecture d'une seconde version de *CASPER* : *CASPER-V2*, basée sur l'utilisation d'un casque semi-transparent. Nous concluons ce chapitre par les résultats de l'évaluation expérimentale de *CASPER-V2*.

Le chapitre VIII conclut nos travaux en dressant un bilan des contributions selon les deux facettes suivantes : conception et réalisation. Nous développons ensuite les perspectives de travaux futurs, concernant *CASPER-V2* à court terme, et visant à la définition d'une boîte à outils pour la réalisation de systèmes mixtes à plus long terme.

Enfin, deux annexes sont fournies en fin du mémoire. L'Annexe A décrit les spécifications techniques du casque SONY LDI-D100/BE utilisé pour *CASPER-V2*, et l'Annexe B consigne les informations fournies aux utilisateurs de *CASPER-V2* lors de l'évaluation expérimentale.



## **Partie A :**

# **LES DEUX FACETTES DE L'ESPACE PROBLEME**

---



# CHAPITRE I

## LES GESTES MEDICO-CHIRURGICAUX ASSISTES PAR ORDINATEUR (GMCAO)



# Sommaire

<b>Chapitre I : Les Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO)</b>	<b>21</b>
<b>1 Introduction aux GMCAO</b>	<b>21</b>
1.1 Objectif des GMCAO	21
1.2 Etapes de fonctionnement des systèmes de GMCAO	22
<b>2 Exemples de systèmes de GMCAO</b>	<b>25</b>
2.1 Systèmes de GMCAO et formes de rendu	26
2.1.1 Rendu visuel	27
2.1.1.1 Des informations de stratégie	27
2.1.1.2 Des informations anatomiques	28
2.1.1.3 La combinaison d'informations anatomiques et de stratégie	32
2.1.1.4 Des informations fonctionnelles	34
2.1.2 Rendu haptique	35
2.1.2.1 Des informations de stratégie	35
2.1.2.2 Des informations anatomiques	37
2.1.2.3 La combinaison d'informations anatomiques et de stratégie	38
2.1.3 Rendu sonore	38
2.2 Systèmes de GMCAO et sources d'informations	39
2.2.1 "Laisser Faire"	40
2.2.2 "Faire"	41
2.2.3 "Faire Faire"	41
2.2.4 "Faire Ensemble"	42
2.3 Synthèse des systèmes de GMCAO	43
<b>3 Approches et limites de la littérature</b>	<b>44</b>
3.1 Classification "Passifs, Semi-Actifs, Actifs"	44
3.2 Avantages et inconvénients	45
<b>4 Bilan de l'analyse des systèmes de GMCAO</b>	<b>47</b>
<b>5 Références</b>	<b>49</b>

Les Gestes Médico-Chirurgicaux Assistés par Ordinateur constituent un domaine en plein essor et cela depuis une quinzaine d'années. Chirurgie Augmentée, Chirurgie Assistée par Ordinateur ("*Computer Assisted Surgery*" - CAS), Interventions Médicales Assistées par Ordinateur ("*Computer Assisted Medical Intervention*" - CAMI) sont d'autres termes désignant le même domaine d'activité : la mise en œuvre d'un ordinateur pour aider dans les pratiques médico-chirurgicales, c'est-à-dire diagnostiques ou thérapeutiques. Cet essor est intimement lié aux avancées et au perfectionnement des technologies. Aussi, l'intérêt que manifestent différents secteurs chirurgicaux à l'égard de tels systèmes est croissant. Tout d'abord utilisés comme aide au diagnostic [Hasman 85, Miller 85], la neurochirurgie est l'un des premiers domaines cliniques où ces systèmes ont été développés afin d'aider le chirurgien à intervenir. En effet, les zones à risques sont très nombreuses et les cibles très petites : l'ordinateur a alors un rôle à jouer pour le geste chirurgical dont la précision requise est extrêmement grande [Lavalée 92]. Désormais les domaines cliniques concernés sont très nombreux : l'orthopédie, la maxillo-faciale, la chirurgie cardiaque, la cardiologie, etc. Le paragraphe 2 présente des systèmes développés pour ces différentes spécialités. Mais il convient auparavant de définir plus complètement ce qu'est un système de GMCAO. Ceci fait l'objet du paragraphe 1 qui décrit les objectifs puis les étapes générales de fonctionnement des systèmes de GMCAO.

## Chapitre I :

# Les Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO)

## 1 Introduction aux GMCAO

### 1.1 Objectif des GMCAO

L'objectif des Gestes Médico-Chirurgicaux Assistés par Ordinateur est clinique : il s'agit d'aider le médecin ou le chirurgien à définir puis réaliser une stratégie médico-chirurgicale optimale. Les gestes ainsi réalisés sont à but diagnostique ou thérapeutique. Le but des systèmes de GMCAO est double : il s'agit tout d'abord de favoriser l'aspect minimalement invasif de l'intervention<sup>1</sup>. Ceci permet de réduire la région de l'action chirurgicale à une zone la plus petite possible autour de celle à traiter. Ainsi les lésions causées au patient sont réduites et par la suite les traumatismes post-opératoires et les temps d'hospitalisation se trouvent réduits. Le second objectif des systèmes de GMCAO est d'accroître la précision du geste par l'apport de moyens complémentaires à ceux dont dispose traditionnellement le clinicien.

Dans le cadre de gestes mini-invasifs, la visibilité de la structure anatomique explorée est diminuée, obligeant le chirurgien à travailler avec une perception limitée des structures sur lesquelles il agit ou qu'il doit éviter. Cette perception peut être artificiellement améliorée par le biais de fibres optiques introduites dans le corps du patient ou grâce à des images médicales telles que des radios ou encore des scanners. De

---

<sup>1</sup> Dans ce mémoire, nous utiliserons plusieurs termes issus du domaine chirurgical. Cependant, les systèmes de GMCAO dépassent le cadre strictement limité à la chirurgie, et intègrent des pratiques diagnostiques aussi bien que thérapeutiques.

plus, le chirurgien est contraint de deviner le tissu qu'il est en train de traverser au lieu de le toucher directement [Cinquin 93]. Il est donc important de définir au préalable une stratégie opératoire basée sur l'exploitation de données médicales multimodales et de connaissances a priori. La fusion de ces données permet alors d'envisager la planification optimale de l'intervention par rapport à un critère donné, comme la position d'une prothèse, mais aussi par rapport aux conséquences fonctionnelles et morphologiques. Le traitement de l'information médicale joue donc un rôle important dans les systèmes de GMCAO. De plus, la planification optimale doit être précisément reproduite. Le geste médico-chirurgical nécessite donc dextérité et précision qui sont diminuées du fait de l'abord mini-invasif.

Devant la richesse de cet objectif clinique, les systèmes de GMCAO prennent tout leur sens : ces systèmes apportent au chirurgien un support au traitement de l'information mais aussi, par le biais de dispositifs électro-mécaniques, le guident lors du geste médico-chirurgical afin de reproduire la stratégie planifiée au plus près. La Figure 1 illustre la relation du chirurgien avec le monde informatique au sein du bloc opératoire.

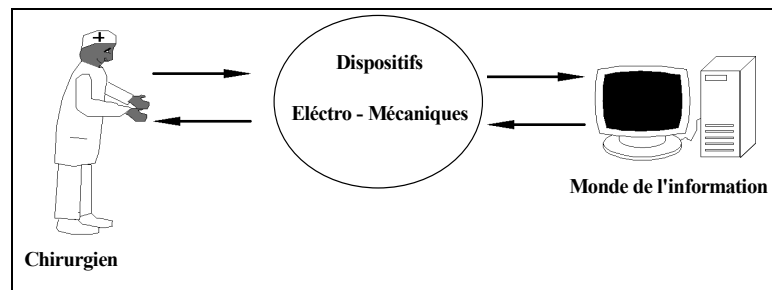


Figure 1 Schématisation de la relation entre le chirurgien et le monde de l'information.

Clinique, novateur et générique, cet objectif implique une problématique scientifique liée à :

- l'acquisition de différents types de données, de sources variées, représentant des informations complémentaires,
- la visualisation et la manipulation de ces données ré-agencées pour assister le chirurgien lors de l'élaboration d'une stratégie opératoire, et
- le guidage de la réalisation de cette stratégie optimale.

Ces trois points correspondent aux trois étapes de fonctionnement d'un système de GMCAO, habituellement dénotées : "Perception - Raisonement - Action". Nous détaillons ces étapes générales au paragraphe suivant.

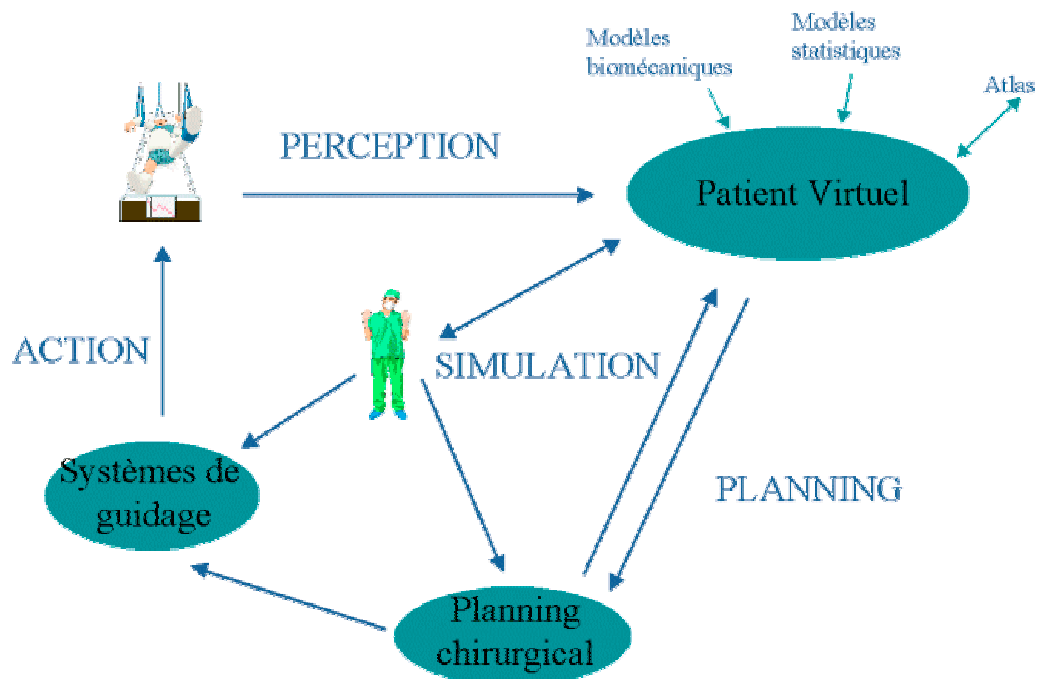
## 1.2 Etapes de fonctionnement des systèmes de GMCAO

Il existe deux approches de définition des étapes générales de fonctionnement d'un système de GMCAO. Ces dernières adoptent des points de vue différents. Tandis que la première est centrée sur les données médicales traitées, la seconde prend en considération l'usage du système par le clinicien. Ces deux approches sont donc complémentaires.

Adoptant le point de vue sur les données, [Evers 98] décrit six étapes de fonctionnement. La première consiste en une phase d'acquisition. Le choix des images à **acquérir** doit être guidé par les aspects fonctionnels et morphologiques d'intérêt pour le diagnostic ou l'intervention. L'acquisition d'images intervient en général au cours d'une phase préopératoire, au cours de laquelle le chirurgien, qui réalisera l'intervention, n'est pas systématiquement présent. Le lieu d'analyse des images n'est donc pas toujours le

même que celui de l'acquisition. De même, un délai temporel entre les phases d'acquisition et d'analyse des images peut apparaître. Par conséquent, une étape de **transmission** d'images intervient entre l'acquisition et l'utilisation de ces images. Cette étape doit prendre en considération ce délai temporel ainsi que la distance géographique entre le lieu d'acquisition et d'analyse des données. Des problèmes de transmission et de numérisation sont courants à ce niveau là du processus. Il convient alors d'appliquer un **pré-traitement** aux images afin d'éliminer toute distorsion et tout artéfact induits principalement par des mouvements du patient lors de l'acquisition et les processus de production et de numérisation des images. Celles-ci sont alors de qualité suffisante pour être manipulées et analysées par le chirurgien. C'est l'étape de **segmentation** au cours de laquelle les zones d'intérêts seront identifiées manuellement ou automatiquement. Différentes images étant utilisées, il est nécessaire de pouvoir superposer dans un même référentiel l'ensemble des informations segmentées sur toutes ces images : c'est la phase de **mise en correspondance**. La dernière étape permet alors la **visualisation** ou plus généralement le rendu, des informations ainsi synthétisées.

L'approche présentée dans [Troccaz 96-a] vise aussi à définir des étapes de fonctionnement des systèmes de GMCAO, mais celles-ci sont liées aux activités successives du ou des cliniciens. Ces étapes forment la boucle "Perception - Raisonement - Action", illustrée par la Figure 2.



**Figure 2** Principes des systèmes de GMCAO : "Perception – Raisonement – Action".

Ce principe en trois étapes débute par l'étape de **perception**. Cette étape a pour but de modéliser le patient afin d'établir un "patient virtuel". Il s'agit donc d'une part **d'acquérir** des images médicales tant fonctionnelles (SPECT, PET, IRM fonctionnelle, etc.) que morphologiques (radiographie, tomographie par rayons X, échographie, endoscopie, etc.) ainsi que des signaux tels que la pression sanguine, un électrocardiogramme ou électroencéphalogramme par exemple. Les données multimodales subissent alors les

étapes de **pré-traitement**, de **segmentation** et de **mise en correspondance** mentionnées par Evers ci-dessus.

A partir de ce patient virtuel et de données complémentaires telles que des atlas anatomiques, des modèles statistiques et/ou tout autre type de connaissance a priori, le clinicien peut définir une stratégie opératoire adaptée : c'est l'étape de **raisonnement**. Cette phase peut être :

- manuelle, ce qui permet au clinicien de tester et modifier sa stratégie plusieurs fois,
- semi-automatique, et dans ce cas, ce sont une ou plusieurs stratégies satisfaisant des contraintes qui sont proposées au clinicien. Ce type de raisonnement est particulièrement adapté dans le cas où les données sont trop complexes à traiter par le clinicien. Par exemple lors de la définition d'un traitement de radiothérapie, irradiant avec une certaine dose une tumeur de forme donnée en évitant au mieux les organes à risques, le système peut proposer une ou plusieurs stratégies satisfaisant les contraintes de dose. Le clinicien réalise ensuite une sélection selon des critères cliniques qui n'auraient pas été intégrés dans le processus d'optimisation de la stratégie.

Une fois la stratégie établie, le système fournit une aide au chirurgien pour l'aider à reproduire cette stratégie. C'est l'étape d'**action**. Plusieurs types d'aide peuvent être fournis pour assister le clinicien : il peut s'agir d'informations visuelles sur les structures anatomiques voisines ou sur la stratégie à reproduire. Il peut s'agir également de robots pouvant contraindre les mouvements de l'outil chirurgical manipulé par le chirurgien ou pouvant réaliser seuls une partie de l'intervention.

Quelle que soit la spécialité médico-chirurgicale du système, les trois étapes "Perception - Raisonnement - Action" sont présentes dans son fonctionnement. Au cours de l'étape d'action, le système reçoit des informations lui permettant de mettre à jour les informations à fournir au clinicien. C'est donc de nouveau une étape de perception pour le système, puis pour le clinicien. Une étape de raisonnement est alors requise par le clinicien pour interpréter correctement les informations fournies et en déduire les nouvelles actions à mener. Les trois étapes forment donc une boucle et ne constituent pas seulement une suite linéaire d'étapes.

Prenons l'exemple d'un système de chirurgie orthopédique. Dans le cas d'une distorsion significative de la colonne vertébrale (scoliose par exemple), il est nécessaire d' "instrumenter" la colonne vertébrale, c'est-à-dire de poser une tige métallique rigide le long de celle-ci pour lui redonner une courbure plus convenable. Cette tige rigide est maintenue en place grâce à un vissage sur les vertèbres. Les vis à positionner sont donc à proximité de structures anatomiques à haut risque telle que la moelle épinière par exemple, que le chirurgien ne peut distinguer lors de l'intervention chirurgicale. L'acquisition (phase de **perception**) d'images scanner en préopératoire permet d'identifier le contour de la vertèbre par segmentation manuelle ou automatique. Le clinicien peut alors définir une position optimale pour les vis à positionner en tenant compte du résultat de la segmentation mais aussi de données médicales complémentaires, telle que la position de vaisseaux sanguins à proximité de certaines vertèbres. C'est la phase de **raisonnement**. Enfin, pendant l'intervention, le repérage de la taille, de la position et de l'orientation de la vertèbre réelle, par palpation de points de la surface de la vertèbre par

exemple, permet d'aligner les données planifiées avec les données réelles. En effet la stratégie a été planifiée sur un modèle de la vertèbre du patient. Ce modèle est basé sur les images médicales de cette vertèbre. Il convient donc, pour rendre perceptible la stratégie dans le bloc opératoire d'établir un lien entre les données de planification et le patient. Un lien envisageable dans ce cas est la vertèbre puisqu'elle est présente sur les images médicales et dans le bloc opératoire. L'acquisition des coordonnées de points constituant la surface de la vertèbre réelle permet de localiser cette vertèbre dans le bloc opératoire. Par le biais d'un changement de repère entre celui de la vertèbre réelle et celui de la vertèbre modélisée, il est possible d'aligner la stratégie planifiée avec la structure anatomique réelle. Le transfert des informations de planification est donc réalisé grâce à la mise en correspondance des données pré- et per-opératoires. Lors de la phase **d'action**, c'est-à-dire l'insertion de la vis dans la vertèbre, l'affichage de données de guidage permet alors au chirurgien de reproduire la stratégie planifiée au plus près.

En présentant les étapes de fonctionnement d'un système de GMCAO, présentée par [Evers 98] ou par [Troccaz 96-a] nous identifions quatre disciplines complémentaires :

- la médecine, pour prendre en compte le patient, sa pathologie et les connaissances médicales nécessaires,
- la physique et le traitement du signal, pour comprendre et savoir acquérir ainsi que manipuler les images médicales,
- l'informatique, pour le traitement des images et leur visualisation, ou plus généralement pour le traitement des informations ainsi que pour l'interaction,
- la robotique, pour concevoir et adapter les dispositifs mécaniques d'interfaces entre le système et le chirurgien.

Ayant cerné les objectifs et les étapes de fonctionnement qui sont générales aux systèmes de GMCAO, il convient maintenant de considérer des exemples de systèmes. Aux paragraphes suivants nous présentons des systèmes en ne considérant que leur phase d'action. Ces systèmes sont soit en cours de développement, soit d'ores et déjà utilisés en bloc opératoire.

## 2 Exemples de systèmes de GMCAO

Comme expliqué au paragraphe précédent, la troisième étape de fonctionnement des systèmes de GMCAO, l'étape d'action, fait intervenir différents dispositifs pour rendre perceptible l'information préopératoire au cours de l'intervention. La nature des informations fournies par ces dispositifs varie d'un système à l'autre : il peut s'agir d'informations relatives à la stratégie opératoire définie, d'informations anatomiques ou encore fonctionnelles relatives à une zone proche de celle de l'intervention. Même si la visualisation de ces informations constitue le cas le plus fréquent, l'usage de sons et de retour d'efforts sont aussi mis en œuvre. Parallèlement, les sources d'informations du système informatique, au cours du déroulement de l'intervention, sont la plupart du temps transparentes aux yeux du clinicien. L'identification de ses mouvements et actions suffisent au système pour en déduire la situation de l'intervention ; à partir de ces mouvements, le système peut procurer les informations ad hoc au clinicien. Néanmoins, certains systèmes requièrent une participation plus prononcée de la part de l'utilisateur.

Cette participation se manifeste sous la forme d'ordres donnés oralement pas le clinicien ou par le biais de transmission de force du clinicien sur le système.

Dans les paragraphes suivants, nous illustrons cette variété des systèmes de GMCAO en illustrant les différents axes présentés dans la Figure 3 ci-dessous. Nous considérerons dans un premier temps la diversité des formes de sorties mises en œuvre par les systèmes de GMCAO. Pour cela nous nous intéresserons aux types de rendus utilisés par les différents dispositifs, ainsi qu'aux types d'informations affichées par ces dispositifs. Puis nous considérerons les différents types de sources d'informations sur lesquels se basent les systèmes de GMCAO.

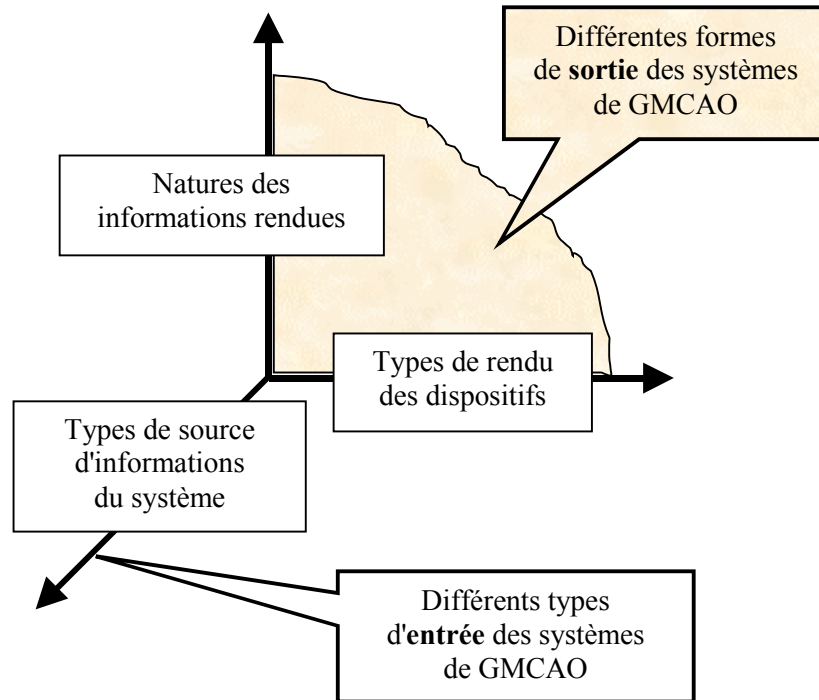


Figure 3 Cadre d'illustration des systèmes de GMCAO.

## 2.1 Systèmes de GMCAO et formes de rendu

Nous avons identifié quatre types principaux d'informations médicales fournies au clinicien. Tout d'abord il est possible d'apporter des informations "prêtes à l'emploi" directement liées à la tâche de l'utilisateur : c'est le cas de l'affichage d'une stratégie à reproduire par exemple. Dans ce cas le clinicien perçoit l'état actuel de l'avancée de l'intervention, par exemple la position d'un outil chirurgical, ainsi que l'état dans lequel il devrait se trouver conformément à la stratégie définie préalablement. Ces informations sont donc des informations relatives à la **stratégie** à reproduire.

Le système peut aussi être moins directif et fournir au chirurgien uniquement des informations qu'il ne peut percevoir directement dans le champ opératoire : dans ce cas, un modèle 3D d'une structure anatomique peut être fourni au chirurgien afin de l'aider à élaborer un modèle mental de la zone à traiter du patient. Il s'agit alors d'informations **anatomiques**.

Le dernier type d'informations fournies se compose d'informations **fonctionnelles** relatives au comportement, le plus souvent dynamique, de structures anatomiques proches ou dans la zone de traitement.

Enfin, notons que les informations anatomiques et de stratégie peuvent naturellement être **combinées**. L'intérêt est de fournir la stratégie à suivre ainsi que des informations anatomiques permettant au clinicien d'estimer plus précisément les risques liés à une déviation éventuelle par rapport à la stratégie optimale.

Nous illustrons ces différents types d'informations en présentant au paragraphe suivant des systèmes de GMCAO existants. Nous mettrons en outre l'accent sur les technologies utilisées pour permettre la présentation de ces informations.

### 2.1.1 Rendu visuel

#### 2.1.1.1 Des informations de stratégie

Un des premiers domaines chirurgicaux à avoir utilisé un ordinateur en bloc opératoire (ou salle d'opération) est celui de la neurochirurgie. Devant intervenir dans des zones réduites, à visibilité faible, il est très utile pour les neurochirurgiens de pouvoir connaître la position de leur outil dans la masse cérébrale, par rapport à des structures cérébrales identifiées sur des IRM par exemple, ou par rapport à une trajectoire à réaliser préalablement définie. Depuis, l'utilisation d'ordinateurs en bloc opératoire visant à présenter des informations de stratégie s'est largement étendue. Ces systèmes sont identifiés sous le nom global de systèmes de navigation : les systèmes développés présentent simultanément la position d'un outil chirurgical ou d'un implant chirurgical, ainsi que la trajectoire ou la position idéale, définie en préopératoire.

Dans un domaine clinique autre, *HipNav* [Simon 97], en est un exemple représentatif. *HipNav* a pour but d'aider le chirurgien à positionner idéalement le composant acétabulaire dans le bassin lors d'une chirurgie de remplacement de la hanche. L'implant fémoral, seconde composante d'une hanche artificielle, est fixé dans le fémur et s'appuie sur l'implant acétabulaire comme le montre la Figure 4. La position de cet implant conditionne la réussite du remplacement de la hanche. Sur la base d'images scanner du pelvis, un modèle 3D du pelvis est présenté au chirurgien qui peut définir manuellement la taille et la position idéales qu'il souhaite donner à l'implant acétabulaire. Lors du positionnement de l'implant sur le patient, un système de guidage est affiché sur un **écran** : il se compose de deux croix mobiles représentant chacune la position en temps réel d'une extrémité de l'implant, et d'une croix fixe représentant la position et l'orientation idéales préalablement définies par le chirurgien. Cette interface de guidage est présentée à la Figure 4. (partie droite).

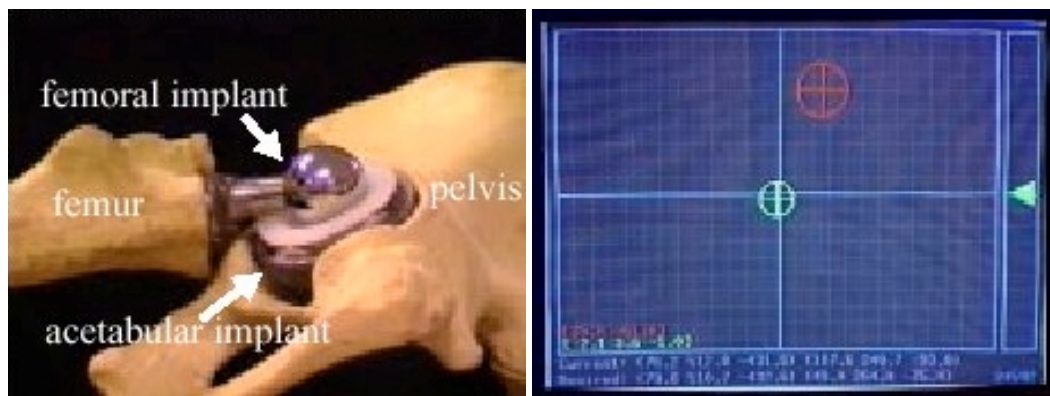
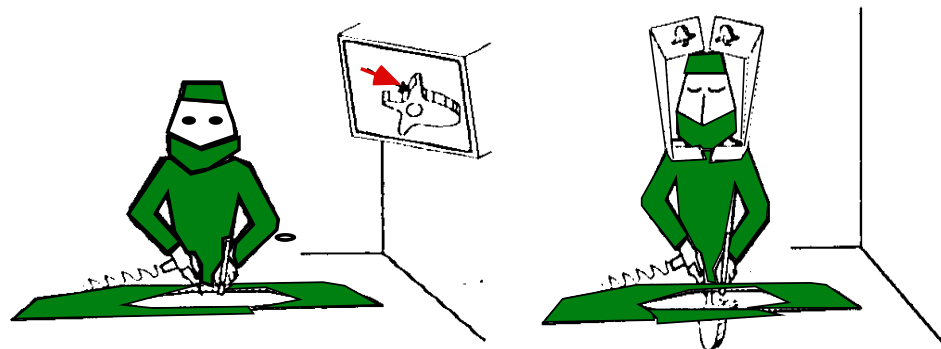


Figure 4 Eléments d'une hanche artificielle (gauche) et interface de guidage de *HipNav* (droite). Photos issues de [Simon 97].



Le système présenté dans [Peuchot 95] est un second exemple de système affichant des informations de stratégie. Dans cette étude, aucun outil n'est suivi. Il ne s'agit donc plus de suivre ou d'indiquer la position d'un outil ni même d'un implant, comme dans le système *HipNav*, mais directement d'indiquer la position idéale que devrait avoir une vertèbre après remise en place de la colonne vertébrale. Cette information est issue de la phase de raisonnement des systèmes de GMCAO : il s'agit d'une information résultant d'une planification de l'intervention chirurgicale. Sans l'aide de ce système, le chirurgien observe sur les images médicales le degré de déviation de la vertèbre et tente de la replacer durant l'intervention. Sa seule possibilité pour vérifier l'exactitude de son geste était une série d'images médicales après l'intervention. Désormais le système de Peuchot lui permet d'avoir sous les yeux et tout au long de l'opération, la position idéale qu'il doit redonner à la vertèbre et la position actuelle de la vertèbre sur laquelle il travaille.

A la différence de *HipNav* qui affichait les informations sur un écran, Peuchot utilise un miroir semi transparent situé au dessus de la zone d'intervention du patient. Le clinicien voit donc la vertèbre réelle au travers du miroir semi-transparent et le reflet de l'information de planning affichée par un écran disposé au dessus du miroir. L'intérêt de cette solution est qu'elle évite au clinicien de devoir regarder alternativement le champ opératoire et l'écran affichant les informations : les deux informations sont superposées au niveau du champ opératoire. La Figure 5 illustre le principe de fonctionnement de ce système.

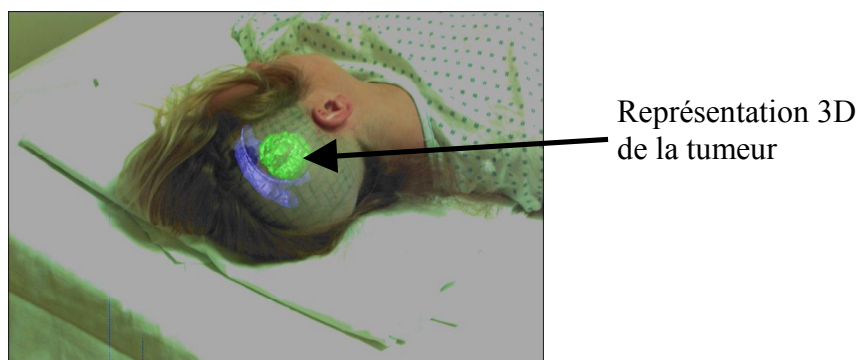


**Figure 5** Schématisation du fonctionnement du système de Peuchot (droite) par rapport à un système de GMCAO basée sur un écran.  
*Schémas issus de [Peuchot 95].*

### 2.1.1.2 Des informations anatomiques

L'évolution des technologies a notamment et considérablement accru les vitesses de calcul et la qualité d'affichage. Grâce à ces évolutions, quelques systèmes de GMCAO procurent au chirurgien directement un rendu volumique 3D des informations médicales acquises précédemment. Ce rendu volumique, représentant la structure anatomique à étudier ou des structures importantes pour le geste (structures à éviter par exemple), demande certes un certain temps de calcul, mais son utilisation procure au chirurgien une sorte de "vue aux rayons X". Il perçoit ainsi des structures anatomiques invisibles à l'œil nu. L'avantage est alors que la phase de planification peut être quelque peu modifiée par le clinicien pendant la réalisation de l'intervention, sans devoir revenir à l'étape explicite de planification (raisonnement) : en effet le système lui présente en temps réel toutes les informations nécessaires à la prise de décision relative à la stratégie opératoire à suivre.

Un des systèmes du MIT [Grimson 96] permet d'optimiser la définition du contour d'une craniotomie. Une craniotomie consiste en l'ouverture du crâne sur une surface définie, afin d'atteindre les structures cérébrales pour retirer une tumeur par exemple. L'objectif est de minimiser la surface de la craniotomie tout en permettant un accès suffisant à la tumeur à traiter. Dans le système de [Grimson 96], l'acquisition d'images IRM permet de reconstruire un modèle 3D de la tumeur. Au cours de l'intervention, une caméra est orientée sur le patient. Superposé au flux vidéo ainsi obtenu, le modèle 3D de la tumeur est affiché. L'alignement de l'image vidéo avec le modèle 3D de la tumeur est assuré par l'utilisation d'un capteur de surface permettant l'acquisition du crâne du patient sur lequel repose la mise en correspondance. La Figure 6 montre la tumeur superposée au flux vidéo du patient. Le chirurgien peut alors marquer sur le crâne du patient le contour suffisant pour accéder à la totalité de la tumeur : il déplace son crayon dans le champ opératoire réel et observe sur un écran de visualisation, la position de son crayon et de la tumeur, ainsi que les parties du contour déjà définies.

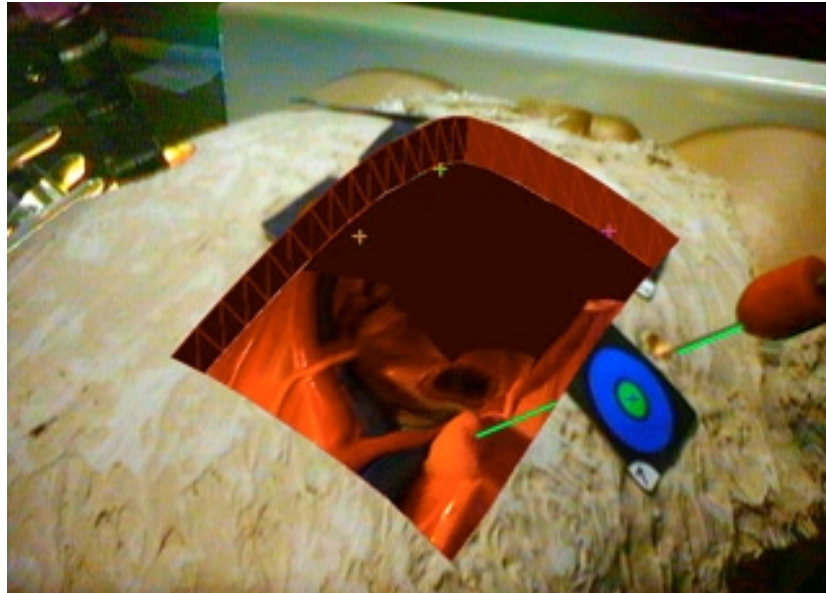


**Figure 6** Informations affichées sur l'écran du système du MIT, optimisant la réalisation d'une craniotomie. Photo issue de [Grimson 96].

Les informations fournies sont donc des modèles 3D de tumeurs et de zones du cerveau particulièrement sensibles. Appelée aussi **vidéo augmentée**, cette forme de visualisation permet au chirurgien de voir au même endroit son outil (le crayon), le patient et des informations permettant de prendre une décision. La grande différence entre le principe de fonctionnement de ce système et celui de [Peuchot 95], est qu'avec le système de [Grimson 96] il n'est plus nécessaire pour le clinicien de regarder le patient directement. En effet tout se passe par l'intermédiaire de l'écran.

Un second exemple illustre la visualisation d'informations anatomiques dans le cadre d'une chirurgie laparoscopique. Classiquement, un trocart est introduit dans le patient, puis une caméra et les outils chirurgicaux sont glissés dans le trocart jusqu'à la zone cible. Le chirurgien manipule alors ces outils en observant sur un écran, l'image que lui envoie la caméra du laparoscope. La solution mise en œuvre par [Fuchs 98], consiste à faire porter au chirurgien un casque opaque. Un modèle anatomique 3D des zones d'intérêt situées à l'intérieur du patient est affiché dans le casque. Pour compléter le modèle 3D, le flux vidéo de la caméra laparoscopique est texturé sur ce modèle 3D. Enfin, l'ensemble est aligné et fusionné avec le champ opératoire complet (patient, outils, assistants, etc.) capturé par une seconde caméra. Il s'agit donc d'une **visualisation virtuelle**. Le chirurgien perçoit donc toutes les informations au même endroit et n'est pas contraint de détourner son regard de son principal centre d'intérêt : le champ opératoire. La Figure 7 illustre le résultat obtenu et perceptible à l'intérieur du casque. L'avantage de l'utilisation d'un

casque opaque est qu'il permet de voir le monde réel sauf aux endroits où les informations virtuelles sont plus importantes. L'usage d'un casque semi-transparent ou d'un dispositif similaire à celui de [Peuchot 95] par exemple, ne permettrait pas d'assurer l'occlusion du monde réel aux endroits où le chirurgien doit percevoir les informations "virtuelles" représentant l'intérieur du patient. Le chirurgien verrait donc une combinaison de l'extérieur et de l'intérieur du patient qui ne l'aiderait pas aussi efficacement que la solution de [Fuchs 98].



**Figure 7** *Vue dans le casque utilisé pour la laparoscopie.*  
*Photo issue de [www.cs.unc.edu/~us/Laparo](http://www.cs.unc.edu/~us/Laparo).*

Cependant, la résolution visuelle du champ opératoire est réduite à celle offerte par le casque ou la caméra alors que dans les précédents systèmes évoqués, le chirurgien pouvait percevoir directement le monde réel. De plus l'utilisation de caméra introduit nécessairement un phénomène de distorsion de l'image qu'il n'est pas toujours aisé de corriger. Enfin, tout utilisateur a tendance à croire ce qu'il voit, d'autant plus facilement que cela lui semble réaliste. Aussi laisser le chirurgien travailler seulement sur une image ou un modèle de son patient peut être discutable du point de vue de la sécurité de fonctionnement.

D'autres systèmes constituent une version adaptée du dispositif utilisé dans une chirurgie traditionnelle pour ajouter des informations relatives à l'anatomie du patient. C'est notamment le cas en endoscopie et laparoscopie. Par exemple dans [Scholz 98], un système permet de rajouter une vue arrière à un endoscope et l'affiche sur le même **écran** que la vue endoscopique normale. Ainsi lors du retrait de l'endoscope le chirurgien peut visualiser les structures anatomiques qu'il va rencontrer et donc adapter en conséquence son axe de retrait ou la force avec laquelle il retire le laparoscope. Dans le domaine de la laparoscopie, [Shahidi 98] présente un laparoscope muni d'une vision stéréoscopique. Ordinairement la vue d'une exploration laparoscopique est monoscopique, mais l'utilisation d'un système informatique enrichit cette vue par la capacité de perception de la profondeur. Enfin, [Frankenthaler 98], [Jolesz 97] et [Mori 95] présentent des systèmes d'endoscopie virtuelle, dont l'image est présentée sur un **écran**. Au lieu d'introduire un endoscope classique, un modèle anatomique 3D du patient est reconstruit à partir

d'images médicales, puis la navigation s'effectue à l'intérieur de ce modèle. Proche de systèmes de simulation, ces trois systèmes permettent néanmoins une exploration dont le but est le diagnostic. L'avantage certain est qu'il permet une plus grande liberté d'examen puisque "l'endoscope virtuel" peut aussi bien explorer l'intérieur que l'extérieur d'une structure anatomique, dans la mesure où celle-ci a été correctement et suffisamment segmentée. De plus cela évite au patient une anesthésie générale, et permet au chirurgien d'examiner le patient sous différents angles, sans risque de lésion d'un organe, et dans des zones que l'endoscope traditionnel n'aurait pu atteindre par manque d'espace.

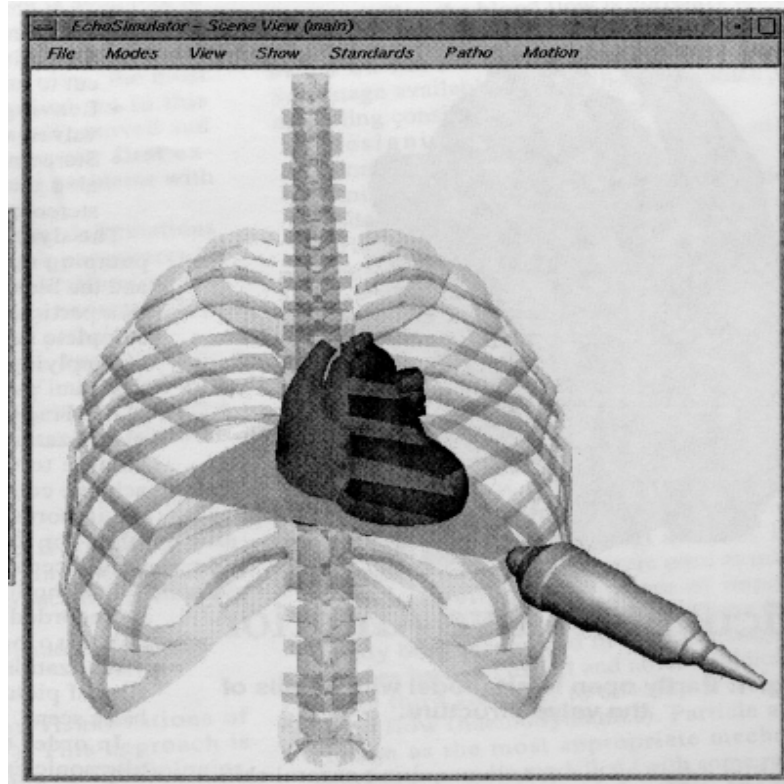
Adapter le dispositif traditionnellement utilisé, est aussi la méthode suivie par [Edwards 95]. Le système de neurochirurgie utilise un microscope, comme souvent en neurochirurgie. Aussi pour ajouter des informations anatomiques à celles perçues dans le champ opératoire au travers du microscope, il a ajouté des LCD semi-transparents au niveau des optiques du microscope. Les informations anatomiques sont alors affichées sur ces LCD et mises en correspondance, c'est-à-dire alignées, avec la vue réelle que le clinicien a au travers du microscope. Comme par défaut le chirurgien regarde au travers du microscope, il perçoit simultanément les informations affichées sur les LCD et le champ opératoire, sans avoir changé son mode opératoire.

Le dernier exemple que nous présentons illustre une forme plus originale d'information anatomique. Destiné à être utilisé par un clinicien en phase d'apprentissage, ce système, appelé *EchoSim* [Berlage 96], facilite la compréhension et l'interprétation d'une image anatomique 2D : une échographie du cœur. Si le principe de ce procédé d'imagerie est relativement simple, l'utilisation des images résultantes l'est beaucoup moins. En effet la bonne compréhension de l'image requiert une certaine habitude, ainsi que la connaissance de paramètres extérieurs : position de la sonde échographique par rapport au patient, connaissance de l'anatomie du cœur, etc. comme le montre la Figure 8.

*EchoSim* ajoute donc une forme particulière de données anatomiques 3D, qui représentent :

- la position de la sonde échographique et son orientation,
- un modèle du cœur et de structures anatomiques voisines comme les cotes par exemple,
- le plan contenant l'image échographique affichée.

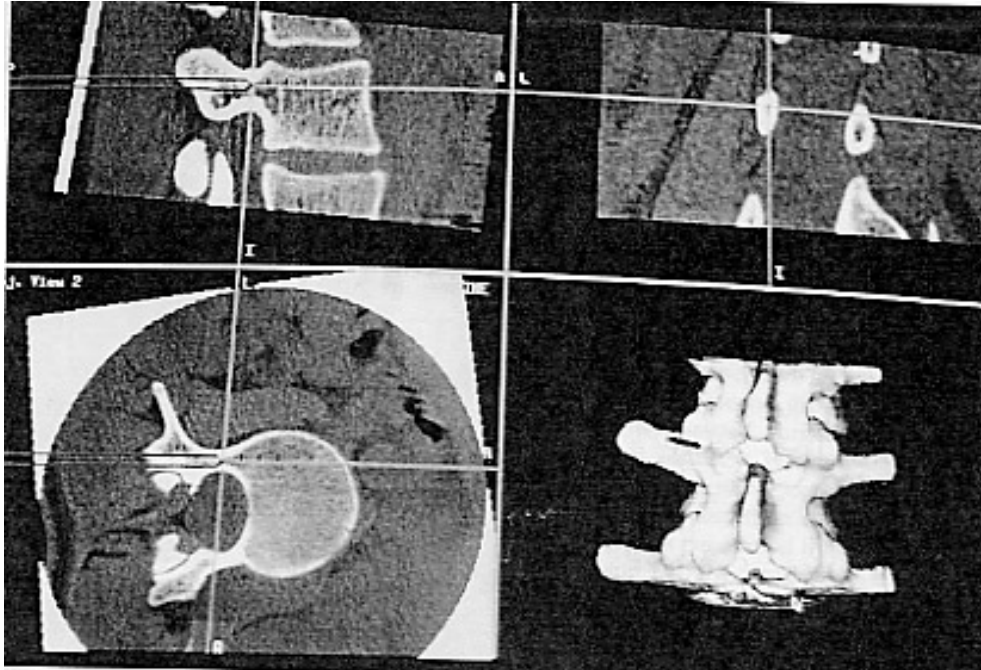
L'image échographique seule étant par nature très parcellaire, sa fusion avec d'autres informations anatomiques procure une information anatomique beaucoup plus complète. La fusion de données anatomiques constitue l'originalité de ce système par rapport aux autres systèmes présentés dans ce même chapitre.



**Figure 8** Données affichées par le système EchoSim.

### 2.1.1.3 La combinaison d'informations anatomiques et de stratégie

Le système de *visée pédiculaire*, très similaire dans le principe d'utilisation au système *HipNav* présenté précédemment, aide le chirurgien à insérer une vis dans le pédicule d'une vertèbre. Une légère déviation de la vis par rapport à une position idéale entraîne le risque de fracture du pédicule, ou plus gravement le risque d'endommager la moelle épinière de la colonne vertébrale. Le guidage fourni permet d'aligner correctement le forêt de la perceuse avec la trajectoire planifiée de la vis. Dans une première version [Merloz 97], le guidage reposait également sur un système de viseur tel que celui de *HipNav*. Néanmoins ce guidage ne permettait pas d'évaluer complètement les risques d'une déviation. Aussi, une seconde version du système [Merloz 00] présente sur un **écran**, la position de l'outil chirurgical par rapport à des données scanner de la vertèbre dans trois coupes différentes (frontale, sagittale et axiale). Sur chacune de ces coupes, la position idéale de l'outil est également affichée. La Figure 9 montre cette interface de guidage qui permet de voir en temps réel la position de l'outil par rapport à la trajectoire planifiée mais aussi par rapport à la structure anatomique visée.



**Figure 9** Informations de guidage pour la visée pédiculaire.

Le système de *visée pédiculaire*, qui comme beaucoup recourt à l'écran, implique donc que le chirurgien regarde alternativement l'écran et le champ opératoire afin de percevoir les informations de guidage d'une part et d'observer les structures anatomiques et son outil d'autre part. Pour pallier cette dissémination des informations, des solutions ont été mises en œuvre pour permettre de superposer les informations virtuelles au champ opératoire. C'est par exemple le cas du système de [Blackwell 98]. Ce système est utilisé pour positionner correctement les composants acétabulaire et fémoral d'une prothèse de hanche. Au cours de cette intervention, aucun microscope ou dispositif comparable au cas précédent n'est utilisé traditionnellement. Amener des informations anatomiques complémentaires au niveau du champ opératoire nécessite donc l'ajout d'un dispositif aussi transparent que possible pour le chirurgien. Le principe de fonctionnement de ce système est très similaire au système de [Peuchot 95] schématisé à la Figure 5. Baptisé "*Image Overlay*" par [Blackwell 98], ce principe se base ici sur la mise en œuvre d'un miroir semi-transparent au travers duquel le chirurgien perçoit son patient. L'image affichée sur un écran disposé au-dessus du miroir se reflète sur ce même miroir. Cette image est composée d'une représentation des informations de guidage et de la position idéale des implants qui permet au clinicien équipé de lunettes polarisées d'avoir une vue 3D stéréoscopique, c'est-à-dire avec perception de la profondeur, des informations. En plus des informations fournies par la première version du système *HipNav* présenté au paragraphe 2.1.1.1, le système indique ici les corrections à apporter à la position des implants en utilisant des flèches en 3D. De plus, la position idéale de l'implant est représentée graphiquement en 3D. La Figure 10 représente une vue schématique du dispositif utilisé.





**Figure 10** Schéma de principe du dispositif utilisé pour l'"Image Overlay" de Blackwell.  
Photo issue de [www.ri.cmu.edu/projects](http://www.ri.cmu.edu/projects).

#### 2.1.1.4 Des informations fonctionnelles

D'autres systèmes se distinguent des précédents par le type de données fournies : il s'agit ici de données fonctionnelles, c'est-à-dire relatives au comportement de structures anatomiques ou à des paramètres dynamiques de ces structures. Par exemple le système présenté dans [Julliard 98] a pour but de montrer à l'écran une carte de positions respectant un certain critère dans le cadre de la reconstruction du ligament antérieur. La position de ce ligament dépend d'un ensemble de modèles qu'il est difficile de gérer sans l'assistance d'un ordinateur. Le rôle du système est donc à la fois de pouvoir prendre en compte l'ensemble des données à considérer, ainsi que fournir des données complémentaires issues de calcul, comme par exemple les variations de longueur supportées par le ligament lors de mouvements de flexion/extension de la jambe. Ces systèmes permettent donc au chirurgien de percevoir plus d'informations qu'une intervention normale ne lui procurerait. Ces informations sont, de plus, utiles pour parfaire l'intervention en cours. Elles permettent donc de fournir des informations inaccessibles par des techniques d'imagerie ou par des examens médicaux classiques nécessaires à la réalisation du planning opératoire lors de l'intervention.

C'est également dans des simulateurs que ce type d'informations peut être ajouté. En effet permettre au clinicien d'apprendre à réaliser un geste sur un simulateur ne le positionne jamais dans un cadre réel d'utilisation du système. [Oppenheimer 01] a proposé un simulateur d'endoscopie dans lequel la représentation de l'écoulement du sang apparaît. Cette représentation est constituée d'un flux vidéo de sang qui est texturé sur l'image de l'endoscope virtuel. Ainsi lors de la réalisation virtuelle d'une intervention sous endoscopie, le réalisme de la scène observée est assuré. Dans ce cas, il ne s'agit donc pas

seulement d'apprendre à reproduire un geste mais aussi de connaître les situations exactes de travail.

Les systèmes que nous venons de présenter proposent tous une présentation visuelle afin de permettre au clinicien de percevoir des informations médicales de différentes natures au cours d'une intervention. Dans le paragraphe suivant, les systèmes présentés se distinguent des précédents par le type de rendu utilisé : il s'agit d'informations tactiles ou plus généralement de retour d'efforts que nous dénoterons par le terme "informations haptiques".

### 2.1.2 Rendu haptique

Le rendu haptique repose en général sur l'utilisation de bras robotisés à retour d'effort. Par exemple, le clinicien agit sur un outil, qui est fixé sur un bras robotisé piloté par un ordinateur. Lorsque le déplacement proposé par le clinicien dépasse une certaine limite l'ordinateur agit sur le bras de sorte qu'il empêche physiquement le clinicien à se déplacer dans cette direction. Ce retour d'effort constitue l'information rendue par l'ordinateur au clinicien. Nous avons choisi ici d'illustrer l'usage du canal haptique de manière plus large, c'est-à-dire non limitée à l'usage de bras robotisés. Nous présentons donc ci-après plusieurs systèmes proposant un retour d'information visuelle complété par des contraintes physiques imposées au clinicien. Nous nous appuyons donc tout comme précédemment sur la nature des informations fournies par le biais du dispositif haptique, et nous mettrons également l'accent sur le niveau de participation du système dans la réalisation de l'intervention.

#### 2.1.2.1 Des informations de stratégie

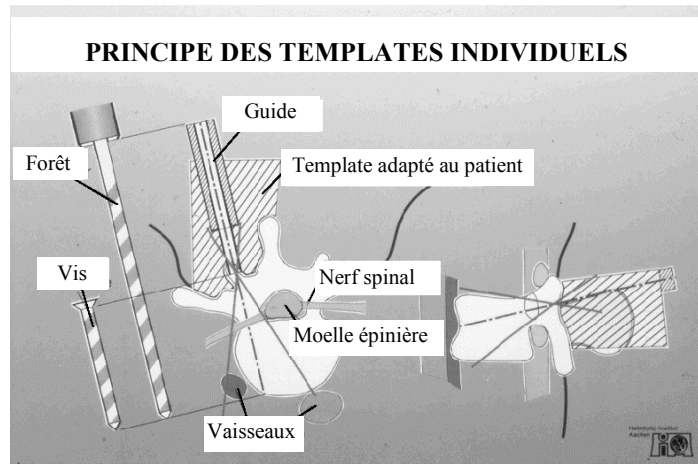
Rendre une stratégie en exploitant le canal haptique humain consiste à matérialiser physiquement la stratégie à réaliser. Une première solution consiste à coupler un guide rigide à un robot qui, piloté par un ordinateur positionnera automatiquement le guide, appelé canule, selon la trajectoire planifiée par le chirurgien. Par exemple, dans le cas d'une fracture du fémur, le chirurgien doit réduire la fracture manuellement puis positionner une vis pour maintenir le fémur dans la position obtenue après réduction de la fracture. La vis doit être positionnée de manière très précise, puisque le volume d'os susceptible de la recevoir est très limité. Traditionnellement fait sur la base d'images médicales, le perçage de l'os en vue du vissage doit parfois être répété pour parvenir à une position plus adéquate. Il en résulte plusieurs trous dans l'os et donc une fragilité de l'os accrue. Le système *Orthosista* [Finlay 95] matérialise dans le champ opératoire la trajectoire optimale. Sur la base d'une trajectoire définie par le chirurgien sur les images médicales, le robot positionne automatiquement un guide dans lequel sera glissé le foret de la perceuse réalisant ainsi le perçage optimal. Le déplacement de l'outil n'est donc que **contraint** par le dispositif physique, et l'initiative des déplacements incombe au clinicien. Le système permet au chirurgien de n'avoir à déplacer son outil que dans une direction, les autres étant contraintes par le guide.

Une autre catégorie de systèmes fait un usage différent des systèmes mécaniques pour matérialiser une stratégie. [Ham 98] rappelle que plusieurs solutions ont été envisagées. Il ajoute également que l'utilisation de robots matérialise plus concrètement les contraintes et les stratégies planifiées, ce qui motive le développement d'autres systèmes de guidage



exploitant un retour haptique. Toutefois, [Ham 98] précise que l'usage de guides portés par un robot, comme *Orthosista*, pallie un certain nombre de problèmes liés à une intervention classique, mais manquent encore de précision. Aussi, des solutions robotisées fonctionnant de manière autonome ont été développées. Dans [Ham 98] un robot permet d'assister le chirurgien pour couper l'os du tibia lors de la pose d'une prothèse de genou. Le chirurgien reste maître des déplacements de l'outil, mais le robot garantit que le plan de coupe obtenu est parfaitement plane. Une des perspectives à ce système permet d'envisager la découpe du tibia selon une forme géométrique plus complexe pour une meilleure adaptation d'une prothèse de genou par exemple. Etant donné que les contraintes géométriques sont exprimées numériquement, le robot peut facilement les mettre en œuvre. L'activité chirurgicale est donc **dynamiquement contrainte** par le système en fonction d'une stratégie définie préalablement par le clinicien. Il est important de noter que le système n'est pas à l'origine des déplacements de l'outil.

Outre les guides mécaniques portés et positionnés par des robots, une autre méthode permettant de matérialiser physiquement une trajectoire consiste à utiliser un gabarit. Par exemple les *Templates* de [Radermacher 97] ou encore le système présenté dans [Fortin 95] reposent tous deux sur l'utilisation d'un gabarit représentant la trajectoire que l'outil doit suivre. Le gabarit a deux fonctions simultanées : il fait la mise en correspondance parce qu'il est le négatif de la structure sur laquelle le clinicien doit intervenir, et il contient la stratégie. Cette stratégie est en fait un trou ou un plan de coupe aménagé dans le gabarit et correspondant à la planification établie. Dans les systèmes du paragraphe précédent, la mise en correspondance est faite par les images et le robot matérialise la stratégie - le gabarit fait les deux simultanément. Dans le cas de [Fortin 95], il s'agit de la trajectoire d'un forêt devant réaliser un trou pour fixer un implant dentaire dans la mâchoire d'un patient. Le moulage dentaire s'adapte à la bouche du patient conformément aux planifications préopératoires. Dans ce gabarit, un trou a été préalablement préparé et correspond à la trajectoire que le forêt doit suivre pour visser idéalement l'implant dentaire. [Radermacher 97] applique ce principe de guidage pour la *visée pédiculaire*. Il propose de réaliser un guide individuel, c'est-à-dire propre au patient à opérer, qui s'adapte à la vertèbre à opérer. Le gabarit de [Radermacher 97] est réalisé à partir des données scanner. Dans ce guide, un trou a été réalisé en suivant les indications données par le chirurgien lors de la phase de raisonnement. Au moment de l'intervention, le *Template* est positionné sur la vertèbre dénudée du patient. Le forêt puis la vis sont ensuite introduits dans le trou réalisé à l'intérieur du *Template*. La Figure 11 illustre le principe de fonctionnement des *Templates*.



**Figure 11** Schéma du principe de fonctionnement des templates.  
Schéma issu de *sttl.hia.rwth-aachen.de*.

Dans ces deux cas, le gabarit agit comme dans *Orthosista* : il **contraint statiquement** le déplacement de l'outil mais le système n'entraîne aucun déplacement de l'outil.

### 2.1.2.2 Des informations anatomiques

Fournir une information anatomique au clinicien par le biais d'un retour haptique consiste à lui faire ressentir la présence d'une structure anatomique à proximité de l'outil que le chirurgien manipule. Illustrons ce principe par un système utilisé en endoscopie. Une endoscopie consiste à réaliser une exploration d'un organe interne au moyen d'une caméra montée sur un endoscope. Cet endoscope est souple ou rigide et est introduit dans le corps du patient. Pour mieux percevoir telle ou telle partie d'une structure anatomique, le chirurgien peut être amené à effectuer une pression sur une structure anatomique afin de la décaler légèrement. De même lors du retrait de l'endoscope, le chirurgien peut exercer une pression sur une structure anatomique sans s'en rendre compte puisque la caméra ne filme que ce qui se trouve devant la tête de l'endoscope. Non seulement, il ne voit pas la structure sur laquelle il est en train d'appuyer, mais en plus il lui est très difficile de percevoir l'intensité de la force exercée. Aussi, le *Force Sensing Retractor* [Poulose 98] pallie ce manque de capacité de perception du chirurgien et l'aide à extraire l'endoscope. Le principe repose sur un calcul par le robot des forces appliquées à différents niveaux de l'endoscope. Le résultat est éventuellement d'empêcher un mouvement de retrait si une des forces appliquées est supérieure au seuil toléré par le système, en débrayant l'endoscope de la poignée que le clinicien manipule par exemple. Dans le cas où le système serait amené à débrayer l'endoscope de la poignée, il agirait de manière **complémentaire** au clinicien. Dans les autres cas, ce système ne fait que compléter les capacités de perception du clinicien.

Le retour de force pour rendre des informations anatomiques est aussi utilisé dans les systèmes destinés à simuler des interventions chirurgicales afin que la situation soit plus réaliste. Les conditions d'apprentissage étant proche des conditions réelles, le temps d'apprentissage sur patient réel et les risques liés à un manque d'expertise sont réduits. Ainsi [Ikuta 98] présente un endoscope virtuel avec retour d'effort. Les chirurgiens peuvent s'entraîner à pratiquer une coloscopie en percevant les contraintes haptiques telles qu'ils les auraient perçues sur un patient réel. Cet entraînement a pour but de les amener à réduire petit à petit les forces qu'ils appliquent lors de la manipulation de

l'endoscope, réduisant du même coup les douleurs causées au patient lors d'une intervention réelle. Ce système vise donc à **apprendre** le geste au chirurgien en situation virtuelle.

### 2.1.2.3 *La combinaison d'informations anatomiques et de stratégie*

Pour le retour haptique il est aussi envisageable de combiner des informations de stratégie et anatomiques. Le dispositif joue alors deux rôles : aider le clinicien à conduire l'outil selon la stratégie prédéfinie et tenir compte de données anatomiques pour éviter des déplacements dangereux.

Un exemple est le système *PADyC* -"*Passive Arm with Dynamic Constraints*", développé par le TIMC à Grenoble et présenté dans [Troccaz 96-b]. Ce système se compose d'un bras articulé dont la vitesse angulaire de chaque raccord est contrôlée par ordinateur. L'outil chirurgical manipulé pour l'intervention est rigidement fixé à l'extrémité du bras articulé de *PADyC*. Innovant de par sa conception technique, n'agit pas directement sur l'outil mais réagit aux mouvements du chirurgien. Enfin les contraintes appliquées par le robot sont programmables. Il propose quatre modes d'expression des contraintes :

- **libre** : aucune contrainte n'est exprimée.
- **position** : les contraintes exprimées aident le chirurgien à atteindre une cible.
- **trajectoire** : les contraintes évitent au chirurgien de quitter une trajectoire prédéfinie, permettant ainsi au chirurgien de reproduire exactement une trajectoire prédéfinie.
- **région** : dans ce cas, le chirurgien peut librement déplacer l'outil, par l'intermédiaire du robot *PADyC*, sans toutefois pouvoir sortir d'une région prédéfinie. Il peut déplacer l'outil en toute liberté en respectant toutefois la présence de zones à risque dont l'accès est interdit.

*ACRobot* [Ho 95] -"*Active Constraint Robot*"- est un système similaire utilisant d'autres solutions techniques. La réaction du robot dépend de la position de l'outil, de la direction du mouvement et des contraintes définies au préalable. Par conséquent, loin des zones à risques, le clinicien a le contrôle intégral du bras robotique et donc de son outil. Par contre à proximité des zones à risques, le robot intervient en imprimant les contraintes physiques nécessaires pour éviter que le chirurgien ne sorte de la zone sûre et peut également dévier le mouvement proposé par le chirurgien afin de produire un mouvement autorisé par les contraintes exprimées précédemment. Ce fonctionnement est donc proche du mode région présenté pour *PADyC* ci-dessus.

### 2.1.3 **Rendu sonore**

Un dernier type de rendu utilisé par des systèmes de GMCAO consiste à utiliser le son pour guider le clinicien. Le système de [Wegner 97] met en œuvre cette forme originale de rendu de l'information. Une déviation de l'outil par rapport à sa trajectoire provoque l'émission de sons dont les caractéristiques telles que la hauteur, le volume et la spatialisation varient. Cette solution permet au clinicien de ne jamais perdre des yeux le champ opératoire. Restant concentré sur son patient, il perçoit en plus les informations relatives à la navigation. De plus, les auteurs soulignent le fait qu'il est plus facile pour le chirurgien de capter des informations sonores en parallèle d'une action qui requiert le sens visuel, que de devoir capter des informations visuelles à deux endroits différents. Néanmoins l'usage du son reste très envahissant pour l'entourage du clinicien et

l'interprétation de l'information n'est pas forcément très aisée et nécessite un temps d'apprentissage et de familiarisation non négligeable.

## 2.2 Systèmes de GMCAO et sources d'informations

Nous venons d'illustrer la diversité des formes de présentation de l'informations fournies par les systèmes de GMCAO. La variété des systèmes de GMCAO s'illustre également par les différents types de sources d'informations captées par le système. Nous étudions ces sources d'informations pour le système en adoptant le point de vue du clinicien et selon deux aspects.

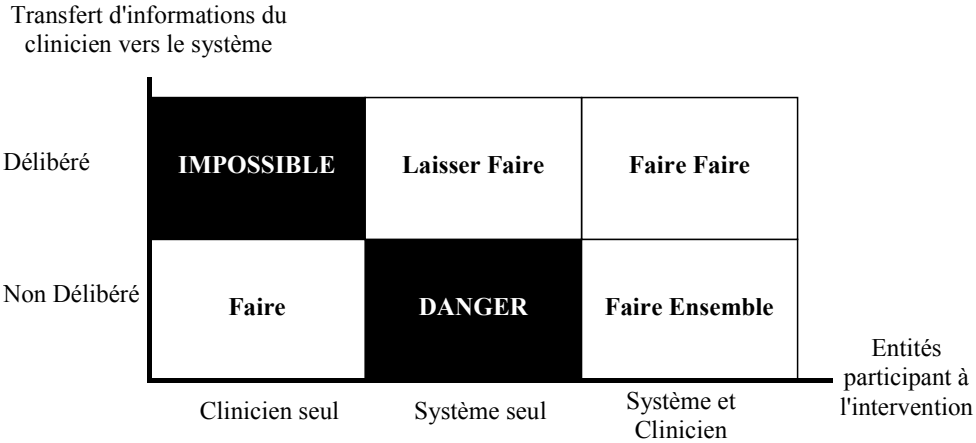
1. Le premier aspect consiste à identifier si l'utilisateur renseigne le système délibérément ou non. Par exemple, lorsque le clinicien saisit une commande au clavier, il fournit délibérément une information au système. A l'opposé, dans la *visée pédiculaire* par exemple, le clinicien ne transmet pas délibérément la position du foret au système. Les diodes fixées sur le foret permettent de le localiser et le clinicien ne fait que déplacer l'outil chirurgical. Le chirurgien n'effectue donc pas d'actions destinées spécifiquement au système (tâche articulatoire) mais effectue sa tâche. Ses mouvements sont source d'informations pour le système.
2. Un second aspect à considérer, concerne la sémantique de l'information fournie par le clinicien aux systèmes. Pour cela nous distinguons trois cas :
  - le clinicien donne des ordres au système qui effectue seul une partie de la tâche en manipulant l'outil chirurgical. Une fois l'ordre énoncé, le système participe seul à la réalisation d'une partie de l'intervention.
  - le clinicien communique avec le système qui réagit et est amené à compléter ou modifier légèrement les actions du clinicien sur le patient. Le système et le clinicien participent tout deux simultanément à la réalisation de l'intervention.
  - le clinicien agit sur le patient et reçoit en complément des informations fournies par le système. Le clinicien agit donc seul.

Il en résulte six types de sources d'informations. Dans le cas où le clinicien serait le seul acteur sur l'outil, il s'agit simplement pour le clinicien de "**faire**" l'action. Dans ce cas, les informations sont transmises de façon non délibérée au système. Transmettre de l'information délibérément au système impliquerait une modification de la tâche. Le clinicien ne se contenterait plus simplement de "faire" ses actions.

Par contre lorsque le système agit seul sur l'outil à la suite d'un transfert d'information explicite, ou délibéré, pour le clinicien, celui-ci "**laisse faire**" la totalité de l'action au système. Si toutefois, le système pouvait agir seul sur l'outil, sans que le clinicien n'ait délibérément donné cet "ordre" au système, il s'agirait là d'une situation dangereuse pour le patient.

Considérons maintenant le cas où les deux agissent sur l'outil. Si l'utilisateur a délibérément demandé au système d'agir conjointement avec lui, alors la forme d'interaction signifie que le clinicien "**fait faire**" une sous-partie de la tâche ou une tâche parallèle au système. Dans le cas contraire, si le clinicien n'a pas délibérément demandé au système de faire une partie de la tâche, cela signifie que les deux acteurs "**font ensemble**" l'action.

Parmi les six types de source d'informations pour le système, quatre seulement sont envisageables pour les systèmes de GMCAO. Ces quatre types et leur traduction en termes de caractérisation de l'interaction induite pour le clinicien sont résumés dans la Figure 12. Nous illustrerons dans les paragraphes suivants les catégories "Laisser Faire", "Faire", "Faire Ensemble" puis "Faire Faire" avec des systèmes de GMCAO existants.

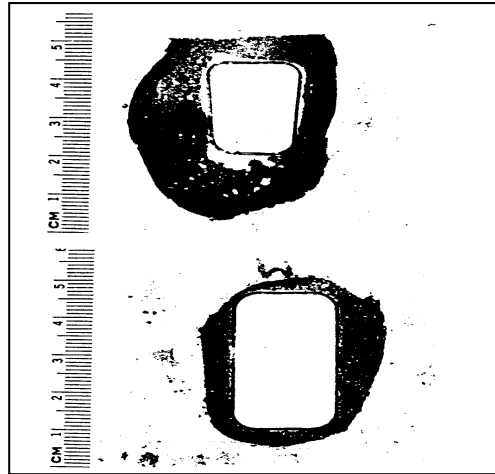


**Figure 12** Types de sources d'informations des systèmes de GMCAO.

### 2.2.1 "Laisser Faire"

Une interaction basée sur le principe du "Laisser Faire" signifie que le clinicien n'a à sa disposition qu'un nombre très réduit de données à fournir au système. Plus précisément il s'agit d'informations de lancement, d'arrêt ou d'arrêt d'urgence. Dans les trois cas, le système fonctionne indépendamment : le clinicien "Laisse Faire" la sous-tâche au système. Concrètement ceci se traduit par la réalisation d'une sous partie de l'intervention chirurgicale de manière totalement autonome par le système. Par exemple *Robodoc* [Mittelstadt 93] est un robot chargé d'usiner une cavité dans l'os du fémur en vue de l'insertion d'une prothèse de hanche. L'utilisation du robot permet, comme l'illustre la Figure 13 ci-dessous, d'obtenir une cavité proprement usinée, dont les contours sont beaucoup plus nets et plus conformes à la forme de la prothèse à positionner, que ceux obtenus lors d'un geste réalisé de manière conventionnelle par le chirurgien. Cet usinage beaucoup plus régulier permet d'envisager la pose de la prothèse de hanche sans utiliser de ciment pour la fixer à l'os. L'absence de ciment limiterait le risque de voir la prothèse se desseller du fémur après un certain nombre d'années.

Le fonctionnement de ce robot est basé sur une définition par le chirurgien de la partie à usiner dans le fémur. Ensuite, après mise en correspondance avec les données scanner grâce à des repères fixés dans l'os et transfert du planning chirurgical, le robot localisé dans l'espace tout comme le patient, usine l'os en respectant les indications données préalablement par le chirurgien. Ce dernier suit sur un écran de contrôle l'exécution de la procédure et peut à tout moment stopper le robot. A titre d'exemple complémentaire nous citons le robot *Minerva* [Glauser 95] qui est chargé de réaliser l'incision, la découpe de l'os crânien, ainsi que la manipulation d'une sonde dans une intervention de neurochirurgie.



**Figure 13** *Différence entre une cavité usinée dans le fémur manuellement (haut) et par Robodoc (bas).*

A la différence des systèmes présentés précédemment, *Robodoc* et *Minerva* agissent donc en autonomie totale. Mais réalisant seuls une partie de l'opération, ces robots ajoutent des risques en terme de sécurité qui imposent des restrictions dans leur utilisation.

### 2.2.2 "Faire"

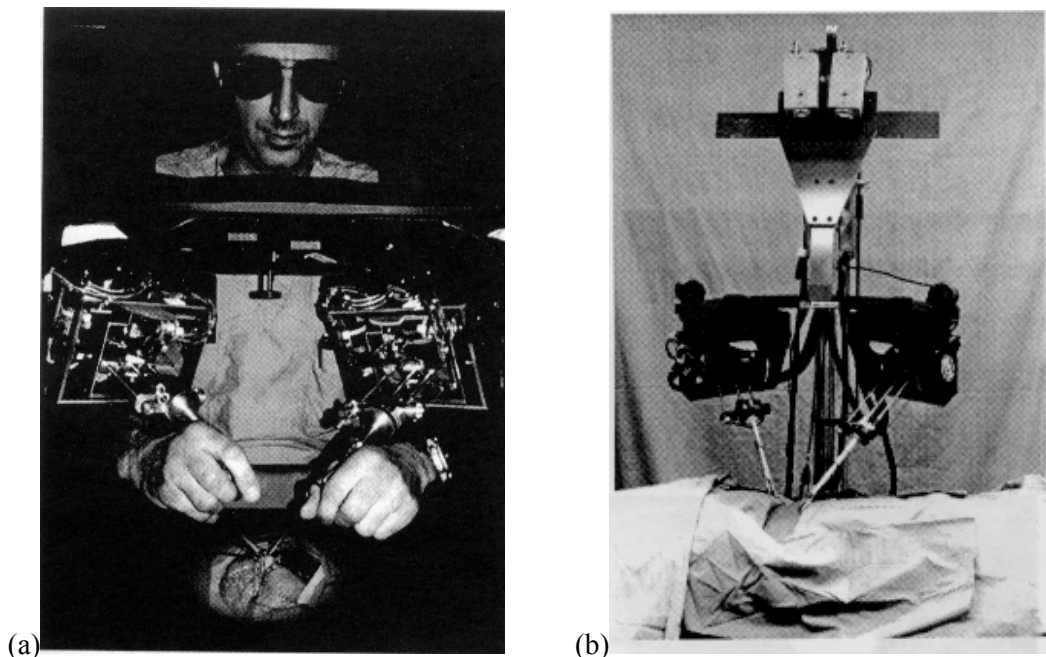
Dans tout système mêlant entités réelles et virtuelles, il est nécessaire de connaître la position des entités réelles pour mettre à jour le rendu des données virtuelles. Il en est de même pour les systèmes de GMCAO qui, quel que soit le système de rendu utilisé, doivent connaître la position de l'outil, du patient, etc. La plupart des systèmes que nous avons présentés aux paragraphes précédents se basent donc sur l'utilisation d'un localisateur 6D qui permet de connaître, par différents procédés, la position et l'orientation d'une entité réelle donnée. Quand ils sont optiques, magnétiques ou ultrasonores, ces localisateurs se composent d'un récepteur et d'un émetteur rigidement fixé à l'entité à localiser (un foret par exemple). Ils peuvent également mettre en œuvre un bras articulé, dont la longueur des segments est connue et les débattements articulaires mesurés en temps réel. En déplaçant ces entités pour intervenir sur le patient, le clinicien communique de manière totalement transparente la nouvelle position de l'entité au système. L'interaction ne demande aucune tâche articuloire particulière. Le clinicien se contente de "Faire" son geste pour fournir des données au système.

### 2.2.3 "Faire Faire"

Au cours d'une intervention chirurgicale, le clinicien est soumis à des contraintes fortes l'empêchant de se déplacer et de manipuler tous les objets présents dans le bloc opératoire. Réduire au strict minimum l'interaction du clinicien avec le système et déléguer le plus possible d'actions articuloires au système constitue le principe du "Faire Faire". Quelques travaux ont par exemple visé à exploiter la parole comme moyen de commandes du système par le clinicien. Ainsi le système *AESOP* ("*Automated Endoscopic System for Optimal Positioning*") [Sackier 93], a servi de support d'intégration d'une interface utilisant la parole comme interface. Dédié à la manipulation de laparoscope, un robot manipule le laparoscope sur ordre du clinicien. Ainsi, l'utilisation du robot libère les mains du clinicien qui n'a plus à déplacer lui-même le

laparoscope. Le langage choisi pour cette interaction est un langage haut niveau et correspond à celui qui serait utilisé entre un clinicien et son assistant. L'objectif est d'éviter au chirurgien l'apprentissage d'un vocabulaire particulier. Le vocabulaire a été volontairement réduit et la grammaire est figée, ce qui permet de fixer le contexte autour des mots. Par exemple, le système peut différencier l'association de mots « à droite » qui lui est destiné, de la phrase « l'outil qui est à droite ». Le clinicien est équipé d'un casque audio avec micro proche de la bouche. Le système fournit un important retour d'information oral, à destination du clinicien et des assistants (haut-parleur supplémentaire placé dans la salle d'opération) : chaque action réalisée est confirmée oralement. Par exemple, le système annoncera « Outil entré », « Limite du champ d'actions atteinte », « Problème mécanique » !

Ce principe est aussi largement utilisé dans les systèmes actuels de télé-opération visant à transmettre à un robot esclave les actions effectuées sur un robot maître. Ainsi, [Green 95] présente un système baptisé *RSU* ("*Remote Surgical Unit*") destiné, entre autres, aux opérations à distance sur un champ de bataille. Le chirurgien agit sur des outils chirurgicaux reliés à un robot à retour d'efforts. Ce robot transmet les informations au système local, qui les achemine au système distant. Le site distant est également muni d'un robot à retour d'effort, capable de transmettre des informations haptiques dans les deux sens (émissions, réception), et, directement orientées sur le champ opératoire, deux caméras filment la scène pour pouvoir la diffuser sous forme stéréoscopique au chirurgien. La Figure 14 ci-après montre à droite la partie du système locale au patient et à gauche le chirurgien en action.



**Figure 14** Exemple de système de télé-opération en fonctionnement.  
(a) : partie locale au chirurgien, (b) partie locale au patient.

#### 2.2.4 "Faire Ensemble"

Pour d'autres systèmes de GMCAO, l'entrée de données par le clinicien requiert la manipulation d'un dispositif. Dans ce cas, le système réagit ou agit en fonction des mouvements explicites du clinicien. Le principe d'interaction est donc ici le "Faire

Ensemble". Il peut s'agir de bras synergiques tels que *PADyC* ou *ACRobot* présentés précédemment, mais ce principe s'applique également aux systèmes de télé-chirurgie réalisant plus que la simple transmission des actions effectuées sur un robot maître. Ainsi un système de télé-chirurgie réalisant un changement d'échelle des mouvements appliqués au robot maître agit avec le clinicien.

### 2.3 Synthèse des systèmes de GMCAO

Le tableau de la Figure 15 résume les différentes possibilités offertes en sortie par un système de GMCAO en regroupant les systèmes présentés dans le paragraphe 2.1. Pour les dispositifs basés sur un rendu visuel, le tableau rappelle les différentes technologies utilisées par les systèmes présentés. Les dispositifs basés sur un retour haptique sont quant à eux complétés par le style de participation du système dans la réalisation de l'intervention. Notre étude ne vise pas l'exhaustivité. Elle souligne la diversité des systèmes de GMCAO existants si l'on considère la forme de sortie des informations fournies par ces systèmes.

Nature de l'information rendue par le système							Type de rendu utilisé par le système
		Retour visuel		Retour haptique		Retour sonore	
Fonctionnelle	Julliard Oppenheimer	Ecran Ecran					
	Fuchs Grimson Shahidi Frankenthaler Jolesz Mori Sholz	Visualisation virtuelle Vidéo augmentée Ecran stéréoscopique Ecran Ecran Ecran	Troccaz Ho	Complémentaire Complémentaire			
Anatomique	Blackwell Berlage Edwards Merloz	Image Overlay Ecran Microscope + LCD Ecran	Ikuta Poulose	Apprentissage Complémentaire			
Stratégie	Peuchot Simon	Image Overlay Ecran	Finlay Fortin Radermacher Van Ham	Contrainte statique Contrainte statique Contrainte statique Contrainte dynamique		Wegner	
		Système	Technologie	Système	Niveau de participation	Système	

Figure 15 Résumé des solutions exploitées par les systèmes de GMCAO pour rendre perceptible de l'information lors de la phase d'action.

Cette variété des systèmes de GMCAO est aussi illustrée par les différents types de source d'informations du système. Le tableau de la Figure 16 illustre les différents types de sources d'information du système.



- Tout système de la Figure 15 présentant un retour visuel	- Mittelstadt - Glauser	- Sackier	- Troccaz - Ho - Green - Fortin - Radermacher
<b>Faire</b>	<b>Laisser Faire</b>	<b>Faire Faire</b>	<b>Faire Ensemble</b>

**Figure 16** Illustration par des systèmes de GMCAO existants des différentes sources d'information possibles lors de la réalisation du geste.

Parmi les systèmes de la Figure 15 présentant un retour haptique, *PADyC* [Troccaz 96-b], *ACRobot* [Ho 95] et le système de [Ham 98] réalisent la tâche conjointement avec le clinicien. Ils "font ensemble" la tâche, de même que les *Templates* de [Radermacher 97] et le système de [Fortin 95]. Tous les systèmes du tableau de la Figure 15 mettant en œuvre un retour d'informations visuelles se basent sur le premier type de source d'information : "faire". Il en est de même pour le système basé sur un retour d'information sonore.

Cependant, *AESOP* [Sackier 93] produit un retour d'informations visuelles anatomiques mais fait partie des systèmes auxquels le clinicien "fait faire" une partie de la tâche. *Robodoc* [Mittelstadt 93] et *Minerva* [Glauser 95] permettent aussi au clinicien de suivre l'avancée de l'outil au moyen d'une représentation graphique affichée sur un écran mais font partie des systèmes auxquels le clinicien "laisse faire" une partie de la tâche.

Prendre en considération le type de source d'information du système est donc complémentaire au type d'informations fournies par le système, notamment dans le cas de rendu visuel.

Devant la très grande variété des systèmes développés dans le cadre des gestes médico-chirurgicaux assistés par ordinateur, une taxonomie seulement est en vigueur pour tenter de dresser des comparaisons entre ces différents systèmes. Cette taxonomie est centrée sur les degrés de passivité de la technologie utilisée pour le guidage. Nous présentons cette taxonomie dans le paragraphe suivant et illustrons les classes identifiées avec les systèmes que nous venons de présenter.

### 3 Approches et limites de la littérature

#### 3.1 Classification "Passifs, Semi-Actifs, Actifs"

La principale classification du domaine caractérise les systèmes du point de vue de la dernière étape de fonctionnement d'un système de GMCAO : la phase d'action, qui est celle que nous avons retenue pour présenter les systèmes de GMCAO dans les paragraphes précédents. Les phases d'acquisition et de raisonnement ne sont pas prises en compte dans la classification de ces systèmes. Cette taxonomie distingue trois types de systèmes selon leur degré d'investissement dans la réalisation de la tâche chirurgicale : les systèmes passifs, actifs et semi-actifs [Lavalée 89].

Avec un système **passif**, l'aide apportée au chirurgien se limite à rendre possible la comparaison entre la stratégie réalisée, par exemple la position de l'outil chirurgical, et la

stratégie idéale, par exemple la position optimale définie lors de la planification. Le chirurgien est donc le seul à agir sur l'outil et peut ou non tenir compte des informations fournies par le système. En ne considérant que le type des sources d'informations des systèmes passifs, il s'agit typiquement des systèmes avec lesquels l'utilisateur "**fait**" la tâche seul. On peut identifier deux types principaux de systèmes passifs : les navigateurs et les systèmes de visée. Dans le premier cas, le système permet la localisation en 6D de l'outil manipulé et le représente sur une image médicale précédemment acquise. Dans le cas des systèmes de visée, le système affiche simultanément la position de l'outil réel et la position de la trajectoire à reproduire.

Parmi les systèmes présentés précédemment, il apparaît très clairement que la première version de *HipNav* (Simon) ainsi que la *visée pédiculaire* (Merloz) sont des systèmes passifs. Le système permettant de définir le contour idéal d'une craniotomie (Grimson), ainsi que le système de visualisation virtuelle (Fuchs) et ceux mettant en œuvre le principe d'"*Image Overlay*" (Blackwell, Peuchot, Edwards) font également partie de cette première classe de système. Enfin les systèmes fournissant des informations fonctionnelles (Julliard) sont aussi des systèmes passifs.

La seconde classe de systèmes regroupe les systèmes dits **semi-actifs**. Ces systèmes contraignent physiquement l'action du chirurgien de façon à l'aider à reproduire le plus exactement possible la stratégie prédéfinie. Néanmoins l'initiative des mouvements appliqués à l'outil est entièrement laissée au chirurgien. Du point de vue du type des sources d'information du système lors de la phase d'action, ces systèmes font référence aux systèmes auxquels le clinicien "**fait faire**" une partie de la tâche ou alors ceux avec lesquels le clinicien et le système "**font ensemble**" la tâche.

Parmi les systèmes présentés précédemment, les systèmes synergiques *Force Sensing Retractor* (Poulose), *PADyC* (Troccaz) et *ACRobot* (Ho) sont des représentants de cette classe. En effet aucun de ces trois systèmes n'est à l'origine exclusive des mouvements de l'outil, mais tous obligent physiquement le clinicien à respecter certaines contraintes. Deux autres systèmes présentés appartiennent à cette classe : les *Templates* (Radermacher) et *Orthosista* (Finlay). L'originalité des *Templates* par rapport aux autres systèmes de cette même classe est qu'ils ne font pas intervenir de robot.

Enfin, la troisième classe rassemble les systèmes **actifs**. Ces systèmes sont capables de réaliser de manière totalement autonome une sous-partie de l'intervention chirurgicale. En ne considérant que le type d'informations fournies au système dans la phase d'action, il apparaît que le clinicien leur "**laisse faire**" la tâche. Leur fonctionnement est supervisé par le chirurgien et une série de capteurs redondants est utilisée pour vérifier à tout instant que la position du robot est bien en correspondance avec la planification de la sous-tâche faite par le chirurgien.

*Robodoc* (Mittelstadt) et *Minerva* (Glaser) sont deux exemples de systèmes actifs. Un système actif peut donc, à l'image de *Minerva*, être chargé de la réalisation de plusieurs sous-tâches successives.

### 3.2 Avantages et inconvénients

La taxonomie "Passifs, Actifs, Semi-actifs" adopte un point de vue suffisamment large pour classer la plupart des systèmes de GMCAO. Nous identifions néanmoins plusieurs

limites. Tout d'abord la classification ne permet pas de clairement positionner les systèmes de télé-chirurgie par rapport aux autres systèmes de GMCAO. En effet tout système de télé-opération requiert nécessairement l'usage d'un robot, que nous qualifions d'esclave, à proximité du patient, ainsi qu'un système maître à proximité du clinicien. Le robot esclave agit donc selon les ordres du clinicien exécutés sur le système maître et envoyés par le système. Dans ce cas, le système esclave fait naturellement partie des systèmes actifs. Pourtant il requiert un ordre du clinicien pour agir. La taxonomie montre donc ses limites, et la prise en compte du type de sources d'information du système en phase d'action apporte des éléments de distinction qui semblent mieux convenir.

Le même problème se pose avec les simulateurs, pour lesquels la phase d'action se déroule sur un patient virtuel. Ils ne réalisent pas une sous partie de la tâche puisqu'ils ne font que fournir de l'information au chirurgien. Ils ne font donc pas partie des systèmes actifs. Ils n'ont pas non plus les caractéristiques des systèmes semi-actifs : à aucun moment ils ne contraignent les mouvements du chirurgien. Ils devraient donc être des systèmes passifs. Mais selon la définition donnée au paragraphe 3.1 ceux-ci ne font que proposer des informations que le clinicien peut choisir de prendre en compte ou non. Pour les simulateurs dispensant de l'information visuelle sans retour d'effort, la classification comme systèmes passifs convient. Par contre si le simulateur fournit une information haptique l'identification comme système passif est remise en cause : le clinicien ne peut plus choisir d'ignorer l'information rendue. Dans ce cas, aucune des trois classes ne convient pour classer ce système.

La classification "Actifs, Passifs, semi-actifs" ne semble donc pas être adaptée aux systèmes particuliers que sont les simulateurs et les systèmes de télé-opérations. Plus généralement, la taxonomie n'est adaptée que pour certains systèmes de guidage. Ceci peut s'expliquer par l'évolution importante qu'a vécue la discipline : lors de la proposition de cette taxonomie, le guidage du clinicien constituait le principal objectif des systèmes de GMCAO. Dans cette optique, les systèmes passifs regroupaient clairement les systèmes non motorisés, à l'opposé des systèmes actifs. Cependant les technologies évoluant, de nouveaux systèmes sont conçus et ne peuvent plus être classés sur la base de cette distinction car le guidage du geste n'est plus le seul objectif des systèmes de GMCAO.

D'autres systèmes identifiés comme étant des systèmes semi-actifs, peuvent de par leur comportement adopter des comportements de systèmes actifs. Par exemple *ACRobot* maintient l'outil déplacé par le chirurgien dans une certaine zone, en s'opposant plus ou moins fortement aux mouvements du clinicien en fonction de la proximité de l'outil chirurgical des zones frontières. Mais dans des cas très proches des frontières, *ACRobot* peut transformer le mouvement proposé par le clinicien de telle manière que l'outil reste à l'intérieur de la zone définie : par exemple au lieu de laisser sortir l'outil de la zone de sécurité, *ACRobot* va lui faire suivre la frontière qui peut être perpendiculaire au mouvement proposé par le clinicien. *ACRobot* ne se contente donc pas de freiner le mouvement, il imprime un mouvement différent de celui proposé par le clinicien. Ce système a priori semi-actif peut donc adopter un comportement actif dans certaines configurations.

*PADyC* présente un problème de classification similaire. Dans le cas d'une utilisation en mode région, lorsque l'utilisateur approche des frontières, *PADyC* se comporte

exactement comme le prévoit la définition des systèmes semi-actifs : il collabore avec le clinicien afin de l'empêcher de sortir de la zone. Par contre lorsque les frontières sont éloignées, le clinicien est totalement libre de faire exécuter les mouvements de son choix à l'outil. *PADyC* se comporte alors typiquement comme un système passif.

Ces deux exemples illustrent donc qu'un système de GMCAO peut appartenir à plusieurs classes de la taxonomie proposée.

Enfin, nous avons choisi de présenter les systèmes de GMCAO en illustrant tout d'abord différentes solutions utilisées pour rendre de l'information, puis les différents types de sources d'informations du système utilisés lors de la phase d'action. Cette distinction n'est pas prise en compte dans la taxonomie "Passifs, Actifs, Semi-Actifs". Cette distinction entre les dispositifs d'entrée et de sortie s'est pourtant révélée intéressante lors du bilan de la présentation des systèmes de GMCAO (cf. paragraphe 2.3). Une approche distinguant dispositifs d'entrée et de sortie est donc complémentaire à une approche basée uniquement sur le niveau de participation du système, comme c'est le cas de la taxonomie "Actifs, Passifs, Semi-actif".

Néanmoins les caractéristiques mises en évidence par cette taxonomie constituent des points clés des systèmes de GMCAO. En précisant le contexte exact d'utilisation de ces termes, leur usage conduit à un système de réflexion cohérent quoiqu'un peu trop influencé par les systèmes existants au moment du développement de la taxonomie. La décomposition en classe mériterait donc d'être plus approfondie et complétée par des aspects permettant de distinguer les dispositifs d'entrée et de sortie mais aussi de prendre en compte les aspects particuliers liés aux systèmes distants et aux simulateurs.

## 4 Bilan de l'analyse des systèmes de GMCAO

L'approche de classification habituelle des systèmes de GMCAO tend donc à s'appuyer sur la technologie mise en œuvre par le système. Ce point de vue se révèle particulièrement restrictif, notamment devant l'émergence des nouvelles utilisations faites des systèmes de GMCAO : opérations à distance et simulateurs.

De plus, la mise en évidence de seulement trois classes distinctes de systèmes, ne facilite pas la comparaison des systèmes entre eux. En effet, tous les systèmes d'une même classe ne répondent pas à des critères bien identifiés. Aussi, lorsque deux systèmes font partie d'une même classe de la taxonomie, il existe encore entre eux des différences importantes. Une conséquence directe de ces moyens de comparaison limités est qu'il n'est pas aisément envisageable d'envisager la conception de nouveaux systèmes en s'inspirant de ceux existants. En effet l'identification des caractéristiques de chaque classe n'étant pas précise, envisager de nouveaux systèmes ne peut pas reposer sur la taxonomie usuelle des systèmes de GMCAO. Or une taxonomie a pour vocation de favoriser la comparaison des systèmes entre eux, mais aussi d'en inférer de nouveaux. Dans le cas des systèmes de GMCAO, la taxonomie est trop restrictive pour permettre ces deux aspects.

Par ailleurs, nous avons mentionné que cette taxonomie est centrée sur le degré d'implication du système dans la réalisation de la tâche. Or nous avons présenté des systèmes, impliqués d'égale manière dans la tâche, qui se distinguaient de par le type d'information échangée entre le système et le clinicien : par exemple son versus

graphique. Il convient donc de savoir privilégier l'une ou l'autre des solutions en fonction des besoins de l'utilisateur.

Recentrer l'analyse des systèmes de GMCAO sur l'utilisateur est donc requis. C'est d'ailleurs dans cet esprit que nous avons analysé les systèmes de GMCAO, en adoptant le point de vue des informations transitant entre le système et le clinicien : à destination du clinicien, ou à destination du système. Ceci permet également de prendre en considération l'utilisabilité des systèmes proposés aux cliniciens. Nous nous tournons donc au chapitre suivant vers le domaine de l'Interaction Homme Machine (IHM) dont l'objectif est l'utilisabilité des systèmes par le biais d'une conception centrée sur les utilisateurs. Dans ce domaine, le paradigme de Réalité Augmentée vise à combiner des entités du monde virtuel et des entités du monde réel. C'est donc ce paradigme que nous prenons comme point de départ pour recentrer les méthodes d'analyse et de classification des systèmes de GMCAO, sur l'interaction entre le clinicien et le système.

## 5 Références

[Berlage 96]

Berlage, T., Fox, T., Grunst, G., Quast, K., J., "*Supporting Ultrasound Diagnosis Using an Animated 3D Model of the Heart*", Acte de la conférence IEEE Multimedia Computing and Systems, Hiroshima, Japon, (1996), 6 pages.

[Blackwell 98]

Blackwell, M., Nikou, C., DiGioia, A., Kanade, T., "*An Image Overlay System for Medical Data Visualization*", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p.232-240.

[Cinquin 93]

P. Cinquin, "*Gestes Médico-Chirurgicaux Assistés par Ordinateur*", Annales de radiologie, Vol. 36 - n° 6/7, Europacs 93, (1993), p. 386-406.

[Edwards 95]

Edwards, P., J., Hill, D., L., G., Hawkes, D., J., Spink, R., Colchester, A., C., F., Strong, A., Gleeson, M., "*Neurosurgical Guidance Using the Stereo Microscope*", Acte de la conférence CVRMed'95, Ayache, N. (ed.), Springer Verlag, Nice, France, (1995), p. 555-564.

[Evers 98]

Evers, H., Glombitza G, Mayer A, Meinzer HP. "*Virtual Surgery: Methods and Areas of Application*", Acte du forum TelecomIT, Frankfurt-am-Main, Allemagne, (1998).

[Finlay 95]

Finlay, P., A., "*Orthosista™: An Active Surgical Localiser for Assisting Orthopaedic Fracture Fixation*", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 203-207.

[Fortin 95]

Fortin, T., Coudert, J., L., Champlébourg, G., Sautot, P., Lavalée, S., "*Computer Assisted Dental Implant Surgery Using Computed Tomography*", dans la revue Journal of Image Guided Surgery, n°1, (1995), p. 53-58.

[Frankenthaler 98]

Frankenthaler, R., Moharir, V., M., Kikinis, R., van Kipshagen, P., Jolesz, F., A., Umans, C., Fried, M., P., "*Virtual Otoscopy*", dans la revue Otolaryngologic Clinics of North America. 31(2), (1998), p. 383-391. Egalement dans SPL Technical Report #76, (1998).

[Fuchs 98]

Fuchs, H., Livingston, M., A., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J., R., Rademacher, P., Drake, S., H., Meyer, A., "*Augmented Reality Visualization for Laparoscopic Surgery*", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p.934-943.

[Glauser 95]

Glauser, D., Fankhauser, H., Epitiaux, M., Hefti, J.,L., Jacottet, A., "*Neurosurgical Robot Minerva. First Results and Current developments*", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 24 -30.

[Green 95]

Green, P., S., Jensen, J., F., Hill, J., W., Shah, A., "*Mobile Telepresence Surgery*", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 97-103.

[Grimson 96]

Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells, W.M., Kikinis R., "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization", dans la revue IEEE Transactions on Medical Imaging, 15(2), (1996), p. 129-140.

[Ham 98]

Ham, Van, G., Denis, K., Sloten, J., V., Audekercke, Van, R., Perre, Van der, G., Schutter, De, J., Aertbelien, E., Demey, S., Bellemans, J., "Machining and Accuracy Studies for a Tibial Knee Implant Using a Force-Controlled Robot", dans la revue Journal of Computer Aided Surgery, n°3, (1998), p. 123-133.

[Hasman 85]

Hasman, A., Lei, J. van der, Kwa, H.Y., Waage, M., "A prototype expert system for the diagnosis and treatment of epilepsy", dans Medical decision making: diagnostic strategies and expertsystems. (Bemmel, J.H. van, Gremy, F., Zvarova, J. eds), North Holland Publ Comp, Amsterdam, (1985), pp. 106-111.

[Ho 95]

Ho, S., C., Hibberd, R., D., Davies, B., L., "Robot Assisted Knee Surgery", dans la revue IEEE Engineering in Medicine and Biology Magazine, Special Issue on Robotics in Surgery, (1995), p. 292-300.

[Ikuta 98]

Ikuta, K., Takeichi, M., Namiki, T., "Virtual Endoscope System with Force Sensation", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p. 293-304.

[Jolesz 97]

Jolesz, F., A., Lorensen, W., E., Shinmoto, H., Atsumi, H., Nakajima, S., Kavanaugh, P., Saiviroonporn, P., Setzer, S., E., Silverman, S., G., Phillips, M., Kikinis, R., "Interactive Virtual Endoscopy", AJR, 169:1229-1235, SPL Technical Report #42, posted March 1997, (1997), 12 pages.

[Julliard 98]

Julliard, R., Lavallée, S., Dessenne, V., "Computer Assisted Reconstruction of the Anterior Cruciate Ligament", dans la revue Clinical Orthopaedics and Related Research, n° 354, (1998), p. 57-64.

[Lavallee 89]

Lavallee, S., "Gestes Médico-Chirurgicaux Assistés par Ordinateur : Application à la Neurochirurgie Stéréotaxique", Thèse de l'Université Joseph Fourier, Grenoble, (1989).

[Lavallee 92]

Lavallée, S., et al., "Image guided Robot : a clinical application in stereotactic neurosurgery", Acte de la conférence IEEE Robotics and Automation, Nice, (1992), p. 618-625.

[Merloz 97]

Merloz, P., Tonnetti, J., Eid, A., Faure, C., Lavallee, S., Troccaz, J., Sautot, P., Hammadeh, A., Cinquin, P., "Computer Assisted Spine Surgery", dans la revue Clinical Orthopaedics and Related Research, n°337, (1997), p. 86-96.

[Merloz 00]

Merloz, P., Lavallee, S., Tonnetti, J., Pittet, L., "Image-Guided Spinal Surgery: Technology, Operative Technique and Clinical Practice", dans la revue Operative Techniques in Orthopaedics, Vol 10, n°1, (2000), p. 56-63.

[Miller 85]

Miller RA, Pople HE, Myers JD. INTERNIST-1, "*An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine*", dans Reggia J and Tuhim S (eds), *Computer-Assisted Medical Decision-Making*, New York: Springer-Verlag (1985).

[Mittelstadt 93]

Mittelstadt, B., Paul, H., Kazanzides, P., Zuhars, J., Williamson, B., Pettitt, R., Cain, P., Kloth, D., Rose, L., Musits, B., "*Development of a Surgical Robot for Cementless Total Hip Replacement*", dans la revue *Robotica*, vol. 11, Cambridge University Press, (1993), p. 553-560.

[Mori 95]

Mori, K., Hasegawa, J., I., Toriwaki, J., I., Anno, H., Katada, K., "*Automated Extraction and Visualization of Bronchus from 3D CT Images of Lung*", Acte de la Conférence CVRMed'95, Nice, France, (1995), p. 542-548.

[Oppenheimer 01]

Oppenheimer, P., Gupta, A., Weghorst, S., Porter, J., R., Sweet, R., "*The Representation of blood flow in Endoscopic Surgical Simulations*", Acte de la conférence MMVR'01, Newport Beach - CA, USA, (2001), p. 365-371.

[Peuchot 95]

Peuchot, B., Tanguy, A., Eude, M., "*Virtual Reality as an Operative Tool During Scoliosis Surgery*", Acte de la conférence CVRMed'95, Nice, France, (1995), p. 549-554.

[Poulose 98]

Poulose, B., Kutka, M., Mendoza-Sagaon, M., Barnes, A., Yang, C., Taylor, R., Talamini, M., "*Human Versus Robotic Organ Retraction During Laparoscopic Nissen Fundoplication*", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p. 197-206.

[Radermacher 97]

Radermacher, K., Portheine, F., Zimolong, A., Eichhorn, C., Staudte, H., W., Rau, G., "*Image Guided Orthopedic Surgery Using Individual Template*", Acte de la conférence CVRMed-MRCAS'97, Grenoble, France, (1997), p. 606-615.

[Sackier 93]

Sackier, J., et al., "*Computer-Aided Guidance for Endoscopic Surgery*", American College of Surgeons, San Francisco, (1993).

[Scholz 98]

Scholz, M., Konen, W., Tombrock, S., Fricke, B., Adams, L., von Düring, M., Hentsch, A., Heuser, L., Harders, A., G., "*Development of an endoscopic navigation system based on digital image processing*", dans la revue *Journal of Computer Aided Surgery*, Volume 3, Issue 3, (1998), p. 134-143.

[Shahidi 98]

Shahidi, R., Wang, B., Epitiaux, M., Grzeszczuk, Adler, J., "*Volumetric Image Guidance Via a Stereotactic Endoscope*", Acte de la Conférence MICCAI'98, Cambridge MA, USA, (1998), p. 241-252.

[Simon 97]

Simon, D. A., B. Jaramaz, M. Blackwell, F. Morgan, A.M. DiGioia, E., Kischell, B. Colgan, T. Kanade, "*Development And Validation Of A Navigational Guidance System For Acetabular Implant Placement.*", Acte de la conférence conjointe CVRMed / MRCAS, (1997), 583-592.

[Troccaz 96-a]

J. Troccaz, S. Lavallée, P. Cinquin, "*Computer augmented surgery*", *Human Movement Science* 15, (1996), p. 445-475.



[Troccaz 96-b]

Troccaz, J., Delnondedieu, Y., "*Semi-Active Guiding Systems in Surgery. A Two-DOF Prototype of the Passive Arm with Dynamic Constraints (PADyC)*", dans la revue *Mechatronics*, Vol 6, n°4, Elsevier Science Ltd., (1996), p. 399-421.

[Wegner 97]

Wegner, Kristen and Karron, D.B. "*Surgical Navigation Using Audio Feedback.*" Acte de la conférence MMVR'97, San Diego, (1997).

# CHAPITRE II

## LA REALITE AUGMENTEE : FUSION DU MONDE REEL ET DU MONDE VIRTUEL

## Sommaire

<b>Chapitre II : La Réalité Augmentée (RA) : Fusion du Monde Réel et du Monde Virtuel</b>	<b>55</b>
<b>1 Introduction à la RA</b>	<b>55</b>
1.1 Principes fondamentaux	55
1.2 Positionnement de la RA par rapport à la RV	57
<b>2 Exemples de systèmes de réalité augmentée existants</b>	<b>58</b>
2.1 RA : Intégration du monde réel dans le monde virtuel	58
2.1.1 Intégration pour les sorties du système (du système vers l'utilisateur)	59
2.1.2 Intégration pour les entrées du système (de l'utilisateur vers le système)	59
2.1.2.1 Le système Illuminating Light	59
2.1.2.2 Le Tangible Geospace	60
2.2 RA : Intégration du monde virtuel dans le monde réel	61
2.2.1 Systèmes de RA mono-utilisateur	61
2.2.1.1 Le Musée Augmenté	61
2.2.1.2 Le système ARC : "Augmented Reality for Construction"	62
2.2.1.3 Le Paper-Based Audio Notebook	63
2.2.1.4 Le système Haptic interfaced telemanipulation	64
2.2.1.5 Le système Chameleon	64
2.2.1.6 Le Digital Desk	64
2.2.2 Systèmes de RA multi-utilisateurs	65
2.2.2.1 La Gesture Cam	66
2.2.2.2 Le système DOLPHIN	66
2.2.2.3 Les médiaspace	66
2.2.2.4 Les systèmes de vidéo-conférence	67
2.2.2.5 Le jeu de Mah-Jong et le système CAVE	67
<b>3 Approches et limites de la littérature</b>	<b>68</b>
3.1 Deux approches concurrentes	68
3.1.1 Approche du domaine de la synthèse d'images	68
3.1.2 Approches de la communauté IHM	69
3.2 Avantages et inconvénients	71
3.2.1 Diversité des aspects pris en compte	71
3.2.2 Contradictions des conclusions	72
3.2.3 Limitations propres à chaque approche	73
<b>4 Bilan des systèmes de RA</b>	<b>75</b>
<b>5 Références</b>	<b>77</b>

Le chapitre précédent a présenté le domaine de la chirurgie assistée par ordinateur et plus généralement des Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO). Nous avons mis en évidence que l'approche de classification est principalement guidée par la technologie utilisée et que le nombre limité de classes, trois, ne permet pas une comparaison très complète des systèmes entre eux. En particulier, les caractéristiques pouvant avoir une influence sur l'interaction du clinicien avec le système n'interviennent pas dans la taxonomie des systèmes de GMCAO. Aussi nous présentons dans ce chapitre des systèmes interactifs du domaine de l'Interaction Homme-Machine (IHM), domaine dont l'objectif est l'utilisabilité des systèmes, c'est-à-dire la qualité de l'interaction Utilisateur-Système. Nous illustrons dans un premier temps la diversité des systèmes combinant l'utilisation d'entités du monde réel et du monde virtuel puis, dans un deuxième temps différentes approches de caractérisation de ces systèmes. Une grande partie de ces systèmes sont qualifiés de systèmes de Réalité Augmentée (RA). Nous rappelons donc tout d'abord les concepts principaux de la RA.

## **Chapitre II :**

# **La Réalité Augmentée (RA) : Fusion du Monde Réel et du Monde Virtuel**

## **1 Introduction à la RA**

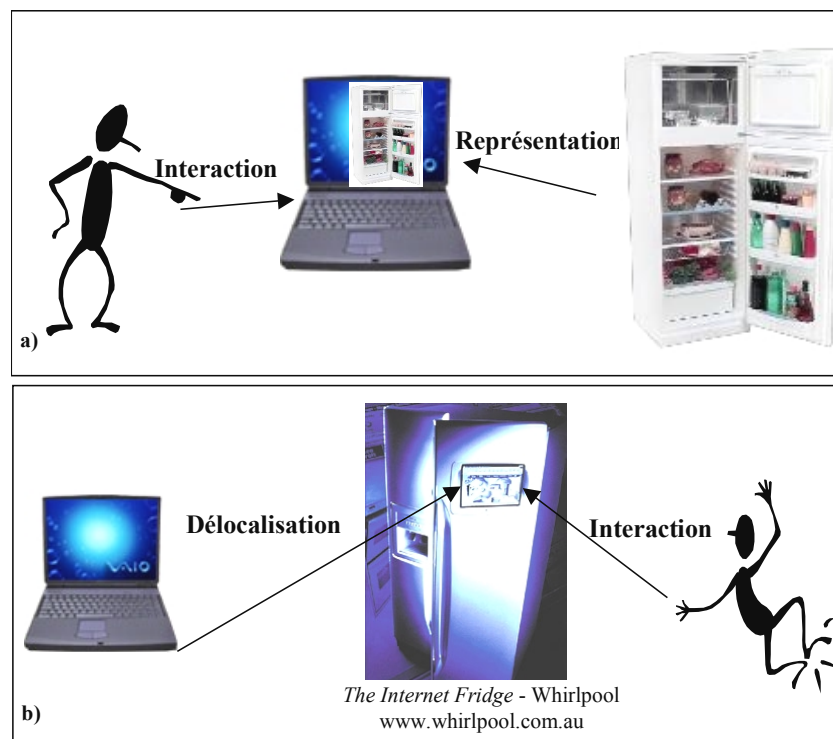
### **1.1 Principes fondamentaux**

L'Interaction Homme-Machine (IHM) se situe habituellement entre un utilisateur et un ordinateur apparaissant souvent comme une boîte noire, dont les entrées et sorties sont essentiellement limitées à un écran, un clavier et une souris. Les périphériques utilisés sont donc à la frontière entre une boîte noire et l'utilisateur. Désormais, l'utilisateur est amené à utiliser des entités de son environnement, en plus de cette boîte noire et de ces périphériques. Il doit donc partager son temps entre deux mondes :

- le monde réel, c'est-à-dire son environnement physique, composé d'objets physiques ;
- le monde virtuel, c'est-à-dire les moyens informatiques mis à sa disposition qui regroupent des moyens de traitements, des moyens de stockage mais aussi des données informatiques.

Le paradigme de la Réalité Augmentée (RA) a pour objectif de favoriser l'interaction de l'utilisateur avec les deux mondes. Ainsi la RA vise à offrir à l'utilisateur la possibilité de bénéficier de moyens informatiques, tout en lui permettant de rester au contact de son environnement réel. La RA privilégie donc les moyens d'interaction naturels de l'utilisateur [Wellner 93], par exemple ses outils habituels dans son environnement réel. Ainsi, en sortant les capacités de l'ordinateur de sa boîte noire habituelle, la frontière entre le monde virtuel, c'est-à-dire les outils informatiques, et le monde réel, dans lequel l'utilisateur maîtrise un grand nombre de moyens d'interaction, tend à disparaître. L'entrelacement simple et transparent des données issues des deux mondes est donc favorisé.

Considérons l'exemple d'un réfrigérateur. Pour gérer informatiquement la liste des courses de produits frais, l'utilisateur devra se détourner du réfrigérateur pour travailler sur son ordinateur. Au mieux il ne disposera que d'une copie de la réalité sur son écran, par exemple une image ou un flux vidéo montrant le contenu de son réfrigérateur si l'on suppose qu'une caméra est située à l'intérieur du réfrigérateur. L'utilisateur complète alors sur ce même écran sa liste de courses. Il est donc contraint à une indirection comme l'illustre la Figure 17.a. A l'opposé, en intégrant des moyens de gestion de la liste de courses au cœur même du réfrigérateur réel, constituant ainsi un "réfrigérateur augmenté", l'utilisateur reste en contact du réfrigérateur réel tout en bénéficiant des capacités de stockage et de manipulation des données, fournies par l'ordinateur (Figure 17.b). Dans le cas du réfrigérateur augmenté, un ordinateur peut être inséré dans la porte de sorte que l'utilisateur établisse sa liste de courses en restant près du réfrigérateur. Ce principe est en phase de concrétisation chez Whirlpool. En effet le *iHome* [Whirlpool 01] est un réfrigérateur dans lequel sont intégrés un écran et un ordinateur, qui permettent aux utilisateurs d'accéder à Internet depuis le réfrigérateur pour acheter en ligne des produits. L'évolution suivante consisterait à automatiser la gestion de la liste de courses. Chaque fois qu'un produit est retiré du réfrigérateur, l'ordinateur met à jour l'état des stocks. Lorsqu'un seuil critique est atteint, le produit est ajouté sur la liste et un signal avertit l'utilisateur.



**Figure 17** Illustration du réfrigérateur augmenté : l'informatique classique (a) vs. le principe de la RA (b).

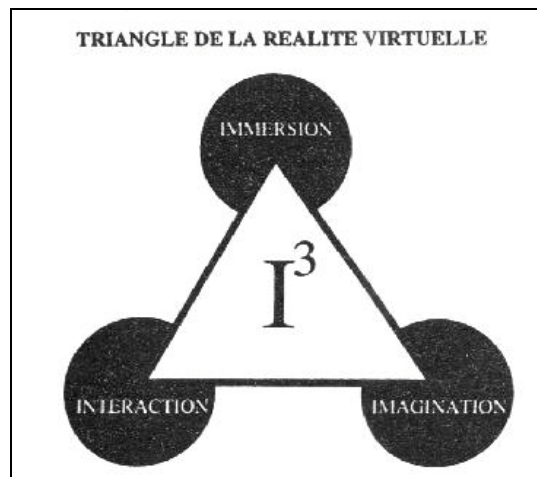
La RA vise donc à permettre à l'utilisateur de rester au contact de l'environnement réel sur lequel il travaille, en associant des capacités de traitements complémentaires, directement à son environnement, et aux objets réels impliqués dans la tâche.

La RA est souvent mise en relation avec la Réalité Virtuelle (RV). Tandis que pour certains la RA est une évolution de la RV, pour d'autres la RA est mise en opposition

avec la RV. Au paragraphe suivant nous définissons la RV et nous la positionnons par rapport à la RV.

## 1.2 Positionnement de la RA par rapport à la RV

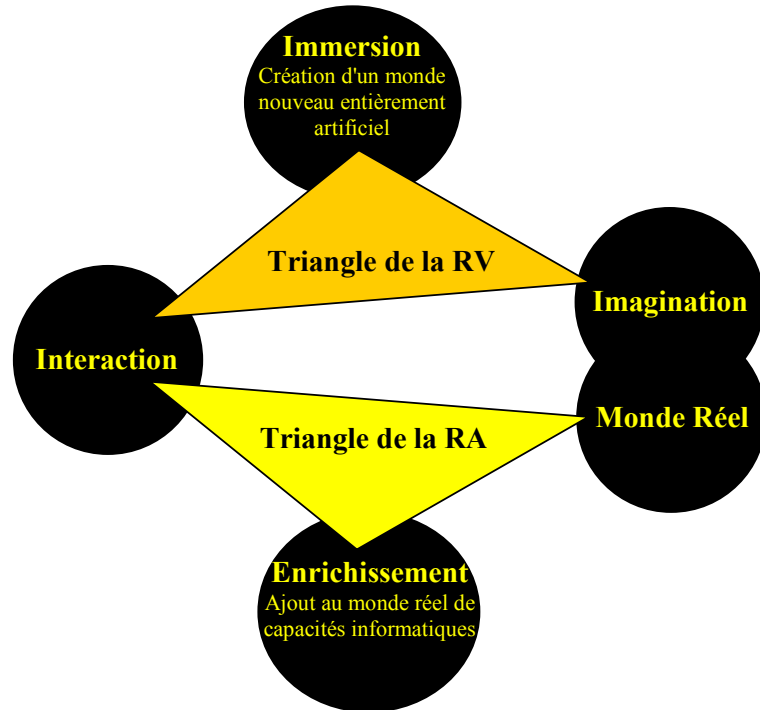
La réalité virtuelle est définie comme un environnement interactif, généré par un ordinateur et dans lequel une personne est immergée [Aukstakalnis 92]. Le monde ainsi généré étant une simulation d'un monde existant, l'illusion doit être donnée à l'utilisateur que les objets ainsi que les capacités d'actions sur ces objets virtuels entourent cet utilisateur. Pour cela l'utilisateur est **immergé** dans un environnement artificiel. Dans cet environnement, l'utilisateur est actif. Par des réactions plausibles et immédiates, la réalité virtuelle doit présenter toutes les similitudes de la réalité : c'est **l'interaction**. Enfin, dans la mesure où le monde virtuel n'est pas une copie exacte du réel, la RV fait appel à **l'imagination** de l'utilisateur. La Figure 18 illustre l'approche de [Burdea 94] selon laquelle la Réalité Virtuelle résulte de l'intégration de ces trois notions que sont l'immersion, l'interaction et l'imagination.



**Figure 18** *Triangle de la réalité virtuelle.*

De notre point de vue, la RA ne nécessite pas obligatoirement la gestion par l'ordinateur d'une scène réaliste, puisque les données ajoutées au monde réel peuvent servir à comprendre le monde réel et se présenter par exemple sous la forme de textes. Ainsi, à la différence de la RV, les données manipulées par l'ordinateur ne sont pas soumises à des contraintes de présentation. De plus en RA, l'utilisateur est maintenu en relation directe avec son environnement réel. Le degré d'immersion de l'utilisateur dans la scène virtuelle produite par le système constitue donc le principal point de divergence entre RA et RV. Par opposition à l'immersion totale de l'utilisateur dans un monde constitué d'objets et d'actions virtuelles, les capacités de traitement des données offertes par l'ordinateur dans les systèmes de RA sont ajoutées à l'environnement réel de l'utilisateur pour l'enrichir. En adaptant le triangle de la RV à la RA, cet enrichissement du monde réel constitue alors l'un des sommets du triangle de la RA. N'étant plus immergé dans un monde reconstitué, l'utilisateur d'un système de RA n'a plus à faire appel à son imagination pour interpréter les données perçues. Il demeure dans son environnement réel, ce qui constitue un point singulier des systèmes de RA et fixe à ce titre un sommet du triangle de RA. Or, comme le montre la Figure 19, c'est sur la base des connaissances relatives au monde réel que se développe l'imagination de l'utilisateur. Le second sommet du triangle de la RA a donc une intersection avec le sommet

Imagination du triangle de la RV. Enfin, un système de RA reste bien évidemment un système interactif. Même si les techniques d'interaction diffèrent, l'interaction constitue néanmoins un sommet commun aux triangles de la RV et de la RA. La Figure 19 ci-dessous positionne le triangle de la RA en regard du triangle de la RV présenté à la Figure 18.



**Figure 19** *Triangles de la Réalité Augmentée (RA) et de la Réalité Virtuelle (RV).*

Après avoir défini le paradigme de la RA, nous l'illustrons par des exemples dans le paragraphe suivant. Ces derniers sont issus de domaines différents afin d'illustrer la diversité des domaines d'application de la RA.

## 2 Exemples de systèmes de réalité augmentée existants

Nous présentons dans ce paragraphe des systèmes qui ont en commun la combinaison des mondes réels et virtuels. Nous distinguons dans la littérature deux types de systèmes de RA, qui relèvent de deux axes de recherche distincts. Le premier axe, que nous illustrons au paragraphe 2.1 combine les deux mondes au sein d'une représentation informatique tandis que le second, objet du paragraphe 2.2, les combine au sein du monde réel.

### 2.1 RA : Intégration du monde réel dans le monde virtuel

Au sein de cet axe, nous identifions deux formes d'intégration. La première consiste à intégrer le monde réel au niveau des sorties, ou représentations virtuelles, produites par le système informatique, tandis que dans la seconde, les entités réelles sont combinées à une représentation virtuelle pour assister la communication de l'utilisateur à destination du système : le monde réel est alors exploité pour manipuler la représentation virtuelle fournit pas le système.

### 2.1.1 Intégration pour les sorties du système (du système vers l'utilisateur)

Cette forme d'intégration est illustrée par le système *Interior Design* présenté dans [Whitaker 95]. Le système vise à aider un décorateur à aménager une pièce. A l'aide du système, l'utilisateur peut créer une représentation 3D informatique d'une pièce à meubler. Dans cette pièce représentée informatiquement, il dispose de modèles 3D de meubles ou de chaises par exemple. Il peut aussi ajouter des photographies d'objets réels qui sont insérés dans la pièce. Une scène affichée à l'écran est présentée à la Figure 20 : celle-ci inclut la photo d'une table réelle et un objet virtuel, la chaise.

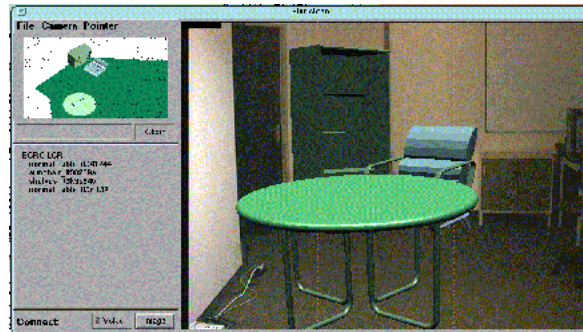


Figure 20 Présentation des données réelles et virtuelles par *Interior Design*.

Notons que les effets spéciaux utilisés dans le cinéma tendent à intégrer de plus en plus de données virtuelles au film. Le principe semble donc proche. Pourtant il est très clair que les trucages cinématographiques ne relèvent pas de la RA : à l'inverse du cinéma, *Interior Design* a pour cible un objet du monde réel, la pièce à aménager. En considérant cet aspect, la réalité de l'aménagement d'une pièce est augmentée par la possibilité de visualiser de façon réaliste le résultat. Dans un film avec des trucages, le seul but est de rendre la scène plus réaliste ou plus fantastique, sans pour autant devoir favoriser un résultat particulier dans le monde réel.

Dans *Interior Design*, la combinaison de données du monde réel, la table, et du monde virtuel, les autres meubles, au sein d'une représentation informatique affichée à l'écran, est bien utilisée en sortie du système. Voyons maintenant comment l'intégration des deux mondes peut être utilisée en entrée du système, c'est-à-dire pour favoriser la communication de l'utilisateur vers le système.

### 2.1.2 Intégration pour les entrées du système (de l'utilisateur vers le système)

Les systèmes illustrant cette formes d'intégration sont décrits dans la littérature comme des interfaces tangibles ("*Tangible User Interface*" - TUI) [Ishii 97]. Leur objectif est de rendre les objets informatiques physiquement accessibles. Considérons deux systèmes : *Illuminating Light* et le *Tangible Geospace*.

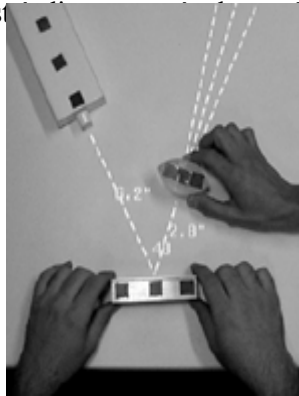
#### 2.1.2.1 Le système *Illuminating Light*

*Illuminating Light* [Underkoffler 98] du projet nommé "*Tangible bits*" est un exemple d'interface tangible. En disposant sur un bureau différents éléments optiques, tels des lentilles et des prismes ainsi qu'une source virtuelle de lumière, l'utilisateur peut visualiser en temps réel le "chemin" suivi par le faisceau lumineux. Etant donné que le tracé est effectué en temps réel, lorsque l'utilisateur déplace un élément optique, le chemin lumineux est automatiquement recalculé et aussitôt projeté sur la table. La Figure



21 ci-dessous montre le système en fonctionnement. On peut noter que les données projetées ne se limitent pas au faisceau lumineux, mais que la longueur du faisceau entre deux éléments par exemple est aussi projetée. Ce système permet d'expérimenter l'impact de la position des éléments optiques sur le chemin suivi par un faisceau lumineux. Les données projetées le sont par l'intermédiaire d'un "I/O-Bulb", dispositif combinant un projecteur et une caméra, cette dernière réalisant la capture de la scène.

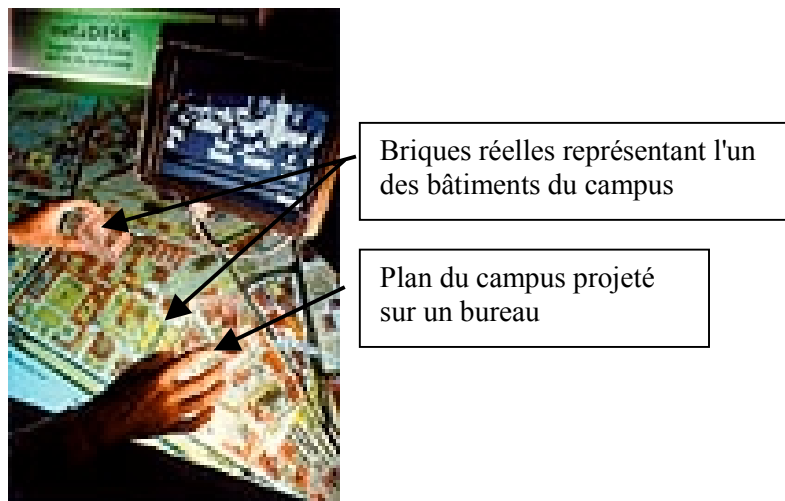
En conclusion, l'intégration d'entités réelles dans le système informatique simulant le comportement d'un faisceau lumineux intervient au niveau de la communication de l'utilisateur vers le système, c'est-à-dire au niveau de l'interaction.



**Figure 21** *Illuminating Light en fonctionnement.*  
Photo issue de [tangible.media.mit.edu](http://tangible.media.mit.edu).

### 2.1.2.2 Le Tangible Geospace

Appartenant au même projet de recherche, le *Tangible Geospace* [Ishii 97] projette un plan du campus du MIT autour d'un ou deux cubes, appelés PHICON<sup>2</sup>, représentant chacun un bâtiment du campus. Une caméra localise sur le bureau l'emplacement des deux PHICON's, transmet les données au système informatique. L'ordinateur adapte alors l'échelle du plan qu'il a en mémoire pour faire coïncider les deux PHICON's avec leur emplacement sur le plan. Enfin le résultat est projeté sur la table où se trouvent les deux PHICON's. Le fonctionnement est illustré à la Figure 22.



**Figure 22** *Tangible Geospace en fonctionnement.*  
Photo issue de [tangible.media.mit.edu](http://tangible.media.mit.edu).

<sup>2</sup> PHICON : "Physical Icon", objet physique pouvant jouer le rôle d'une icône.

La manipulation du plan, fichier informatique, repose sur l'utilisation de cubes, objets réels. L'intégration d'objets réels dans les systèmes informatiques contribue, comme dans l'exemple précédent, au niveau de la communication de l'utilisateur vers le système.

Voyons maintenant une forme différente de la RA, adoptée dans le domaine de l'IHM : il s'agit de combiner les données en provenance des deux mondes, tout en restant en contact direct avec le monde physique réel. Cette approche consiste donc à intégrer des données virtuelles au monde physique, c'est donc le cas inverse de ce que nous venons de présenter.

## **2.2 RA : Intégration du monde virtuel dans le monde réel**

Dans ce paragraphe, nous présentons des systèmes qui visent tous à intégrer le monde informatique dans le monde réel. Il est néanmoins important de noter que tous les systèmes présentés ne sont pas caractérisés comme étant des systèmes de RA dans la littérature du domaine. Nous organisons la présentation de ces systèmes selon une caractéristique classique du domaine des IHM : leur aspect mono- ou multi- utilisateurs.

### **2.2.1 Systèmes de RA mono-utilisateur**

L'utilisation de ces différents systèmes repose sur différents sens de perception : l'audition, la vue, ou le toucher. Les dispositifs utilisés sont également très variés : il peut s'agir de casques semi-transparents, d'ordinateurs de poche, ou encore de projection de données sur des objets réels. Au paragraphe suivant, nous soulignons cette diversité des possibilités, par des exemples de systèmes.

#### **2.2.1.1 Le Musée Augmenté**

Présenté dans [Rekimoto 95] comme une application de la NavyCam de Sony, le *Musée Augmenté* affiche des données dans un casque semi-transparent porté par un visiteur dans un musée. Le système est basé sur la lecture d'un code-barre disposé dans un coin des œuvres, qui entraîne l'affichage de données textuelles dans le casque. Ces données sont relatives au tableau que l'utilisateur est en train de contempler et concernent l'auteur de l'œuvre, ou encore des caractéristiques techniques du tableau. Les données proposées sont configurables par l'utilisateur, c'est-à-dire qu'il peut choisir le(s) type(s) de données à afficher en fonction de ses connaissances et ses intérêts. La Figure 23 montre la vue d'un utilisateur face à un tableau du musée. A tout moment au cours de la visite, l'utilisateur voit les œuvres réelles ainsi que des données complémentaires, affichées dans le casque et restituées par l'ordinateur. Mise à part la phase de configuration du système, l'interaction entre l'utilisateur et le système est totalement transparente aux visiteurs.



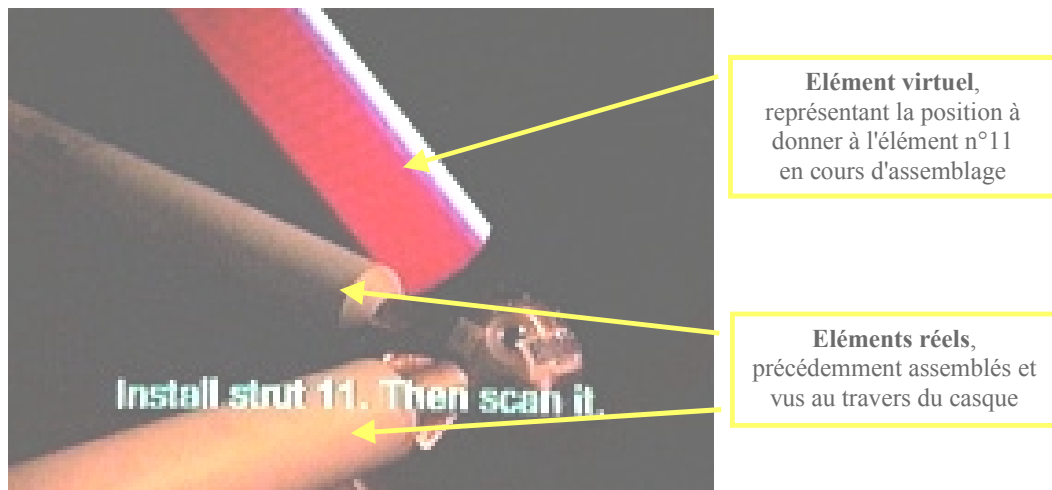
**Figure 23** *Vue d'un utilisateur au travers du casque du Musée Augmenté.  
Photo issue de [Rekimoto 95].*

Notons que ce système peut être adapté à d'autres tâches semblables. Par exemple, un système peut afficher dans le casque semi-transparent les nouveaux livres ou magazines d'un rayon de la bibliothèque lorsque l'utilisateur se place devant ce rayon [NaviCam].

### 2.2.1.2 *Le système ARC : "Augmented Reality for Construction"*

Egalement basé sur l'utilisation d'un casque semi-transparent, ce système présenté dans [Webster 96], vise à simplifier un travail d'assemblage. Afin de tenir compte de différentes caractéristiques physiques et techniques et de respecter les plans de conception de la structure globale, *ARC* procure à l'utilisateur des données facilitant le montage de la structure. Ces données sont de deux sortes : textuelles et graphiques. Pour percevoir ces données, l'utilisateur porte un casque semi-transparent. Les données graphiques affichées sont exactement juxtaposées à la partie de la structure déjà en place afin de montrer à l'utilisateur comment positionner le nouvel élément par rapport aux parties de la structure déjà en place. En effet, l'utilisateur et la structure font l'objet d'un suivi par caméra, ou plus généralement par un localisateur. Ce localisateur transmet ces données au système informatique qui peut alors mettre à jour des données affichées dans le casque semi-transparent. La Figure 24 montre la vue de l'utilisateur au travers du casque : dans cette démonstration, l'élément 11 de la structure est en train d'être mis en place, sa position est indiquée en surimpression du reste de la structure déjà assemblée. L'élément 11 réel n'est pas visible car il n'est pas encore présent dans le champ de vision de l'utilisateur.

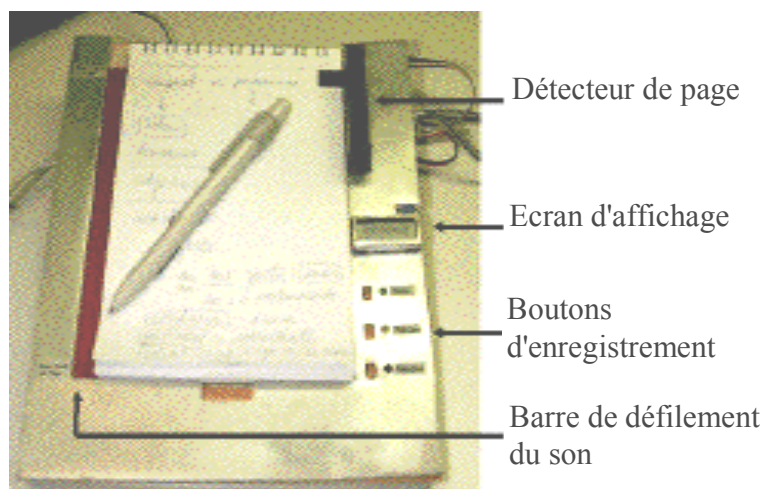
Une fois l'élément de la structure correctement positionné, l'utilisateur renseigne le système en scannant un code-barre fixé sur cet élément. L'utilisation du scanner constitue d'une part le seul moyen d'identifier précisément l'élément qui vient d'être posé et matérialise ainsi la fin de la tâche visant à l'assemblage de cet élément avec le reste de la structure. Le système peut alors mettre à jour les données affichées dans le casque en indiquant le numéro et la position du prochain élément, ou plus généralement la tâche suivante à accomplir. Une autre version de ce système met en œuvre un autre média : des données sonores transmises par l'intermédiaire d'un casque stéréo à l'utilisateur.



**Figure 24** Vue d'un utilisateur du système ARC au travers du casque semi-transparent.  
Photo issue de [www.cs.columbia.edu](http://www.cs.columbia.edu).

### 2.2.1.3 Le Paper-Based Audio Notebook

[Stifelman 96] présente un bloc-notes réalisant la capture et la restitution de données sonores. Le *Paper-Based Audio Notebook*, repose comme *ARC* sur l'utilisation de données sonores. L'un des aspects particuliers de ce système est qu'il permet d'augmenter les notes prises manuellement, lors d'un cours par exemple, en proposant l'accès à l'enregistrement audio de la totalité du cours ou de l'entretien. Ainsi, lorsque l'utilisateur consulte ses notes manuelles, il peut réentendre la conversation à un moment précis : la bande audio est indexée sur les numéros de pages et éventuellement sur d'autres points de repère de la page. Par simple désignation d'un point de repère sur le papier, la bande sonore est lue à partir de l'instant correspondant de la bande audio. La Figure 25 montre l'appareillage constituant cette version évoluée d'un bloc-notes. A l'opposé du système précédent, l'interaction avec le système est entièrement laissée à l'initiative de l'utilisateur qui, à l'aide de son stylo peut ajouter des notes, en modifier et activer la lecture de la bande sonore. Dans le système précédent, les données affichées par le système étaient automatiquement mises à jour par le système informatique.



**Figure 25** Dispositif du Paper-Based Audio Notebook.  
Photo issue de [Stifelman 96].

#### 2.2.1.4 *Le système Haptic interfaced telemanipulation*

Le système de télémanipulation, présenté dans [Chen 98] repose sur le retour d'effort pour favoriser l'interaction de l'utilisateur avec un robot distant. A chaque instant, l'utilisateur contrôle les déplacements du robot par le biais de manettes de commandes. L'ajout de retour d'effort permet à l'utilisateur de percevoir les forces qui s'opposent aux déplacements du robot. Il complète ainsi le retour vidéo utilisé seul dans une version précédente. La qualité de réalisation de sa tâche est donc accrue par l'ajout de données haptiques. Ces données virtuelles sont donc basées sur un retour d'effort et viennent compléter l'environnement réel de l'utilisateur : en effet l'utilisateur ne se contente plus de diriger et de voir le robot distant, il ressent également les contraintes que le robot rencontre sur le terrain. L'intégration du virtuel dans le réel enrichit l'interaction entre l'utilisateur et le robot distant.

#### 2.2.1.5 *Le système Chameleon*

Le système *Chameleon* se distingue des autres systèmes présentés par le dispositif de rendu utilisé pour afficher des données textuelles. En effet, ce système vise à combiner l'utilisation d'une partie du monde réel et d'un Palm-Top, ordinateur de poche, ce dernier étant localisé dans l'espace [Fitzmaurice 93]. En fonction de la position du Palm-Top, le système actualise les données affichées sur l'écran du Palm-Top. L'environnement de l'utilisateur est alors enrichi. Par exemple l'utilisateur du Palm-Top se trouve devant une carte géographique : le système de localisation du Palm-Top permet de déduire la ville au niveau de laquelle le Palm-Top est placé. Le système recherche dans la base de données les données relatives à cette ville (nombre d'habitants, surface, activités) et les affiche sur l'écran du Palm-Top.

A la différence du *Musée Augmenté* par exemple, les données réelles et virtuelles ne sont pas exactement superposées puisque les surfaces de perception sont différentes. D'ailleurs, le dispositif de sortie, le Palm-Top, dissimule aux yeux de l'utilisateur une partie de l'objet sur lequel il travaille, en l'occurrence la carte géographique. L'utilisateur est donc contraint de quitter des yeux la carte géographique réelle pour percevoir les données fournies par le système. De plus, ce système semble moins transparent pour l'utilisateur que le *Musée Augmenté*, car il requiert une action explicite de l'utilisateur (déplacer le Palm-Top) chaque fois qu'il souhaite percevoir une donnée complémentaire à un endroit donné. Néanmoins la combinaison d'un système informatique, d'un dispositif d'affichage particulier et d'un processus de localisation permet d'enrichir le monde réel, en l'occurrence la carte géographique affichée sur le mur. De plus, la présence d'un ordinateur, qui maintient les données relatives aux différentes villes lui est complètement occultée. Sans ces dispositifs, l'utilisateur aurait dû se déplacer explicitement sur un ordinateur et spécifier manuellement la ville qui l'intéresse. La perception de la carte géographique n'aurait alors pu se faire que par le biais d'une indirection, comme expliquée au paragraphe 1.1.

#### 2.2.1.6 *Le Digital Desk*

Le dernier système considéré, le *Digital Desk* [Wellner 93], est l'un des premiers systèmes de RA développés dans le domaine des IHM. Il permet la manipulation de documents physiques et électroniques sur une même surface : un bureau. Surmonté d'une caméra et d'un projecteur, ce bureau présente à l'utilisateur des documents électroniques et réels. L'utilisateur peut manipuler indifféremment les deux types de documents,

effectuer des "copier/coller", enregistrer ou imprimer ces données. La photo de la Figure 26 représente le dispositif requis par ce système.



**Figure 26** *Eléments composant le Digital Desk.*

A la différence des systèmes précédemment présentés, le *Digital Desk* n'utilise aucun dispositif d'affichage dédié : le bureau fait office d'écran. En effet, c'est un projecteur vidéo qui rend perceptible les données virtuelles, textuelles ou graphiques, directement sur la surface de travail réelle de l'utilisateur.

Surfaces d'objets réels, écrans d'ordinateur de poche, écrans de casque semi-transparent, sont autant de surfaces de rendu qui soulignent la grande diversité de systèmes de RA. Les sens de perception requis lors de l'interaction contribuent aussi à cette diversité. Nous allons voir au paragraphe suivant, que l'aspect multi-utilisateurs vient encore ajouter de nouvelles solutions de conception de systèmes de RA.

### **2.2.2 Systèmes de RA multi-utilisateurs**

Les systèmes multi-utilisateurs sont plus communément désignés sous le terme de systèmes collaboratifs. La distinction des situations collaboratives s'appuie habituellement sur un schéma de classification classique [Ellis 91]. Ce schéma caractérise une interaction collaborative et s'articule autour de deux axes : l'espace et le temps. Le premier axe, l'espace, met en évidence deux types d'interactions liés à la situation de l'utilisateur par rapport au système : on parle donc d'interaction locale ou distante. Le second axe, le temps, identifie le caractère synchrone ou asynchrone de l'interaction de chaque utilisateur avec le système. Dans le cas des systèmes de RA, une autre distance est à prendre en compte : la distance des utilisateurs par rapport à l'objet de la tâche, c'est-à-dire l'environnement physique où se réalise la tâche. Le schéma présenté dans [Renevier 01] prend en considération cette nouvelle distance et présente trois classes distinctes pour les systèmes collaboratifs mettant en œuvre le paradigme de RA. La première classe regroupe des systèmes où seul un utilisateur évolue dans un environnement dans lequel sont ajoutées des données issues d'autres utilisateurs distants. La seconde classe regroupe les systèmes dans lesquels l'objet de la tâche est réparti entre plusieurs sites distants. Enfin la dernière classe de systèmes de RA collaboratifs rassemble les systèmes dans lesquels tous les utilisateurs sont réunis autour de l'objet de la tâche. Nous présentons dans ce paragraphe plusieurs systèmes pour illustrer chacune de ces trois catégories.

### 2.2.2.1 *La Gesture Cam*

La *GestureCam* [Kuzuoka 94] illustre la première catégorie de systèmes de RA collaboratifs [Renevier 01] : en effet, ce système est utilisé pour permettre à un technicien de réparer un vélo en recevant l'aide d'un expert distant au moyen de données sonores et de données affichées dans un casque semi-transparent. L'environnement de travail du technicien est donc augmenté par l'ajout de données produites par des utilisateurs distants. Le système fournit à l'expert une vue à peu près similaire à celle de l'opérateur. L'opérateur peut donc agir en suivant les instructions et en utilisant des références spatiales communes aux deux utilisateurs. Du côté de l'expert, les données filmées par la caméra de l'opérateur sont affichées. Une caméra est située dans l'axe de cet écran et enregistre les mouvements faits par l'expert. Ces mouvements sont ensuite retransmis à l'opérateur qui, au travers du casque semi-transparent voit ce que l'expert lui désigne. L'expert peut donc fournir des instructions orales ainsi que gestuelles.

### 2.2.2.2 *Le système DOLPHIN*

La seconde classe identifiée dans [Renevier 01] regroupant les systèmes dans lesquels l'objet de la tâche est réparti entre plusieurs sites distants est illustrée par tous les systèmes mettant en œuvre de multiples tableaux électroniques distants. Le système *DOLPHIN* [Streitz 94] en est un exemple. Ce système se compose de tableaux blancs sur lesquels les utilisateurs peuvent écrire avec des feutres traditionnels. Pour chaque tableau, une caméra et un projecteur permettent d'établir un lien avec les autres tableaux blancs, distants. Les modifications apportées à un tableau blanc, manuellement ou par le biais de claviers et de souris, sont donc reflétés sur tous les tableaux blancs distants.

### 2.2.2.3 *Les médiaspace*

Cette seconde classe peut également être illustrée par les médiaspace. Ces systèmes enrichissent des environnements de travail existants mais distants, en ajoutant des capacités de communication audio et/ou vidéo entre différents lieux de connexion au médiaspace. Un lieu de connexion se définit par un bureau physique (une pièce dans un bâtiment), un écran sur lequel s'affichent les données du médiaspace et une caméra reliée à un ordinateur permettant la capture d'images. Dans l'exemple du médiaspace *CoMedi* [Coutaz 98], les données affichées simultanément dans chaque lieu de connexion, fournissent une vue de l'espace de travail de chacun des autres utilisateurs connectés au médiaspace. Des données textuelles complémentaires peuvent être affichées ainsi qu'un niveau d'occupation de la personne connectée en ce lieu. Un médiaspace fournit donc une aide à la localisation d'une personne, renforce la sensation d'appartenance à un groupe et la sensation d'accessibilité aux autres membres du groupe. Un médiaspace peut aussi aider à mettre un utilisateur en relation plus proche avec l'environnement de travail d'un autre utilisateur. C'est le cas du *Multiple Target Video* [Gaver 93], qui transmet plusieurs vues d'un même lieu de connexion, mais sous des angles différents.

Dans le cas des médiaspace, l'apport du monde informatique au monde réel ne permet pas exactement de percevoir des données complémentaires à l'environnement physique réel de l'utilisateur. Par contre, un médiaspace permet d'agrandir l'environnement physique de l'utilisateur en créant une sorte de grand bureau unique, composé de différents lieux de connexion.



#### 2.2.2.4 Les systèmes de vidéo-conférence

Enfin, les systèmes de vidéo-conférence relèvent également de cette seconde catégorie de systèmes de RA collaboratifs, dans lesquels l'objet réel est réparti en différents lieux physiques distants. L'augmentation provient de la possibilité pour les participants d'être distants de l'interlocuteur principal, tout en pouvant assister et intervenir au cours de la conférence. A la différence des médiaspace, ils sont en général utilisés dans des salles de réunions, c'est-à-dire avec un large nombre de participants en face de la zone d'affichage des données, et peuvent également permettre le partage de documents électroniques.

Par exemple, le système *MAJIC* [Okada 94], projette les images des participants en grandeur réelle sur un grand écran incurvé. La métaphore ainsi mise en œuvre est celle d'une réunion se déroulant autour d'une table, alors que les participants sont tous distants. Dans ce système, tous les utilisateurs ont une même référence spatiale, ce qui rend la conversation très proche d'une conversation de type face-à-face. Là encore, l'environnement physique des utilisateurs n'est pas modifié, mais la création d'une table de réunion composée de bureaux géographiquement distants permet d'agrandir l'environnement physique de chacun des utilisateurs.

#### 2.2.2.5 Le jeu de Mah-Jong et le système CAVE

Le jeu de *Mah-Jong* [Szalavari 98] illustre la dernière classe de systèmes de RA collaboratifs [Renevier 01], rassemblant les systèmes dans lesquels tous les utilisateurs sont réunis autour de l'objet de la tâche. Ce système utilise des casques semi-transparents et consiste en une table de jeu sur laquelle les joueurs doivent déposer tour à tour leur combinaison de "dominos" virtuels. La table ainsi que les combinaisons virtuelles déjà jouées sont donc communes à tous, mais leurs "dominos" et les combinaisons qu'ils préparent, doivent rester cachés des autres joueurs : ils sont perceptibles au niveau d'un PIP (Personal Interaction Panel) au travers de leur propre casque semi-transparent, comme l'illustre la photo de la Figure 27. Au travers du casque un utilisateur peut donc voir le plateau de jeu commun avec les combinaisons virtuelles déjà déposées, ainsi que son espace de jeu privé, sur lequel se trouvent ses dominos et les combinaisons qu'il est en train de préparer. La manipulation des dominos virtuels dans l'espace privé de chacun ou sur la table de jeu est réalisée par le biais de l'utilisation de stylo repéré dans l'espace par un localisateur magnétique. Ce système permet donc de mettre en œuvre une activité collaborative en utilisant le principe de réalité augmentée sur le plateau de jeu commun.



Figure 27 Vue au travers du casque de l'un des joueurs de Mah-Jongg.



C'est aussi le cas des systèmes qui immergent plusieurs utilisateurs dans un environnement artificiel généré par un système informatique. Dans ce monde artificiel, la communication se déroule normalement entre les participants situés côte à côte. C'est par exemple le système *CAVE* [Cruz-Neira 93]. Le principal inconvénient est qu'il immerge les utilisateurs dans un monde qui, même s'il correspond à l'image captée par une caméra dans le monde extérieur, n'est pas le monde réel : tous les utilisateurs ont une seule et même vue alors qu'en réalité, ceux-ci auraient chacun leur angle de vue et leur perspective. Néanmoins la réalité de la communication face-à-face est augmentée par le système.

### 3 Approches et limites de la littérature

Au paragraphe précédent, nous avons présenté des systèmes dont la caractéristique commune est de combiner les mondes réel et virtuel. La diversité des systèmes présentés souligne la variété des possibilités. Dans ce vaste espace des possibilités, nous avons discerné l'approche qui consiste à exploiter le monde réel au sein du monde informatique et celle opposée qui consiste à intégrer le monde informatique dans le monde réel. Néanmoins cette dichotomie n'est pas suffisante pour rendre compte pleinement de la diversité des possibilités. Il est nécessaire de définir un cadre de classification permettant de regrouper les systèmes présentant des similarités et de bien distinguer ceux dont le fonctionnement ou le rôle est différent. Par la suite, ce cadre de classification permet de fédérer les approches d'étude dans le domaine de la RA. Dans cet objectif, les approches de la littérature visent à mieux définir ce concept afin de pouvoir envisager les différentes formes possibles que peut prendre un système de RA. Malgré la relative jeunesse du paradigme, deux approches concurrentes constituent les bases de la RA.

Il s'agit d'une part de l'approche du domaine de la synthèse d'images qui caractérise les systèmes de RA adoptant une représentation informatique mêlant virtuel et réel, et d'autre part des approches du domaine de l'IHM, orientées sur les caractéristiques liées à l'interaction entre l'utilisateur et le système.

#### 3.1 Deux approches concurrentes

##### 3.1.1 Approche du domaine de la synthèse d'images

Dans le domaine de la synthèse d'images, nous regroupons l'ensemble des systèmes qui combinent une partie du monde réel et du monde informatique au sein d'une représentation informatique. L'utilisation du réel combiné au virtuel est incontestable puisqu'elle permet par exemple de concevoir des décorations d'intérieurs. Mais cette combinaison se fait sur un écran d'ordinateur : la photo d'une pièce réelle est "augmentée" par des modèles de fauteuils par exemple. La scène observée est bien une combinaison des deux mondes mais la perception du monde réel induit une indirection. C'est dans ce cadre que sont présentés des travaux sur la réalité mixte [Milgram 94]. Ces travaux concernent uniquement les dispositifs d'affichage permettant de combiner des données réelles et virtuelles.

L'approche de caractérisation, présentée dans [Milgram 94], prend en compte quatre principes permettant de visualiser une donnée de réalité augmentée. Le premier caractérise la **provenance** de l'image, est-ce une vidéo de la réalité ou bien une

modélisation informatique d'une partie du monde réel. Un second considère le niveau d'**indirection** subie par l'utilisateur pour percevoir les données relatives au monde réel, autrement dit, le fait que l'utilisateur perçoit la réalité directement ou via un dispositif d'affichage électronique. Le troisième principe est le **point de vue** de l'utilisateur sur la partie du monde réel qu'il perçoit. Enfin le dernier, prend en compte la capacité du dispositif d'affichage à conserver les **dimensions** et proportions du monde réel. Selon Milgram ces quatre principes sont insatisfaisants dans la mesure où la classification résultante devient rapidement hasardeuse car les concepts à la base de ces distinctions ne sont pas suffisamment précis. Par exemple, la provenance tend à identifier le caractère réel ou virtuel de l'image visualisée. Mais la qualité de rendu des scènes 3D ne permettra bientôt plus de distinguer les images calculées par ordinateur des images photographiées dans la réalité. En conséquence, l'objectif principal d'une taxonomie, fédérer les approches d'études, n'est pas rempli puisque les axes de classification deviennent vagues.

[Milgram 94] transforme alors l'approche initiale afin de produire une taxonomie plus formelle, permettant de répondre à trois questions de conception que les auteurs jugent essentielles :

- Que savons-nous du monde affiché ?
- Pouvons-nous l'afficher de façon réaliste ?
- Dans quelle mesure l'utilisateur se sent-il présent dans ce monde affiché ?

Il propose alors une taxonomie qui s'articule autour de trois continua. Le premier, **Extent of World Knowledge**, vise à caractériser la quantité de savoir que l'ordinateur maintient de la scène affichée : par exemple, un flux vidéo se trouverait à l'extrême minimum de ce continuum, tandis que le résultat graphique d'une simulation numérique serait placé à l'extrême opposé. Il s'agit donc de distinguer les systèmes en fonction de leur connaissance de la scène. Le second, **Reproduction Fidelity**, caractérise un système en fonction de sa qualité de rendu. A l'extrême inférieur se trouve un flux vidéo monoscopique, noir et blanc de faible résolution, tandis qu'à l'extrême supérieur se trouve une vidéo stéréoscopique de très haute résolution. Enfin le dernier axe, **Extent of Presence Metaphor**, caractérise la sensation que l'utilisateur aura d'être présent dans la scène. Ceci regroupe les notions d'immersion et de point de vue.

Cette taxonomie identifie les caractéristiques des dispositifs d'affichage à prendre en compte lors de la conception et de la réalisation de systèmes de réalité augmentée. Aucune restriction relative au niveau d'immersion que le système doit fournir n'est apportée. Le but de l'utilisation de ces systèmes n'étant pas véritablement défini non plus, le spectre des systèmes, affichant une combinaison de données du monde réel et virtuel, et donc couverts par cette taxonomie, est très large. Nous avons d'ailleurs évoqué le cinéma et ses trucages pour illustrer le fait que la limite entre système de RA et systèmes interactifs s'avère assez difficile à établir selon cette première approche.

### 3.1.2 Approches de la communauté IHM

Les approches de caractérisation de la RA dans la communauté IHM visent à distinguer les différentes manières d'ajouter en temps réel des données ou des services informatiques dans le monde réel. Il s'agit de distinguer les méthodes permettant d'accroître la capacité de l'utilisateur, à percevoir des données ou à exécuter des tâches,

sans le couper de son environnement physique. Ces approches regroupent des systèmes fusionnant des données du monde réel et du monde virtuel, mono- ou multi-utilisateurs.

En partant de deux principes communs, briser la frontière réel/virtuel, et aider l'utilisateur à réaliser une tâche, les approches existantes tentent d'expliquer comment le monde réel est augmenté ou plus exactement comment la RA peut permettre d'accroître, à un moment donné, les capacités perceptuelles de l'utilisateur engagé dans une tâche.

Une première approche vise à affiner ce terme de capacités perceptuelles. Deux approches privilégient la vue puisqu'elles considèrent que les données ajoutées par l'usage de la RA sont des données graphiques tridimensionnelles [Azuma 97], ou une combinaison de graphiques et de données textuelles [Webster 97]. Une troisième approche, par contre, ne considère pas que la vue constitue la seule capacité perceptuelle susceptible d'être enrichie par un système de RA : en effet une extension du système *KARMA*, qui se baserait sur l'ajout d'un module de parole, ou au moins un module sonore pour guider l'utilisateur, est envisagée [Feiner 93]. Ces données sonores compléteront les données graphiques affichées dans le casque semi-transparent. Un nouveau type de dispositif est d'ailleurs envisagé : le "hear-through". Enfin, l'utilisation de données haptiques fournies par un système est aussi considérée comme une augmentation des capacités perceptuelles [Noma 96]. Le toucher fait donc également partie des capacités perceptuelles susceptibles d'être augmentées par un système de RA.

Le sens perceptif requis (vue, son, toucher) pour percevoir les données virtuelles ainsi que leurs formes (textuelle, 2D, 3D) constituent donc deux traits caractéristiques des systèmes de RA.

Tandis que ces approches s'intéressent à la forme de l'augmentation, une autre approche consiste à étudier la cible de l'augmentation [Mackay 96]. Trois cibles candidates à l'augmentation sont identifiées :

- l'utilisateur : il est porteur d'un dispositif lui permettant de percevoir plus de données relatives à des objets réels.
- les objets : ils sont porteurs de dispositifs d'entrée/sortie informatiques, comme par exemple des capteurs de pressions.
- l'environnement : des dispositifs indépendants de l'utilisateur et des objets fournissent et rassemblent des données relatives à l'environnement.

Ces termes sont réutilisés pour décrire les Strip-Paper augmentés ("augmented flight strips") [Mackay 98]. Dans cette approche, quelle que soit la cible augmentée, le but du système est d'améliorer un univers familier en ajoutant les avantages de l'informatique. Les méthodes ou moyens techniques permettant de combiner des données réelles et virtuelles constituent une seconde caractéristique importante pour les systèmes de RA. Devant le développement grandissant des technologies, elle permet en effet d'envisager plusieurs solutions techniques pour parvenir à assister l'utilisateur dans la réalisation de sa tâche. Néanmoins, il semble délicat de fixer rigoureusement une frontière entre l'environnement et les objets. Nous reviendrons sur cette limitation dans le paragraphe 3.2.

Une autre approche adopte un point de vue centré sur la tâche de l'utilisateur [Ahlers 97]. Elle s'abstrait donc des technologies utilisées ainsi que du type de données fournies

et demeure à un niveau d'abstraction élevé. En effet, selon cette approche, un système de RA doit aider l'utilisateur à réaliser des tâches du monde réel, en fournissant dans la réalité physique de l'utilisateur des données complémentaires et relatives à la tâche. Par exemple le système met en évidence dans un rayonnement le livre que l'utilisateur recherche. L'indication virtuelle fournie par le système, une flèche par exemple, est exactement superposée au livre réel, dans l'environnement physique de l'utilisateur. Il ne s'agit pas seulement d'un numéro de livre et d'étagère à localiser dans une bibliothèque.

Par rapport aux deux approches précédentes, cette dernière tente de caractériser les données fournies du point de vue de son contenu et de sa localisation : seules des données en congruence avec la tâche doivent être présentées dans l'environnement réel de l'utilisateur.

Enfin une dernière approche identifie comme caractéristique des systèmes de RA, le moyen de pallier les différents problèmes liés à l'interaction homme-machine [Rekimoto 95]. Ces problèmes, sont liés au nombre grandissant de tâches à accomplir, mais surtout à la distance existante entre les concepts manipulés dans le monde réel et le monde virtuel. Par exemple, un fichier peut être mis à jour sur un ordinateur, mais la version imprimée de ce fichier ne sera pas mise à jour en même temps : l'utilisateur doit explicitement demander une nouvelle impression du document. Cette distance complique considérablement la manipulation de ces deux mondes. [Rekimoto 95] voit donc dans la RA, le moyen de réduire cette distance, en remplaçant l'utilisateur dans son environnement réel tout en enrichissant ce monde par des données informatiques. Ainsi les capacités d'interaction de l'utilisateur avec le monde réel s'en trouvent élargies et simplifiées.

Toutes les approches, du domaine de la communauté IHM, tentant de caractériser les systèmes de RA, adoptent donc un point de vue centré sur l'interaction homme-système, alors que l'approche du domaine de la synthèse d'images (cf. paragraphe 3.1.1) ne caractérise que les dispositifs d'affichage. Cependant, elles se distinguent les unes des autres par la prise en compte de différents aspects de l'interaction. En synthèse, nous identifions quatre aspects influençant l'interaction d'un système de RA et contribuant actuellement à leur caractérisation:

- le type des données complémentaires fournies à l'utilisateur,
- les technologies utilisées pour combiner des données réelles et virtuelles,
- l'adéquation des données fournies à la tâche courante ainsi que leur lieu de perception,
- la capacité du système à pallier le gouffre existant entre les parties du monde réel et virtuel utilisés pour la réalisation de la tâche.

## **3.2 Avantages et inconvénients**

### **3.2.1 Diversité des aspects pris en compte**

Dans le cas du musée augmenté, les ajouts de textes et de données complémentaires à la scène réelle placent résolument ce système dans la réalité augmentée compte tenu des premières approches du domaine des IHM présentées [Feiner 93], [Webster 97], [Azuma 97]. Du point de vue de l'approche de [Mackay 96], l'utilisation du casque semi-transparent "augmente" l'utilisateur, de même que le code-barre "augmente" l'œuvre d'art.

Pour deux raisons différentes, l'une perceptuelle, l'autre plus basée sur le matériel, ces approches considèrent donc ce système comme faisant partie de la RA.

Si l'on considère le système *ARC*, la différence des aspects pris en compte par les approches de caractérisation de la RA est plus flagrante. Tout comme pour le système précédent, l'affichage de graphiques complémentaires permet de qualifier ce système de système de RA [Feiner 93], [Webster 97], [Azuma 97]. *ARC* réduit clairement le gouffre existant entre le manuel de construction et le réel, ce qui illustre les approches de caractérisation introduites par [Rekimoto 95] et [Ahlers 97]. Enfin, l'approche identifiant la cible candidate à l'augmentation [Mackay 96] établit que l'utilisateur est augmenté par le casque semi-transparent. Par conséquent, comme précédemment, ce système apparaît comme étant de la RA, pour des raisons perceptuelles ou matérielles selon l'approche considérée. Mais la différence résidant dans l'analyse ne s'arrête pas là. Pour pouvoir fournir les bonnes données à l'utilisateur, en fonction des éléments manipulés, le système dépend de l'utilisateur : il doit identifier l'élément de la structure à l'aide d'un lecteur de code-barre. L'utilisateur du système réalise donc deux tâches consécutives, qu'il convient de distinguer pour caractériser le système :

- lors de l'identification de l'élément, l'augmentation de l'utilisateur par le lecteur de code-barre positionne ce système parmi la RA [Mackay 96]. Par contre, le système ne fournit aucune donnée complémentaire à l'utilisateur, ce qui est contradictoire avec le positionnement de *ARC* parmi les systèmes de RA selon toutes les autres approches présentées.
- lors de l'assemblage, l'utilisateur est augmenté par le casque semi-transparent et des données complémentaires lui sont données. Ceci constitue les critères respectifs des approches précédemment contradictoires. Pour cette tâche précise, *ARC* est donc reconnu unanimement comme système de RA.

Nous avons donc illustré que l'approche choisie implique des angles d'analyse différents. Nous avons également souligné que la granularité de l'étude n'est pas le système mais la tâche entreprise par l'utilisateur.

### 3.2.2 Contradictions des conclusions

Dans le cas du *TransBoard* [Ishii 97], les différentes approches de la littérature ne conduisent pas à la même conclusion. Sorte de tableau magique, ou tableau augmenté, l'apparition d'un code barre magnétique sur sa surface, entraîne le stockage informatique de l'ensemble des données inscrites sur ce tableau. Ce stockage est rendu possible par le repérage du tableau par une caméra, reliée à un système informatique qui reconnaît l'apparition d'un PHICON porteur d'un code-barre. L'image du tableau ainsi mémorisée peut être ultérieurement rechargée par la seule utilisation de ce code-barre, sans avoir à entrer manuellement l'adresse informatique du fichier.

L'augmentation d'un objet (le tableau) par une brique portant un code-barre constitue une caractéristique des systèmes de RA [Mackay 96]. A l'opposé, considérer l'apport du système à l'utilisateur, en terme de réalisation de la tâche [Ahlers 97], ou en terme de données ajoutées [Feiner 93], [Webster 97], [Azuma 97] conduit à une conclusion opposée. En effet, le système ne facilite pas l'exécution d'une tâche sur une composante de l'environnement matériel de l'utilisateur, puisque la tâche est une sauvegarde

informatique de données, c'est-à-dire une tâche dans le monde informatique. De plus, aucune donnée n'est présentée à l'utilisateur par ce système.

Ajoutons enfin que le terme "réalité augmentée" semble ici particulièrement inapproprié puisque la réalité n'a pas changé et que la tâche était informatique. Le système facilite l'exécution d'une tâche purement informatique en proposant l'utilisation d'éléments physiques. Ce système se rapproche donc plus de l'approche de la RA suivie dans le domaine de la synthèse d'images [Milgram 94]. D'ailleurs, selon cette dernière approche, le *TransBoard* est un système de RA puisque le dispositif d'affichage combine l'affichage de données virtuelles et réelles. Ce système est donc, en l'état actuel des approches, difficile à cerner.

Pour généraliser, notons que, les interfaces tangibles posent toutes ce même problème : il y a manipulation d'objets réels sans participation à une tâche dans l'environnement réel de l'utilisateur. Certaines approches présentées dans le paragraphe 3.1 en font des systèmes de RA, puisque les interfaces tangibles facilitent l'interaction de l'utilisateur avec le système [Rekimoto 95] et qu'elles utilisent un dispositif d'affichage combinant réel et virtuel [Milgram 94]. Il nous semble que ces systèmes ne puissent pas vraiment être considérés comme augmentant la réalité. La cible est le système informatique (sauvegarde, chargement ou manipulation de données informatiques) et l'utilisation d'un objet du monde réel s'inscrit dans la volonté de définir de nouvelles modalités d'interaction dans un système interactif. Ces systèmes sont donc plus proches des systèmes multimodaux que des systèmes de RA.

### 3.2.3 Limitations propres à chaque approche

L'objectif d'une classification est de fédérer les discussions, de favoriser les comparaisons entre plusieurs systèmes. Une taxonomie doit aussi faciliter l'extrapolation de nouveaux types de systèmes à partir des systèmes existants. Une taxonomie repose donc sur des traits caractéristiques qu'il convient de considérer lors de la conception d'un système. Nous mettons en évidence dans les paragraphes suivants des points propres à chaque approche et qui sont en contradiction avec l'objectif que nous attribuons à une classification.

Les premières approches présentées [Azuma 97], [Feiner 93], [Noma 96], [Webster 97], focalisent leur attention sur le type des données présentées par le système. L'interaction est néanmoins aussi influencée par les dispositifs physiques utilisés. Par exemple, un affichage de données sur un écran d'ordinateur signifie que l'utilisateur doit regarder l'écran. Supposons en outre que lors de sa tâche, l'utilisateur ait à regarder une zone du monde réel, distante de l'écran, il en résulterait que l'utilisateur doit régulièrement quitter des yeux ce sur quoi il travaille pour percevoir les données à l'écran. Une alternative à ce dispositif étant par exemple un casque semi-transparent, l'utilisateur pourrait percevoir exactement la même donnée exprimée de la même façon, mais n'aurait plus besoin de quitter des yeux ce sur quoi il travaille. Le dispositif utilisé peut donc clairement avoir un impact sur le déroulement de l'interaction. Considérer le type de données fournies par le système n'est donc pas suffisant pour caractériser complètement l'interaction dans un système de RA, puisque pour un type de données identiques et une

tâche identique, deux solutions techniques différentes conduisent à des situations d'interaction distinctes.

L'approche consistant en l'identification de l'entité du système qui est augmentée [Mackay 96] constitue un cadre général de caractérisation. De ce cadre, il n'est toutefois pas possible de déduire une taxonomie permettant de distinguer différents systèmes de RA. En effet, identifier l'entité qui a été augmentée, revient à identifier la technologie mise en œuvre pour modifier cette entité. Or plusieurs techniques peuvent être utilisées pour produire exactement les mêmes services à l'utilisateur. Par exemple, une caméra dotée d'un système de reconnaissance de formes est apte à identifier dans une pièce la position d'un étui à lunettes. Cet étui n'a subi aucune modification, il s'agit donc d'une augmentation de l'environnement. Par contre, si l'on instrumente l'étui à lunettes avec un capteur de position, type GPS ou diodes infrarouges), l'étui produit à destination du système des données relatives à sa propre position dans la pièce. Il s'agit ici de l'augmentation d'un objet. L'approche distingue donc deux moyens permettant de réaliser exactement la même chose du point de vue de l'utilisateur : connaître la position de l'étui à lunettes. La technologie utilisée a bien sûr un impact sur la réalisation d'un système. Mais l'exemple précédent montre qu'en limitant la caractérisation d'un système à la technologie qu'il utilise, ne permet pas de regrouper deux systèmes ayant la même fonctionnalité. Le type de dispositif est une caractéristique non négligeable d'un système de RA, mais il s'avère qu'il n'est pas suffisant et qu'il doit être complété par des caractéristiques complémentaires.

Les approches de caractérisation de la RA issues du domaine des IHM ne semblent donc pas répondre de manière individuelle à une caractérisation rigoureuse et complète. Toutefois, les différentes caractéristiques exposées contribuent à la description de ces systèmes. Ces approches sont donc complémentaires les unes des autres car elles adoptent des points de vue différents.

Du point de vue de l'approche du domaine de la synthèse d'images [Milgram 94], l'approche se limite à la description des systèmes basés sur un affichage combiné d'images réelles et virtuelles. Le retour d'effort et l'usage de son ne s'intègrent pas dans cette approche. De plus, le premier facteur, "*Extent of World Knowledge*" (*EWK*), qui permet de caractériser le niveau de connaissance que le système a de la scène affichée, permet en fait de déduire les fonctionnalités offertes par le système interactif. Par exemple, si le *EWK* d'un logiciel de dessin 3D est élevé, le système sera capable de faire des calculs de volumes. Par contre si ce *EWK* est faible, le système ne pourra qu'afficher la scène. Le point de vue de cette approche est donc centré sur l'ordinateur, mais il est aussi dépendant de la technologie utilisée, puisque des algorithmes de type segmentation automatique ou détection de contour peuvent être insérés dans le système. Par conséquent, le système peut avoir une connaissance aussi pointue d'un bureau réel filmé que d'un bureau entièrement modélisé par un logiciel de CAO et rendu sous forme de "fil de fer". Pourtant l'utilisateur pourra aisément faire la distinction entre les deux. Ce premier facteur décrit donc plutôt les capacités logicielles du système informatique que l'impact du système sur la perception visuelle. Il nous semble que cet aspect peut effectivement avoir une conséquence sur les choix de conception de la partie logicielle du

système, sans cependant avoir un impact sur l'analyse de l'utilisabilité de l'interaction induite par le système. Il en est de même pour son troisième facteur, "*Reproduction Fidelity*", qui est lié aux propriétés techniques du dispositif d'affichage. Notons toutefois que la prise en compte du point de vue choisi pour la représentation des données est essentielle dans l'interaction, comme nous l'illustrerons dans le chapitre VII.

Observons au travers de toutes ces approches le système de chirurgie du MIT, présenté au chapitre précédent, et qui aide le chirurgien à définir le contour d'une craniotomie à effectuer sur un patient. Les premières approches présentées mettent en évidence l'utilisation d'un flux vidéo [Azuma 97], [Feiner 93], [Noma 96], [Webster 97]. Une autre approche met en évidence le fait que l'utilisateur ne porte aucun dispositif particulier [Mackay 96]. L'approche du domaine de la synthèse d'images [Milgram 94] quant à elle, caractérise le flux vidéo produit. Aucune de ces trois approches ne semble retenir la situation particulière liée au contexte d'interaction. En effet, ce système est utilisé en milieu chirurgical. Le chirurgien est donc amené à regarder son patient via l'écran. Cet écran lui propose certes une vue du patient complétée par des données anatomiques, mais à aucun moment le chirurgien ne pourra percevoir directement son patient. Or la prise en compte du contexte d'interaction est un facteur crucial dans l'étude d'un système de RA.

En effet, les systèmes de RA sont des systèmes interactifs et l'utilisateur doit être considéré comme l'un des aspects principaux. Evaluer l'utilisabilité de l'interface proposée ou des dispositifs d'interaction mis en jeu sont des aspects reconnus comme étant nécessaires et inévitables lors de la conception et du développement de systèmes interactifs. Or aucune des approches de caractérisation présentée dans la littérature ne permettent de prendre en compte l'aptitude de l'utilisateur à utiliser correctement et efficacement le système.

## 4 Bilan des systèmes de RA

La revue des différentes approches pour caractériser la RA, met en exergue différents points de vue que le concepteur peut choisir d'adopter pour aborder l'analyse ou la conception d'un système. De notre revue de l'existant, nous distinguons trois points de vue : l'utilisateur, les dispositifs techniques, encore appelé le média et la forme des données ajoutées au monde réel par le système. Chaque approche de caractérisation de la RA s'appuie sur l'un de ces points de vue et propose un ensemble cohérent de caractéristiques ayant un impact sur l'utilisation d'un système de RA. Parmi ces caractéristiques, nous retenons les types de données ajoutées au monde réel, de même que les dispositifs choisis et leurs caractéristiques intrinsèques : ces deux ensembles de caractéristiques constituent les deux piliers de caractérisation d'un système de RA. A ceux-ci s'ajoutent les éléments de caractérisation du style de présentation décrits dans l'approche de [Milgram 94], qui caractérise les systèmes de RA dans lesquels l'utilisateur agit sur une représentation virtuelle et non dans le monde réel : le point de vue et le respect des échelles dans la présentation des données à l'utilisateur constituent deux caractéristiques pertinentes et complémentaires à celles déjà retenues.

Cependant, chaque approche de caractérisation révèle plusieurs limitations importantes. Il ressort entre autre qu'aucune des approches étudiées ne prend en compte le contexte d'interaction dans son ensemble. L'analyse de la complexité de l'interaction



Utilisateur-Système, liée à l'ajout de données dans l'environnement réel de l'utilisateur ou encore liée à la manipulation simultanée des mondes réel et virtuel, n'est pas non plus prise en compte par ces approches. Enfin nous avons montré que la caractérisation des systèmes existants selon ces approches mène parfois à des conclusions contradictoires.

Devant cette absence de consensus, chaque communauté, chaque concepteur choisit son point de vue de la RA pour travailler. Prenons pour preuve les systèmes de GMCAO étudiés au Chapitre I : leur conception repose sur les trois étapes de fonctionnement spécifiques au domaine : acquisition, raisonnement, action. Devant les larges possibilités qu'offre ce paradigme d'interaction, les études ne trouvent que très peu de points d'accord et cet axe de recherche reste un grand terrain vague à défricher.

Néanmoins, plusieurs formes de RA sont mises en évidence par notre revue des systèmes existants et des approches de caractérisation. Parmi les formes de RA identifiées, l'une regroupe les systèmes basés sur l'intégration de données du monde réel dans le monde virtuel et une autre regroupe ceux qui sont basés sur l'intégration de données virtuelles dans l'environnement physique réel de l'utilisateur. Le seul point commun entre toutes ces approches réside dans la combinaison du monde réel et du monde virtuel. C'est pourquoi plutôt que d'utiliser le terme de RA, nous adoptons dans la suite de ce mémoire le terme de **système mixte** pour désigner tout système interactif mêlant une partie du monde réel et une partie du monde virtuel.

A l'image de ce terme, qui regroupe différentes formes de RA, il convient d'établir un cadre de classification unifiant et prolongeant les approches de caractérisation existantes. L'identification de catégories de systèmes n'apporte ni modèle ni recommandation pour la conception de systèmes. Néanmoins, un cadre de classification permet d'envisager de nouveaux systèmes en explorant des catégories encore vierges ainsi que de caractériser et comparer les systèmes existants. De plus le rapprochement de systèmes au sein d'une catégorie peut permettre la mise en place d'éléments de conception et de réalisation récurrents, et propres à une catégorie donnée. Pour ce double objectif de classification et de conception, nous présentons au chapitre suivant une taxonomie des systèmes mixtes.

## 5 Références

[Ahlers 97]

Klinker, G., Ahlers, K., H. Breen, D., Chevalier, P.-Y., Crampton, C., Greer, D., S., Koller, D., Kramer, A., Rose, E., Tuceryan, M., Whitaker, R., "Confluence of Computer Vision and Interactive Graphics for Augmented Reality", dans la revue Presence: Teleoperators and Virtual Environments (Special issue on Augmented Reality), 6, 4, (1997), p. 433-451.

[Aukstakalnis 92]

Aukstakalnis, S., Blatner, D., Silicon, "*Mirage - The Art and Science of Virtual Reality*", Berkeley, CA, Peachpit Press, (1992).

[Azuma 97]

Azuma, R. T., "*A survey of Augmented Reality.*", dans la revue Presence: Teleoperators and Virtual Environments 6, 4, 1997, p.355-385

[Burdea 94]

Burdea, G., Coiffet, P., "*Virtual Reality Technology*", New York: John Wiley & Sons, (1994).

[Chen 98]

Chen, N., Yu, W., and Pretlove, J.R. "*Haptic interfaced telemanipulation*", Acte du Symposium SPIE on Intelligent Systems and Advanced Manufacturing, Telemanipulator and Telepresence Technologies V, vol. 3524, Boston, (1998).

[Coutaz 98]

Coutaz, J., Berard, F., Carraux, E., Crowlay, J., "*Early Experience with the Mediaspace CoMedi*", Acte de la Conférence EHCI'98, (1998).

[Cruz-Neira 93]

Cruz-Neira, C., Sandin, D., J., De-Fanti, T., A., "*Surround Screen Projection-Based Virtual Reality: The Design and Implementation of the Cave*", Acte de la Conférence SIGGRAPH'93, (1993), p. 135-142.

[Ellis 91]

Ellis, Gibbs, Rein, "*Groupware: Some Issues and Experiences*", communication de l'ACM, 34(1), (1991), p. 35-58.

[Feiner 93]

Feiner, S., MacIntyre, B., Seligmann, D., "*Knowledge-Based Augmented Reality*", Communication de l'ACM n°7, (1993), p. 53-61.

[Fitzmaurice 93]

Fitzmaurice, G., "*Situated Information Spaces and PalmTop Computers*", Communication de l'ACM n°7, (1993), p.39-51.

[Gaver 93]

Gaver, W., Sellen, A., Heath, C., Luff, P., "*One is Not Enough: Multiple Views in a MediaSpace*", Acte de la conférence INTERCHI'93, (1993), p. 335-341.

[Ishii 97]

Ishii, H., Ullmer, B., "*Tangible Bits : Towards Seamless Interfaces between People, Bits and Atoms*", Acte de la conférence CHI'97, Atlanta, (1997), p. 234-241.

[Kuzuoka 94]

Kuzuoka, H., Kosuge, T., Tanaka, M., "*GestureCam: a Video Communication System for Sympathetic Remote Collaboration*", Acte de la conférence CSCW'94, Chapel Hill - NC, USA, (1994), p. 35-43.

[Mackay 96]

Mackay, W.E., "*Réalité Augmentée : le Meilleur des Deux Mondes*", La Recherche n°285, (1996), p. 80-84.

[Mackay 98]

Mackay, W.E., Fayard, A.-L., Frobert, L., Médini, L., "*Reinventing the Familiar : an Augmented Reality Design Space for Air Traffic Control*", Acte de la conférence CHI'98, Los Angeles, (1998), p558-565.

[Milgram 94]

Milgram, P., Kishino, F., "*A Taxonomy of Mixed Reality Visual Displays*", *Transactions on Information Systems*, E77-D(12), (1994), p.1321-1329.

[NaviCam]

URL : <http://www.csl.sony.co.jp/person/rekimoto/navi.html>

[Noma 96]

Noma, H., Miyasato, T., Kishino, F., "*A palmtop display for dextrous manipulation with haptic sensation*", Acte de la conférence Human factors in computing systems, 1996, p.126-133.

[Okada 94]

Okada, K.-I., Maeda, F., Ichikawaa, Y., Matsushita, Y., "*Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design*", Acte de la conférence CSCW'94, (1994), p. 385-393.

[Rekimoto 95]

Rekimoto, J., Nagao, K., "*The World through the Computer : Computer Augmented Interaction with Real World Environments*", Acte du symposium UIST'95, (1995).

[Renevier 01]

Renevier, P., Nigay, L., "*Mobile Collaborative Augmented Reality: the Augmented Stroll*", à paraître dans les actes de la conférence EHCI'01, Toronto, Canada, (2001).

[Stifelman 96]

Stifelman, L., J., "*Augmenting real-world objects: a paper-based audio notebook*", dans le Tome II des actes de la conférence CHI '96, 1996, p. 199 - 200.

[Streitz 94]

Streitz, N., A., Geißler, J., Haake, J., M., Hol, J., "*DOLPHIN: Integrated Meeting Support across Local and Remote Desktop Environments and LiveBoards*", Acte de la conférence CSCW'94, (1994), p. 345-358.

[Szalavari 98]

Szalavari, Z., Eckstein, E., Gervautz, M., "*Collaborative Gaming in Augmented Reality*", Acte du symposium Virtual Reality Software and Technology, (1998), p. 195-204.

[Underkoffler 98]

Underkoffler, J., Ishii, H., "*Illuminating Light : An Optical Design Tool with a Luminous-Tangible Interface*", Acte de la conférence CHI'98, Los Angeles, (1998), p. 542-549.

[Webster 96]

Webster, A., Feiner, S., MacIntyre B., Massie, W., Krueger, T., "*Augmented Reality in Architectural Construction, Inspection, and Renovation*", présenté au congrès ASCE Computers in Civil Engineering, (1996), p. 913-919.

[Webster 97]

Webster, A., & al., "*A Touring Machine : Prototyping 3D Mobile Reality Systems for Exploring the Urban Environment*", Acte de la conférence ISWC'97, Cambridge, (1997), p. 74-81.

[Wellner 93]

Wellner, P., "*Interacting with paper on DigitalDesk*", *Communication de l'ACM*, (36), 7, (1993).

[Whirlpool 01]

URL : [http://www.whirlpool.com.au/WhatsNew/Internet\\_fridge.htm](http://www.whirlpool.com.au/WhatsNew/Internet_fridge.htm)

[Whitaker 95]

Whitaker, R., T., Crampton, C., Breen, D., E., Tuceryan, M., Rose, E., "*Object Calibration for Augmented Reality*", Acte de la conférence Eurographics'95, Vol. 14, 3, (1995), p. 15-28.



# **CHAPITRE III**

## **NOTRE TAXONOMIE DES SYSTEMES MIXTES**

## Sommaire

<b>Chapitre III : Notre Taxonomie des Systèmes Mixtes</b>	<b>83</b>
<b>1 Démarche d'analyse</b>	<b>83</b>
<b>2 Deux caractéristiques intrinsèques aux tâches</b>	<b>83</b>
2.1 Objet de la tâche	83
2.2 Type d'augmentation	86
<b>3 Illustration de notre taxonomie</b>	<b>87</b>
<b>4 Bilan de notre taxonomie</b>	<b>92</b>
<b>5 Références</b>	<b>96</b>

Nous avons vu au chapitre II que les approches de la littérature visant à caractériser les systèmes de mixtes mettent principalement en évidence le type des données fournies et les dispositifs d'interaction utilisés. Pour sa part, le domaine des Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO) propose une approche de caractérisation selon le degré d'investissement du système dans la réalisation de la tâche (chapitre I).

Nous avons démontré dans le chapitre I que de nombreux systèmes de GMCAO appartenaient à la même classe alors qu'ils présentaient des aspects différents du point de vue de l'interaction entre le clinicien et le système. Au chapitre II, nous avons également montré que la combinaison de deux critères, type de données et technologie, qui constituent les fondements des approches de caractérisation des systèmes mixtes, n'est pas non plus suffisante.

Un bilan analytique des approches existantes nous amène donc à proposer dans ce chapitre une nouvelle taxonomie pour les systèmes mixtes, dont les systèmes de GMCAO font partie. Cette taxonomie est destinée à permettre la classification, la comparaison, et aussi l'exploration de nouvelles solutions de conception pour des systèmes combinant une partie du monde réel ainsi qu'une partie du monde virtuel.

## **Chapitre III :**

### **Notre Taxonomie des Systèmes Mixtes**

#### **1 Démarche d'analyse**

Notre objectif est de définir une taxonomie des systèmes mixtes en adoptant le point de vue de l'interaction de l'utilisateur avec le système. La revue des systèmes existants des chapitres I et II souligne le fait que les systèmes mixtes impliquent un changement dans le mode d'interaction de l'utilisateur avec le monde réel ou le monde informatique. Plusieurs cas sont possibles : avec *KARMA*, l'interaction de l'utilisateur avec le monde réel est enrichie par la présentation de données, avec le *Tangible Geospace* l'usage de briques physiques simplifie l'interaction avec le système informatique, avec le *Digital Desk* les capacités d'interaction avec le monde réel sont complétées par l'ajout de services tels que le copier/coller. Nous avons donc choisi de centrer notre classification sur l'interaction : pour cela nous étudions un système mixte en considérant la réalisation d'une tâche. En effet l'interaction est différente d'une tâche à l'autre. Aussi un même système pourra appartenir à des classes différentes dans notre taxonomie, selon la tâche considérée. Aussi, les deux axes de classification que nous adoptons expriment donc chacun une caractéristique de la tâche considérée pour étudier le système. Le paragraphe suivant présente ces deux caractéristiques : l'objet de la tâche et le type d'augmentation. Nous illustrons ensuite le cadre de classification ainsi obtenu en positionnant les systèmes présentés dans les chapitres I et II.

#### **2 Deux caractéristiques intrinsèques aux tâches**

##### **2.1 Objet de la tâche**

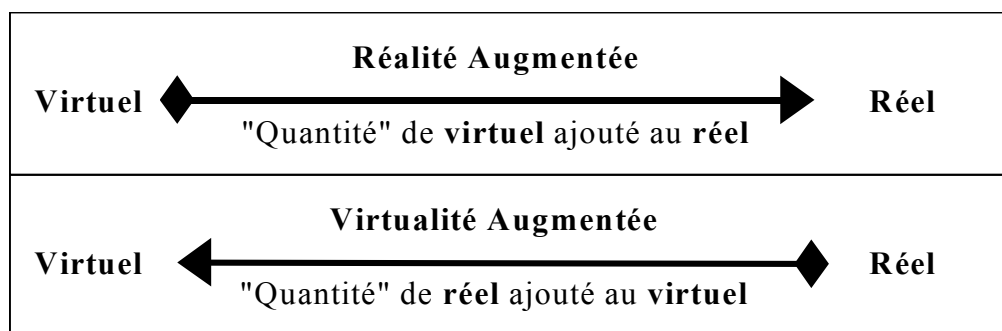
L'objectif des systèmes mixtes est de briser la frontière existant entre le monde réel, incluant l'utilisateur et son environnement, et le monde informatique. Une caractéristique



fondamentale des systèmes mixtes étant que des éléments du monde réel interviennent dans l'interaction, l'existence de ces deux mondes, réel et informatique, implique que l'objet de la tâche, c'est-à-dire l'entité sans laquelle la tâche n'a pas de sens, peut appartenir à l'un des deux mondes. Cette distinction est illustrée par le *Tangible Geospace*, où l'objet de la tâche est le plan numérisé du campus et l'interaction est basée sur la manipulation de briques physiques, et le *Musée Augmenté* dans lequel l'objet de la tâche est l'œuvre d'art, entité réelle.

Prendre en compte le fait que l'objet de la tâche puisse être réel ou virtuel est à mettre en relation avec l'approche exposée au chapitre II, visant à distinguer le cas où les données affichées sont principalement issues d'images réelles, du cas où elles sont construites par calculs [Milgram 94]. Un continuum allant du tout réel au tout virtuel en passant par la réalité augmentée, la virtualité augmentée puis la réalité virtuelle, est défini pour classer les dispositifs d'affichage de RA. Dans ce continuum, la distinction, réel ou virtuel, vise à caractériser un seul et même concept, le flux vidéo. Aussi, un seul continuum est suffisant. Cette notion de continuum exprime le fait qu'un ensemble continu de situations se trouve entre les deux extrêmes, l'un réel et l'autre virtuel.

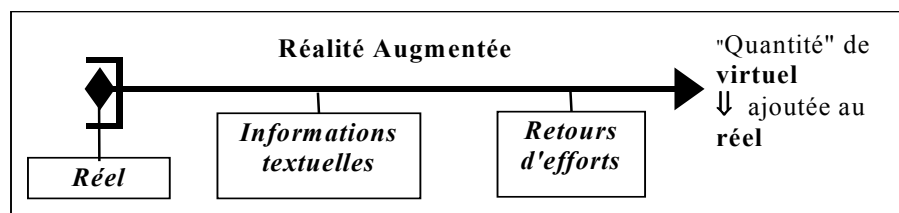
Dans notre distinction entre objet de la tâche réel et virtuel, les deux types de tâches identifiés sont bien différents : dans un cas une entité réelle est au centre de l'interaction, dans l'autre c'est une entité virtuelle. Nous avons donc défini deux continua, la **Réalité Augmentée** et la **Virtualité Augmentée**, présentés à la Figure 28. Dans les deux cas, plusieurs situations d'interaction peuvent intervenir. Ainsi l'utilisateur peut manipuler une entité réelle sans l'aide du système ou au contraire agir en synergie avec le système mixte sur l'entité réelle. De la même manière, l'utilisateur peut interagir avec une entité virtuelle sans intervention de l'environnement réel, c'est alors une IHM classique ou au contraire, l'environnement réel peut jouer un rôle dans la manipulation de l'entité virtuelle. Pour rendre compte de ces différentes situations d'interaction, nous affinons les deux continua en ajoutant une caractéristique quantitative : l'intrusion d'un monde dans l'autre.



**Figure 28** *Systèmes mixtes : deux continua pour distinguer "Réalité Augmentée" et "Virtualité Augmentée".*

Ainsi, le continuum de **Réalité Augmentée** permet de caractériser des systèmes dont l'objet de la tâche est dans le monde réel, comme la plupart des systèmes de chirurgie assistée par ordinateur (GMCAO) : la réalité est constituée de l'environnement réel de l'utilisateur. L'interaction avec l'environnement réel est augmentée par les capacités apportées par l'ordinateur et des actions dans le monde réel sont nécessaires pour réaliser la tâche. A l'origine de ce continuum se trouvent des situations dans lesquelles la réalisation des tâches ne s'appuie sur aucune des données que pourrait fournir le système. Il s'agit donc de situations d'interaction avec le monde réel qui ne requièrent pas de

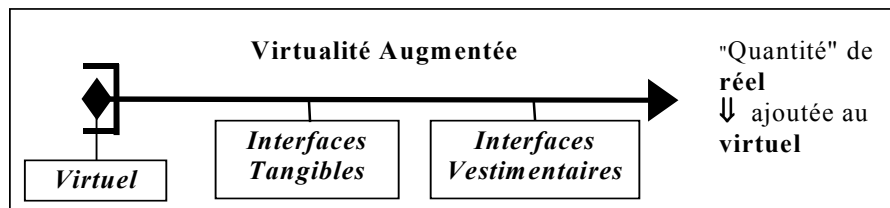
système informatique. Par conséquent l'origine du continuum ne fait pas partie des systèmes mixtes. Les premiers systèmes mixtes pris en compte par ce continuum sont des systèmes faiblement impliqués dans la réalisation de la tâche et participant à la tâche sous une forme qui pourrait être ignorée par l'utilisateur. A l'opposé, se trouvent les systèmes qui fournissent une aide importante pour la réalisation de la tâche. Entre ces deux extrêmes, la Réalité Augmentée peut insérer des éléments du monde virtuel de différentes façons dans le monde réel : le *Musée Augmenté* se base par exemple sur l'insertion de texte, tandis que d'autres systèmes contraignent l'interaction de l'utilisateur en lui fournissant un retour d'effort l'empêchant de sortir de certaines limites. L'évaluation de la quantité de virtuel apportée par le système reste relativement subjective et représente le niveau de participation du monde virtuel dans la réalisation de la tâche par l'utilisateur. On retrouve donc ici l'approche retenue pour la taxonomie des systèmes de GMCAO qui distingue les systèmes passifs, sans contrainte pour l'utilisateur, les systèmes semi-actifs, qui interagissent avec l'utilisateur, et les systèmes actifs réalisant de manière autonome une sous-partie de la tâche de l'utilisateur. Ainsi, il apparaît qu'un retour d'effort contraint le geste de l'utilisateur lors de la réalisation de la tâche. Par contre l'utilisation d'un casque semi-transparent par exemple ne fait qu'afficher des données que l'utilisateur peut choisir d'ignorer. Un système informatique produisant un retour d'effort fourni par le biais d'un bras robotisé est donc plus impliqué dans la réalisation de la tâche qu'un système produisant simplement une donnée textuelle. La Figure 29 illustre ce continuum de Réalité Augmentée avec plusieurs formes d'enrichissement de l'interaction de l'utilisateur avec une entité du monde réel.



**Figure 29** Continuum de "Réalité Augmentée" et différentes façons d'enrichir la réalisation d'une tâche dont l'objet est une entité réelle.

Le continuum noté **Virtualité Augmentée** est dédié aux systèmes dont l'objet de la tâche réside dans le monde informatique. Les systèmes considérés visent à rendre l'interaction plus "réaliste" en sortie ou plus naturelle en entrée. La virtualité représente le monde artificiel créé par l'ordinateur, ses systèmes informatiques et l'interaction de l'utilisateur avec cette virtualité est augmentée par l'utilisation d'objets appartenant à l'environnement réel de l'utilisateur. L'environnement physique permet de fournir à l'utilisateur des moyens d'interaction, en entrée ou en sortie, plus familiers ou alors plus transparents. Tout comme pour la Réalité Augmentée, il est possible d'insérer du réel dans un système de Virtualité Augmentée de différentes façons : les interfaces tangibles (TUI) se basent par exemple sur la manipulation de briques physiques, alors que l'informatique vestimentaire exploite les habits de l'utilisateur. L'approche du domaine de la synthèse d'images quant à elle insère des images réelles dans le modèle 3D d'une pièce affichée sur un écran. A l'origine de ce continuum se trouvent donc les tâches réalisées à l'aide de systèmes sans l'intervention d'éléments de l'environnement réel de l'utilisateur. Dans ce cas, l'interaction est exclusivement conduite à l'aide de dispositifs standards d'entrée-sortie (souris, clavier) et constitue une interaction homme-machine traditionnelle.

Comme pour la Réalité Augmentée, l'origine du continuum de la Virtualité Augmentée ne fait pas partie des systèmes mixtes. A l'extrême opposé, se trouvent les tâches dont l'objet est virtuel et dont la réalisation s'appuie largement sur des objets du monde réel, comme par exemple les IHM vestimentaires. Là encore cette notion de quantité demeure subjective et représente le niveau d'implication de l'environnement réel dans le monde virtuel. Les TUI par exemples ne s'appuient que sur des briques du monde réel. Par contre les IHM vestimentaires exploitent les habits de l'utilisateur, mais aussi ses mouvements. Dans ce dernier cas la part du monde réel présente sur laquelle repose l'interaction est plus importantes que les seules briques des TUI. La Figure 30 illustre ce continuum en le complétant par les deux formes d'implication de l'utilisateur que nous avons évoquées.



**Figure 30** Continuum de "Virtualité Augmentée" et différents types d'interaction pour enrichir la réalisation d'une tâche dont l'objet est une entité virtuelle.

Les deux continua présentés, "Réalité Augmentée" (RA) et "Virtualité Augmentée" (VA) distinguent des systèmes mixtes qui respectivement augmentent soit l'interaction avec le monde réel grâce à l'ordinateur, soit l'interaction avec l'ordinateur en exploitant des éléments du monde réel. Nous nous intéressons maintenant aux types d'augmentation de l'interaction fournis par le système : nous identifions donc l'apport de l'utilisation du système dans l'interaction Utilisateur-Système lors de la réalisation d'une tâche. Le paragraphe suivant présente les deux types d'augmentation envisagés.

## 2.2 Type d'augmentation

Considérons tout d'abord deux systèmes illustratifs. Dans le système *AESOP* [Sackier 93], système de chirurgie augmentée présenté précédemment, le robot joue le rôle d'un assistant, et permet au chirurgien d'exécuter une nouvelle action dans le monde réel : le pilotage du laparoscope. Un autre système de GMCAO utilisé pour aider le clinicien à définir le contour d'une craniotomie optimale [Grimson 96], matérialise la position et le modèle 3D de la tumeur dans le crâne du patient. Cette donnée est donc fusionnée avec le patient réel. La différence entre ces deux systèmes réside dans le type d'augmentation : *AESOP* augmente les capacités d'action du chirurgien, tandis que le système pour la *craniotomie* augmente la quantité de données pertinentes perceptibles par le chirurgien.

Selon la théorie de l'action de Norman [Norman 86], l'accomplissement d'une tâche présente deux gouffres pouvant se révéler problématiques : le gouffre de l'exécution et le gouffre de l'évaluation. L'aide fournie par le système mixte peut donc adresser l'un ou l'autre de ces deux gouffres. Pour reprendre nos deux exemples ci-dessus, le système d'aide à la *craniotomie* réduit le gouffre de l'évaluation en apportant dans le champ opératoire les données nécessaires à la définition du contour optimal de la craniotomie requise. Pour sa part *AESOP* comble le gouffre de l'exécution en libérant les mains du chirurgien lors de la manipulation du robot "assistant". Ceci constitue le second axe de notre taxonomie et identifie donc deux types d'augmentation : l'exécution augmentée et l'évaluation augmentée.

- **l'exécution augmentée** : Un système mixte augmente l'exécution dans le cas où il réduit le gouffre de l'exécution par le biais de l'ajout d'actions nouvelles ou par l'amélioration de la qualité de réalisation de la tâche considérée.

Un système mixte, dont l'objet de la tâche est réel, augmentant l'exécution permet à l'utilisateur d'effectuer des tâches dans le monde réel d'une façon nouvelle, comme le système *AESOP*. L'ouverture automatique d'une porte après avoir vérifié automatiquement l'identité d'une personne ainsi que ses droits d'accès, par le biais d'un badge par exemple, constitue également un exemple d'exécution augmentée. L'utilisateur n'a aucun code à taper, ni aucun verrou à ouvrir. C'est le principe des *Active Badge* [Want 92].

Un système mixte, dont l'objet de la tâche est virtuel, augmentant l'exécution consiste à permettre à l'utilisateur d'interagir avec l'ordinateur de façon nouvelle grâce à des éléments du monde réel. C'est le cas du *Tangible Geospace* [Ishii 97].
- **l'évaluation augmentée** : Un système mixte augmente l'évaluation dans le cas où il réduit le gouffre de l'évaluation par le biais de l'ajout de données qui ne sont pas perceptibles lors de la réalisation de la tâche sans le système. Il s'agit sans nul doute du type d'augmentation le plus communément mis en œuvre dans les systèmes actuels : de nombreux systèmes présentés aux deux chapitres précédents illustrent ce type d'augmentation.

Pour un système mixte, dont l'objet de la tâche est réel, des données pertinentes pour la tâche sont ajoutées à l'environnement de l'utilisateur. Le système d'aide à la définition du contour d'une *craniotomie* ou encore le *Musée Augmenté* [Rekimoto 95] en sont des exemples.

Un système mixte, dont l'objet de la tâche est virtuel, augmente la perception en rendant des scènes virtuelles plus réalistes en insérant des images du monde réel par exemple. C'est notamment le cas dans le système *Interior Design* [Whitaker 95].

Pour classer un système mixte dans notre espace, il est donc tout d'abord crucial de fixer la tâche à étudier. L'identification de la tâche permet de déterminer rapidement si l'objet de cette tâche est réel ou virtuel. Il convient ensuite de s'interroger sur le ou les types d'augmentation induits par le système. En effet les deux phases de réalisation de la tâche selon Norman sont indépendantes. Il est donc naturel d'envisager la possibilité qu'un même système réduise simultanément les deux gouffres d'exécution et d'évaluation au cours de la réalisation d'une tâche donnée. Le paragraphe suivant illustre notre espace de classification avec les systèmes mixtes présentés dans les chapitres I et II, puis résume l'analyse taxinomique de ces exemples dans notre espace de classification.

### 3 Illustration de notre taxonomie

Dans ce paragraphe, nous illustrons tout d'abord notre espace taxinomique avec les systèmes de réalité augmentée (RA) et de virtualité augmentée (VA) présentés dans le chapitre I, c'est-à-dire les systèmes de GMCAO. Nous illustrons ensuite ce cadre de classification avec les systèmes présentés au chapitre II, les systèmes développés dans le domaine des IHM.

A la Figure 31, nous illustrons notre cadre de classification avec les systèmes mixtes de GMCAO, présentés au chapitre I. Nous détaillons l'élaboration de cette classification dans la suite de ce paragraphe.

Parmi les systèmes présentés au chapitre I, les systèmes d'endoscopie virtuelle de Frankenthaler, Jolesz, Mori, Oppenheimer et Ikuta sont des systèmes permettant de simuler une endoscopie. De même *EchoSim* permet à un clinicien d'apprendre à interpréter une image échographique ainsi que les organes situés autour du cœur. Dans tous ces cas, le clinicien travaille sur des données médicales stockées dans l'ordinateur. L'objet de la tâche est donc virtuel. Ces systèmes relèvent donc tous de la classe Virtualité Augmentée. Pour les quatre premiers, il s'agit d'endoscopies virtuelles qui ne font que présenter des données au clinicien : au cours de la simulation, seule l'évaluation du clinicien est augmentée par le système. Il en va de même pour *EchoSim*. Au lieu de n'observer qu'une image échographique, le système lui permet d'avoir connaissance de paramètres et données réelles complémentaires. Par contre l'endoscopie d'Ikuta se distingue des précédents systèmes dans la mesure où il fournit un retour d'effort. L'apprentissage n'est donc pas basé uniquement sur un retour d'informations visuelles. Le système participe aussi à l'apprentissage du geste puisqu'il peut contraindre ou limiter les déplacements du clinicien apprenant. Ce système de VA augmente donc simultanément l'évaluation et l'exécution du clinicien.

Tous les autres systèmes présentés dans le chapitre I sont des systèmes de Réalité Augmentée, leur but étant de participer à une intervention chirurgicale ou de fournir des données de guidage pendant que le chirurgien soigne un patient. Le patient constitue une entité du monde réel. Analysons donc pour chacun, l'apport du système par rapport aux deux phases de réalisation de la tâche : évaluation et exécution.

*AESOP* réalise les mouvements d'un laparoscope à la place du chirurgien. Quant au *Force Sensing Retractor* (FSR) il permet de détecter l'application de forces trop importantes sur le patient et améliore la réalisation de la tâche du chirurgien en suspendant l'application de ces forces. Les *Templates* et les gabarits du système d'implantologie dentaire fournissent un guide physique au clinicien, lui laissant une seule dimension de mobilité. Cette dimension est en plus contrainte en longueur, dans la mesure où l'outil chirurgical introduit dans le guide ne peut pas dépasser une certaine position. En conséquence, aucune donnée de guidage visuelle n'est requise. Ces quatre systèmes font donc partie des systèmes qui augmentent l'exécution de la tâche du chirurgien sans augmenter la phase d'évaluation de la réalisation de la tâche.

*Robodoc* et *Minerva* réalisent une partie de l'intervention à la place du chirurgien en produisant un résultat plus précis et en permettant à l'utilisateur de visualiser l'avancée de l'outil chirurgical. *PADyC* et *ACRobot* n'agissent pas de manière autonome sur le patient, mais permettent au clinicien de percevoir la trajectoire à suivre, grâce à un retour d'effort. Dans une moindre mesure que *Robodoc* et *Minerva*, les deux systèmes participent aussi à la tâche puisqu'ils la contraignent physiquement, par moment au moins. Enfin ils procurent également un retour visuel relatif à l'avancée de l'outil à l'intérieur du patient. Ces quatre systèmes enrichissent donc les deux phases de réalisation de la tâche du clinicien : la perception et l'exécution. Le robot de Van Ham, *Orthosista* et le *Remote Sensor Unit* (RSU) ont des comportements un peu différents des quatre derniers systèmes. Le fonctionnement d'*Orthosista* est similaire au *Templates* et au système d'implantologie

dentaire, mais il procure en plus des données quant à l'avancée de l'outil à l'intérieur du patient. Mis à part le type de contraintes appliquées, le robot de Ham agit de la même façon qu'*Orthosista*. Quant au *RSU*, il réalise à distance le déplacement des outils chirurgicaux et procure un retour d'effort au chirurgien. Compte tenu de ces observations, ces trois systèmes augmentent simultanément l'exécution et l'évaluation. Ils appartiennent donc à la même classe que *Robodoc*, *Minerva*, *PADyC* et *ACRobot*.

Enfin, tous les autres systèmes de GMCAO présentés au chapitre I offrent des données visuelles au clinicien. Ces données peuvent être des informations de stratégie, anatomiques ou fonctionnelles. Elles sont alors présentées sur un écran (*HipNav*, *visée pédiculaire*, *Craniotomie*, Julliard), ou par le biais du principe d'*Image Overlay* comme l'utilisent Peuchot et Blackwell pour la seconde version de *HipNav*. Le système de Edwards se base quant à lui sur l'utilisation du *microscope* traditionnellement utilisé au cours de ses interventions et Fuchs utilise un casque opaque dans lequel est affiché un flux vidéo contenant une image du patient ainsi que les données de guidage. Les données fournies peuvent aussi enrichir ou améliorer la vue habituelle que le clinicien aurait de structures anatomiques. C'est notamment le cas du laparoscope avec vue arrière fournie par Scholz ou encore du laparoscope à vision stéréoscopique de Shahidi. Dans tous ces cas, le système fournit des données au moyen de dispositifs d'affichage variés. Seule la phase d'évaluation est donc augmentée par le système.

Objet de la Tâche		monde réel (Réalité Augmentée)	<i>AESOP</i> <i>FSR</i> (*) <i>Template</i> , <i>ID</i> (***)	<i>ACRobot</i> , <i>PADyC</i> <i>Robodoc</i> , <i>Minerva</i> <i>Orthosista</i> , <i>Ham</i> <i>RSU</i> (**)	
		monde virtuel (Virtualité Augmentée)	Aucun système identifié	<i>Endoscopies</i> <i>virtuelles</i> , <i>EchoSim</i>	<i>Simulateur</i> <i>d'endoscopie à</i> <i>retour d'effort</i>
			Exécution	Evaluation	Exec. + Eval.
		Type d'Augmentation			

*HipNav*, *Craniotomie*,  
*Visée Pédiculaire*, *Image Overlay*,  
*Guidage Audio*, *Microscope*,  
*Julliard*, *Fuchs*, *Scholz*, *Shahidi*

(\*) *FSR* : *Force Sensing Retractor* (\*\*) *RSU* : *Remote Sensor Unit* (\*\*\*) *ID* : *Implantologie Dentaire*

**Figure 31** Positionnement de systèmes mixtes du domaine des GMCAO dans notre cadre de classification.

Nous constatons à la Figure 31 qu'aucun système de GMCAO n'illustre la classe des systèmes de Virtualité Augmentée dans lesquels l'exécution est augmentée par l'usage d'entités du monde réel. Plus exactement, aucun système permettant une activité chirurgicale n'appartient à cette classe. Néanmoins [Hinckley 94] présente un système dont le but est la recherche, dans un volume de données anatomiques tridimensionnelles, d'un plan de coupe optimal. Au lieu de devoir manipuler une souris pour positionner un plan en 3D dans le modèle affiché à l'écran ou un formulaire lui demandant de préciser les coordonnées des points du plan, le clinicien manipule une plaque en verre localisée dans l'espace, qui représente le plan de coupe. Une boule en résine, également repérée dans l'espace, représente le modèle affiché à l'écran. En déplaçant la plaque de verre par

rapport à la boule en résine, le plan de coupe est mis à jour à l'écran. Le clinicien peut donc visualiser très facilement plusieurs plans avant de choisir celui qui lui semble être optimal.

Étudions maintenant les systèmes mixtes présentés au chapitre II, les systèmes mixtes du domaine de l'IHM.

Avec les systèmes tels que les *Active Badge*, *KARMA*, le *Musée Augmenté*, *Arc*, *Chameleon*, *Gesture Cam* et *l'Audio Notebook*, la tâche est facilement identifiable. Les *Active Badge* permettent l'ouverture d'une porte, dans *KARMA* il s'agit de réparer l'imprimante, le *Musée Augmenté* aide l'utilisateur à observer une œuvre d'art. *Chameleon* complète une carte géographique, *Gesture Cam* aide un technicien à réparer un vélo et *l'Audio Notebook* complète la relecture de notes manuscrites. Dans ces sept systèmes, l'objet de la tâche se situe dans le monde réel. Au niveau du type de l'augmentation, les *Active Badge* permettent l'ouverture de la porte à un utilisateur autorisé sans que il/elle n'ait besoin de presser sur un bouton ou de composer un code d'accès. Ce système est, selon nos définitions, un système de Réalité Augmentée qui enrichit la méthode de réalisation de la tâche : il augmente donc l'exécution. Par contre, les six autres systèmes ajoutent des données au monde réel pour tenter d'adresser le gouffre de l'évaluation de Norman. Ce sont donc des systèmes de Réalité Augmentée augmentant l'évaluation de l'utilisateur. Enfin, dans les *Mediaspace*, la tâche consiste à être en contact avec d'autres membres du même groupe de travail, répartis dans différents endroits. L'objet de la tâche est donc réel, il s'agit des bureaux respectifs de chacun. Les *Mediaspace* permettent également aux utilisateurs de percevoir les bureaux distants : il s'agit donc également de systèmes de réalité augmentée, qui augmentent l'évaluation.

Le *Tangible Geospace*, *Illuminating Light*, *Cave* et *Interior Design* sont dédiés à des tâches dont l'objet est résolument dans le monde virtuel. Le *Tangible Geospace* permet à l'utilisateur de manipuler les caractéristiques d'affichage d'un fichier de données tel qu'un plan par exemple. *Illuminating Light* présente le comportement et les caractéristiques d'un faisceau lumineux théorique, compte tenu de la position d'éléments optiques. *Cave* projette des données scientifiques 3D stéréoscopiques et *Interior Design* combine des modèles 3D de meubles avec des photos réelles de meubles existants. Au niveau de l'augmentation proposée, le *Tangible Geospace* et *Illuminating Light* permettent de manipuler physiquement des briques représentant respectivement des bâtiments du plan projeté et des éléments optiques. *Cave* quant à lui permet de conserver l'interaction sociale naturelle existante entre les utilisateurs immergés dans un monde virtuel et *Interior Design* rend la scène présentée plus réaliste. En résumé, ces quatre systèmes sont donc des cas de Virtualité Augmentée, mais le *Tangible Geospace* et *Illuminating Light* augmentent l'exécution d'une tâche informatique alors que les deux autres augmentent l'évaluation.

Considérons maintenant le *Digital Desk* et le jeu de *Mah-Jong*. Dans le *Digital Desk*, nous considérons ici, le cas où l'utilisateur dispose sur son bureau d'un formulaire à remplir. Ce formulaire peut être partiellement rempli à la main et complété par des parties de documents électroniques. Au final, il est donc composé d'encre réelle et d'encre virtuelle et doit être imprimé. La tâche consiste à remplir le formulaire en utilisant éventuellement des parties de documents virtuels. Ce formulaire rempli ne peut donc que résulter de l'impression d'un fichier contenant une image du formulaire original et de

l'encre réelle portée sur ce formulaire ainsi que des parties virtuelles attachées sur ce document. Par conséquent le formulaire est un support pour spécifier des données au système. L'objet de la tâche est donc constitué du fichier informatique dont l'utilisateur devra par la suite demander l'impression. Ce système est donc un système de Virtualité Augmentée. Comme il repose sur des formulaires réels pour créer le fichier, le système fourni donc à l'utilisateur une nouvelle façon d'entrer des données et à ce titre, augmente donc l'exécution. La possibilité d'utiliser un stylo réel et une écriture traditionnelle renforce cet aspect. En parallèle, l'utilisateur visualise le résultat des actions de copier/coller directement sur le bureau réel. Le système utilise donc le monde réel pour rendre perceptibles des données. Selon nos définitions précédentes, il s'agit donc d'un système qui augmente également l'évaluation de l'utilisateur. Dans le contexte de la tâche considérée, le *Digital Desk* est donc un système de Virtualité Augmentée qui augmente simultanément les phases de perception et d'exécution. Dans le cas du jeu de *Mah-Jong*, le but de la tâche consiste à manipuler les dominos virtuels. L'objet de la tâche est donc constitué des dominos virtuels joués ou à jouer : l'objet de la tâche est donc virtuel. La manipulation de ces dominos étant basée sur l'utilisation de stylets et de PIP -Personal Interaction Panel- réels, l'exécution de la tâche est augmentée. Enfin, la perception de ces données est faite au niveau d'un objet du monde réel, la table de jeu ou le PIP de chacun. Par conséquent ce système mixte est un système de Virtualité Augmentée qui augmente simultanément les phases de exécution et d'évaluation lors de la réalisation de la tâche de manipulation des dominos.

En synthèse, les deux caractéristiques "Objet de la tâche" et "Type d'augmentation" permettent d'identifier six classes de systèmes que la revue des systèmes mixtes présentés au chapitre II, nous permet d'illustrer à la Figure 32. Toutefois, nous n'avons pas identifié de systèmes qui permettraient d'illustrer la dernière classe. Ce système serait de la Réalité Augmentée (objet de la tâche réel), augmenterait la perception de l'utilisateur (par ajout de données visuelles, haptiques, sonores) ainsi que l'exécution (perfectionnement de la tâche ou ajout de nouvelles actions contribuant à la réalisation de la tâche). Nous présentons ci-après une façon empirique d'imaginer un tel système.

<b>Objet de la Tâche</b>	monde réel (Réalité Augmentée)	<i>Active Badge</i>	<i>Karma, Chameleon Arc, GestureCam Musée Augmenté Audio Notebook, Mediaspace</i>	Aucun système identifié ...
	monde virtuel (Virtualité Augmentée)	<i>Tangible Geospace Illuminating Light</i>	<i>Interior Design CAVE</i>	<i>DigitalDesk Mah Jong</i>
		<b>Exécution</b>	<b>Evaluation</b>	<b>Exec. + Eval.</b>
<b>Type d'Augmentation</b>				

**Figure 32** Positionnement de systèmes mixtes du domaine des IHM dans notre cadre de classification.

Cette approche consiste à partir d'un système de VA augmentant l'exécution et l'évaluation. Il convient alors que l'objet de la tâche réalisée à l'aide de ce système soit réel. Par exemple dans le cas du *Digital Desk*, l'objet de la tâche est virtuel car la



production du formulaire complété passe nécessairement par la création d'un fichier informatique. Il suffirait donc que le formulaire soit modifié directement sans passer par la génération d'un fichier. Pour se faire, il faudrait envisager la réalisation d'un projecteur d'encre en remplacement du projecteur vidéo actuellement utilisé : le système serait alors le *DigitalDesk-Jet*, par analogie aux imprimantes Deskjet. La faisabilité technique et l'utilisabilité de notre *DigitalDesk-Jet*, ne sont bien entendu pas garanties par cette classification... Remarquons également que si le *Digital Desk* devient par exemple une borne d'information pour apprendre aux visiteurs à remplir leur feuille d'impôts, le formulaire violet n°2042 et les données de la carte d'identité du visiteur sont requis par le *Digital Desk*. Le *Digital Desk* peut copier/coller les éléments de l'identité requis sur le feuillet violet, puis indiquer au visiteur quelles cases compléter manuellement. Dans ce cas, la tâche n'est qu'une démonstration. Elle consiste donc à remplir le feuillet et visualiser le résultat avec le visiteur. Il n'y a donc pas besoin d'imprimer le résultat et l'objet de la tâche est dans ce cas un objet réel. En considérant cette tâche, qui d'ailleurs est exactement celle mise en œuvre par les tableaux blanc électroniques, le *Digital Desk* fait alors partie des systèmes de Réalité Augmentée qui augmentent l'évaluation et l'exécution. Il appartient donc pour cette tâche à la même classe que le *DigitalDesk-Jet*, et illustre la possibilité qu'un système a de migrer d'une classe à l'autre selon la tâche considérée.

Enfin, les derniers exemples que nous citerons pour illustrer la classe des systèmes de réalité augmentée qui augmentent l'exécution et l'évaluation, sont les systèmes de télémanipulation. En effet ils incluent nécessairement un robot distant. De ce fait, ils augmentent l'exécution de l'utilisateur. De plus, l'utilisateur reçoit un retour d'informations visuelles portant sur une entité distante. Les systèmes de télémanipulation permettent donc d'augmenter simultanément l'évaluation et l'exécution.

## 4 Bilan de notre taxonomie

Par rapport aux approches existantes visant à caractériser les systèmes de RA, notre taxonomie permet de prendre en compte l'utilisateur et son interaction avec le système utilisé. L'avantage de cette approche est qu'elle caractérise la principale modification induite par le système mixte : la réalisation de la tâche. En effet, celle-ci est complétée dans les systèmes mixtes en exploitant l'environnement réel de l'utilisateur ainsi que les comportements habituels de l'utilisateur dans son environnement.

De plus, notre taxonomie permet d'identifier deux types bien différents de systèmes mixtes : la réalité augmentée, où l'objet de la tâche est une entité de l'environnement physique de l'utilisateur, c'est-à-dire du monde réel, et la virtualité augmentée où l'objet de la tâche est une entité virtuelle dont la manipulation ou la perception s'appuie sur des outils du monde réel.

Dans notre approche, le point crucial réside donc dans l'identification de la tâche supportée par le système et que l'on souhaite analyser ou concevoir. En effet selon la tâche étudiée, le système peut migrer d'une classe vers une autre. Prenons par exemple le cas des *Active Badge*. La tâche que nous avons étudiée précédemment visait à permettre l'ouverture automatique d'une porte lorsque l'utilisateur se présente devant celle-ci. Supposons maintenant que ce même badge entraîne le lancement automatique du "mailer" de l'utilisateur lorsqu'il entre dans son bureau et qu'il a reçu un courrier urgent. L'objet de la tâche n'est donc plus la porte, mais le courrier que l'utilisateur doit lire : c'est une entité

virtuelle, par conséquent les *Active Badge* font partie, dans ce cas, des systèmes de Virtualité Augmentée.

Par ailleurs, l'identification du type d'augmentation mis en œuvre dans le système permet facilement de comparer deux systèmes et d'envisager des versions différentes d'un système donné. C'est ce que nous avons illustré par le biais du *Digital Desk-Jet*. De la même manière, l'identification de l'objet de la tâche met en évidence l'entité sur laquelle l'utilisateur est amené à agir. Ceci est particulièrement intéressant dans le domaine de la chirurgie assistée par ordinateur. En effet les systèmes de simulation et maintenant d'enseignement, sont de plus en plus largement basés sur l'utilisation de structures anatomiques réelles. Par exemple, la manipulation d'un bassin réel couplée à l'utilisation d'un casque semi-transparent, permet aux chirurgiens novices de visualiser en 3D la position des organes autour de l'os. L'os étant réel, les cliniciens peuvent manipuler les données par simple rotation de l'os ; ils peuvent également visualiser le passage "virtuel" d'un foret au travers des différents tissus et structures anatomiques environnantes avant d'atteindre l'os réel. Les systèmes sont dans ce cas typiquement des systèmes de Virtualité Augmentée, mais l'analogie avec les systèmes de Réalité Augmentée possiblement existants permet d'établir un lien rapide et de faciliter la re-conception de ce système en identifiant rapidement les points de différence.

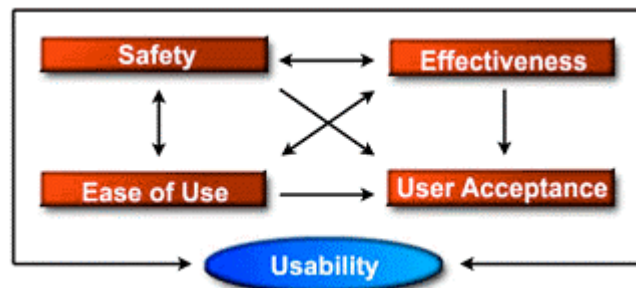
Une taxonomie ne permet pas que de distinguer les systèmes, de les regrouper et de permettre leur comparaison selon un point de vue établi. Elle peut également permettre de concevoir de nouveaux systèmes, soit en remplissant des classes non encore illustrées par des systèmes actuels, soit en faisant évoluer des systèmes existants d'une classe vers une autre, leur ajoutant ainsi toutes les caractéristiques de la nouvelle classe. Cependant, au niveau taxinomique, l'utilisabilité des solutions envisagées ne peut être garantie. En effet, une fois les caractéristiques des nouveaux systèmes identifiées, il convient de s'intéresser plus en détail à l'interaction induite et à ses caractéristiques. En se référant à un cycle de vie du logiciel, c'est lors de la phase de conception générale que de telles analyses peuvent être amorcées.

Dans cet objectif, nous présentons des éléments de conception au chapitre suivant. L'établissement d'une méthode de conception pour les systèmes mixtes n'est sans doute pas crucial pour des prototypes de laboratoire. En effet ces derniers constituent des études de techniques spécifiques au domaine d'utilisation telles que la segmentation d'images ou encore l'alignement des données virtuelles sur le monde réel. A l'opposé, les besoins en terme de conception sont particulièrement nécessaires dans le cas de systèmes dont l'application ou le domaine d'utilisation sont concrets et doivent répondre à des critères d'ergonomie et d'utilisabilité précis. C'est par exemple le cas des systèmes de GMCAO en ce qui concerne le développement des modalités d'interaction entre le chirurgien et le reste du système. En effet, avant de pouvoir utiliser un système de GMCAO, celui-ci doit suivre un protocole de validation clinique, prouvant son efficacité clinique et sa sécurité. Ces évaluations doivent avoir lieu sur modèles physiques, puis sur cadavres et éventuellement sur animaux vivants et enfin sur patient. Depuis peu, l'augmentation du nombre de ces systèmes réellement utilisés en bloc opératoire incite les concepteurs et les cliniciens à s'interroger sur les responsabilités des différents intervenants en cas de problèmes graves causés au patient. A cet égard, la "*Food and Drug Administration*" -

FDA- recommande donc, au cours de la conception du système, de compléter la prise en compte des aspects de sécurité et d'efficacité par des aspects liés à l'utilisateur final du produit : facilité d'utilisation et acceptabilité par l'utilisateur. L'ensemble de ces quatre éléments à prendre en compte lors de la conception est regroupé par la FDA sous le terme d'utilisabilité [Babbit 01] et se compose des aspects suivants :

- "**Effectiveness**" : le système doit permettre d'obtenir des résultats significativement meilleurs pour le patient sur le plan thérapeutique.
- "**Safety**" : le fonctionnement du système doit être garanti sans risque pour le patient et les membres de l'équipe médicale.
- "**Ease of use**" : le système interactif doit être facile d'usage et intuitif.
- "**Acceptance by the user**" : l'interface du système doit être satisfaisante pour le clinicien. On retrouve ici le terme de "*user-friendly*", initialement utilisé dans le domaine des IHM pour caractériser l'utilisabilité des interfaces homme-machine.

La prise en compte de ces quatre aspects est résumée à la Figure 33.



**Figure 33** *Quatre composantes de l'utilisabilité.*  
Schéma issu de <http://www.usabilitymdx.com>.

Notre approche résolument centrée sur l'utilisateur concerne les deux derniers points soulignés par la FDA : facilité d'utilisation ("*Ease of Use*") et acceptabilité par l'utilisateur ("*User Acceptance*"). Dans ce cadre, la conception ergonomique de l'interaction est requise.

Or, la conception des IHM de GMCAO, et en particulier la phase de développement de la présentation des IHM, prennent relativement peu en compte l'utilisateur final du système. Il est donc nécessaire de proposer une aide à la conception des interfaces de ces systèmes. Mise en œuvre concrète de la réalité augmentée, les principes de conception des interfaces de GMCAO doivent s'appuyer sur les principes de conception de la RA. Cependant nous avons montré, au chapitre II, l'existence d'un flou dans la caractérisation des systèmes de RA. Par conséquent, les principes de conception de la RA ne sont que très partiellement développés. Dans ce contexte il convient de compléter notre approche taxinomique par des éléments de conception. Ces éléments doivent s'appuyer tant sur les approches ergonomiques du milieu médical, que sur des approches de conception du domaine des IHM.

Dans les chapitres suivants, nous apportons des réponses de conception en présentant ASUR, une notation de description des systèmes mixtes, centrée sur la tâche et l'interaction de l'utilisateur avec le système au cours de la réalisation de cette tâche.

De plus nous présentons des éléments de conception basés sur l'identification puis l'expression dans ASUR de propriétés ergonomiques pertinentes pour l'étude de

l'utilisabilité d'un système mixte. Notre approche de conception vise donc à combler le manque actuel de la réalité augmentée en terme de conception. Comme mentionné précédemment, notre domaine applicatif est la chirurgie assistée par ordinateur, qui constitue un des domaines mettant en œuvre concrètement le paradigme de réalité augmentée. Néanmoins notre approche méthodologique reste généralisable et dédiée aux systèmes mixtes, incluant donc la réalité augmentée et la virtualité augmentée.

## 5 Références

[Babbit 01]

Babbitt, B., "*Medical Device Usability Engineering*", Workshop de la conférence MMVR'2001, Newport Beach - CA, USA, (2001).

[Grimson 96]

Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells, W.M., Kikinis R., "*An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization*", IEEE Transactions on Medical Imaging, 15(2), (1996), p. 129-140.

[Hinckley 94]

Hinckley, K., et al., "*Passive Real-World Interface Props for Neurosurgical Visualization*", Actes de la conférence CHI'94, (1994), p.452-458.

[Ishii 97]

Ishii, H., Ullmer, B., "*Tangible Bits : Towards Seamless Interfaces between People, Bits and Atoms*", Acte de la conférence CHI'97, Atlanta, (1997), p. 234-241.

[Norman 86]

Norman, D., "*Cognitive Engineering*", In *User Centered System Design, New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, (1986), p.31-61.

[Milgram 94]

Milgram, P., Kishino, F., "*A Taxonomy of Mixed Reality Visual Displays*", In *IEICE Transactions on Information Systems*, E77-D(12), (1994), p.1321-1329.

[Rekimoto 95]

Rekimoto, J., Nagao, K., "*The World through the Computer : Computer Augmented Interaction with Real World Environments*", Acte du symposium UIST'95, (1995).

[Sackier 93]

Sackier, J., et al., "*Computer-Aided Guidance for Endoscopic Surgery*", American College of Surgeons, San Francisco, (1993).

[Want 92]

Want, R., Hopper, A., Falcao, V., Gibbons, J., "*The Active Badge Location System*", dans *ACM Transactions on Information Systems*, 10(1), (1992), p. 91-102.

[Whitaker 95]

Whitaker, R., T., Crampton, C., Breen, D., E., Tuceryan, M., Rose, E., "*Object Calibration for Augmented Reality*". Acte de la conférence Eurographics'95, Vol. 14, **3**, (1995), p. 15-28.

## **Partie B :**

# **L'ESPACE SOLUTION : CONCEPTION ET REALISATION**

---



# **CHAPITRE IV**

## **ASUR : SUPPORT A LA CARACTERISATION DES SYSTEMES MIXTES**



# Sommaire

<b>Chapitre IV : ASUR : Support à la Caractérisation des Systèmes Mixtes</b>		<b>101</b>
<b>1 Composants</b>		<b>102</b>
1.1 Utilisateur et Système informatique		102
1.2 Objets réels		102
1.3 Lien entre le monde réel et monde virtuel		102
1.4 Règles d'utilisation et illustration		103
<b>2 Relations</b>		<b>104</b>
<b>3 Caractéristiques de l'interaction Utilisateur-Système Mixte</b>		<b>106</b>
3.1 Caractéristiques des composants		106
3.2 Caractéristiques des relations		108
<b>4 Exemples de descriptions ASUR</b>		<b>113</b>
4.1 Visée Pédiculaire, ACRobot et Craniotomie Assistée par Ordinateur		113
4.1.1 Composants et Relations ASUR		113
4.1.2 Caractéristiques des composants		114
4.1.3 Caractéristiques des relations		115
4.1.4 Apports de la description ASUR		116
4.2 KARMA et le Musée Augmenté		117
4.2.1 Composants et relations ASUR		117
4.2.2 Caractéristiques des composants		118
4.2.3 Caractéristiques des relations		118
4.2.4 Apports de la description ASUR		119
4.3 Tangible Geospace et Digital Desk		119
4.3.1 Composants et relations ASUR		119
4.3.2 Caractéristiques des composants		120
4.3.3 Caractéristiques des relations		121
4.3.4 Apports de la description ASUR		121
<b>5 Intérêts et limitations</b>		<b>122</b>
<b>6 Références</b>		<b>125</b>

Suite au bilan analytique que nous avons dressé dans les chapitres I et II, illustrant la diversité des approches de caractérisation des systèmes mixtes, nous avons présenté au chapitre précédent un cadre fédérateur pour l'étude et la classification des systèmes mixtes. Notre cadre de classification, centré sur la tâche de l'utilisateur, met particulièrement en évidence l'existence de deux types d'interaction entre l'utilisateur et le système : d'une part une interaction avec des entités du monde réel, d'autre part une interaction avec des entités du monde virtuel. Cette dualité de l'interaction est un point original, propre aux systèmes mixtes. Aussi lors de la conception, il convient de cerner exactement ce nouveau type d'interaction.

Dans cet objectif, nous avons développé la notation **ASUR**, acronyme désignant "*Adaptor*", "*System*", "*User*", "*Real entities*". Notation de description des composants d'un système interactif, ASUR décrit également l'interaction induite par le système. Cette description s'appuie non seulement sur les éléments constituant les approches de caractérisation des systèmes mixtes présentées dans les chapitres I et II, mais aussi sur des caractéristiques complémentaires propres aux deux facettes de l'interaction.

Dans ce chapitre, nous présentons la notation ASUR, ses composants, ses relations et les caractéristiques utiles à l'analyse de l'interaction Utilisateur-Système Mixte. Nous décrivons ensuite plusieurs systèmes mixtes avec la notation ASUR afin d'illustrer son utilisation.

## Chapitre IV :

# ASUR : Support à la Caractérisation des Systèmes Mixtes

Notre notation ASUR, dont l'objectif est de décrire un système mixte, met en évidence les différentes entités composant le système. Par rapport à un cycle de vie du logiciel qui comprend deux étapes principales, la conception puis le développement logiciel, la mise en œuvre d'ASUR intervient clairement dans les phases de conception. Par conséquent, aucune attention particulière n'est portée aux composants logiciels qui devront être ultérieurement conçus et réalisés. Par contre, la description d'un système avec ASUR reflète la phase d'utilisation du système, pour laquelle ASUR identifie les composants matériels intervenant dans l'interaction, ainsi que l'échange de données entre les composants matériels et informatiques. La caractérisation de ces échanges de données entre les composants d'un système fournit donc une description du système en cours d'utilisation.

Nous décrivons, dans le paragraphe 1, les composants ASUR puis, dans le paragraphe 2, les relations entre composants. Dans le paragraphe 3 nous caractérisons de manière plus précise les composants et les relations jouant un rôle dans l'interaction Utilisateur-Système Mixte. En effet nous proposons un ensemble de caractéristiques pertinentes pour les composants et relations impliqués dans l'interaction d'un utilisateur avec un système mixte. Enfin nous illustrons ASUR en décrivant plusieurs systèmes mixtes. Nous mettrons en évidence pour chaque système décrit, les apports de leur caractérisation avec la notation ASUR par rapport à leur caractérisation avec les approches existantes, présentées dans les chapitres I et II.

# 1 Composants

Les premiers systèmes interactifs se limitaient à la mise en présence d'un ordinateur et d'un utilisateur. Pourtant les systèmes de GMCAO présentés au chapitre I, ainsi que les systèmes mixtes présentés au chapitre II, illustrent pleinement que cette configuration réduite est désormais obsolète. Un système mixte ne peut être réduit à un couple (ordinateur, utilisateur). Nous avons mis en évidence l'existence de quatre types de composants que nous présentons ci-après.

## 1.1 Utilisateur et Système informatique

Les deux principaux composants de notre espace sont :

- l'Utilisateur (noté **U**, pour "*User*") du système interactif et,
- le Système informatique (noté **S**, pour "*computer System*").

Plus rigoureusement, nous caractérisons le système informatique (composant **S**) comme un ensemble de dispositifs physiques, capable d'acquérir, de stocker, de transformer et de fournir des données numériques (bits). De plus, le modèle sur lequel se basent les transformations est adaptable, ce qui signifie qu'une personne peut modifier la stratégie de transformation des données. Il s'agit donc de l'unité centrale d'un ordinateur et de ses composants matériels et logiciels. Le système informatique est par exemple chargé du traitement des images, du calcul du rendu volumique ou surfacique.

## 1.2 Objets réels

Les systèmes mixtes intègrent dans la réalisation de la tâche une partie du monde réel : cette dernière doit apparaître dans la caractérisation. Cette partie du monde réel apparaît dans notre notation sous la forme du composant **objet Réel** (noté **R**, pour **Real objects**). Cependant, tous les objets du monde réel utilisés dans les systèmes mixtes n'ont pas toujours le même rôle. Par exemple dans le cas du *Musée Augmenté*, présenté dans le chapitre II, l'œuvre d'art vue au travers du casque constitue la cible de la tâche, ce avec quoi l'utilisateur interagit réellement. Au contraire, dans le *Tangible Geospace*, également présenté dans le chapitre II, les briques que l'utilisateur déplace pour modifier l'échelle de projection du plan ne constituent qu'un outil permettant de réaliser une tâche.

Parmi les objets réels, nous distinguons ceux jouant le rôle :

- d'outil (notés **R<sub>tool</sub>** pour "*Real Tool*"),
- de cible de la tâche, ou encore objet de la tâche (notés **R<sub>task</sub>** pour "*Real object of Task*").

## 1.3 Lien entre le monde réel et monde virtuel

Parmi les trois types de composants présentés dans les deux paragraphes précédents, il est intéressant de noter que le système informatique (composant **S**) définit le monde virtuel alors que la personne utilisant le système (composant **U**) et les objets réels (composants **R<sub>task</sub>** et **R<sub>tool</sub>**) appartiennent au monde réel. Parce que c'est dans le couplage de ces deux types de composants que réside l'intérêt d'un système mixte, il convient de considérer un composant complémentaire, permettant de réaliser les transferts de données d'un monde vers l'autre : il s'agit de notre quatrième composant, l'**Adaptateur** (noté **A** pour "*Adaptor*").

Elément original de notre notation, il réalise une passerelle entre le monde réel et le monde informatique, tout comme le font d'ailleurs les dispositifs électromécaniques évoqués dans le domaine de la GMCAO au chapitre I. A ce titre, un adaptateur peut donc être un guide mécanique, un localisateur mécanique (bras de robots) ou optique, mais aussi un dispositif d'interaction plus traditionnel tel qu'un clavier, une souris ou encore un écran. Nous distinguons deux types d'adaptateurs. Pour cela, nous adoptons le point de vue du système informatique et distinguons :

- les adaptateurs d'entrée, notés  $A_{in}$  (pour "*Adaptator for Input*"), qui transmettent des données du monde réel vers le monde informatique,
- les adaptateurs de sortie, notés  $A_{out}$  (Output "*Adaptator for Output*"), qui transmettent des données du monde informatique vers le monde réel.

Notons que certaines solutions techniques permettent de regrouper les deux types d'adaptateurs en un seul dispositif. C'est par exemple le cas de l'I/O-Bulb : ce dispositif regroupe une caméra et un projecteur vidéo. Intrinsèquement ce dispositif se compose pourtant d'un adaptateur d'entrée, la caméra, et d'un adaptateur de sortie, le projecteur.

## 1.4 Règles d'utilisation et illustration

L'identification des adaptateurs (composants  $A_{in}$  et  $A_{out}$ ) et des objets réels (composant  $R_{tool}$  et  $R_{task}$ ) dépend fortement de la tâche à accomplir. En effet selon la tâche en cours, une même entité réelle peut jouer le rôle d'outil ou d'objet de la tâche, voire même dans certains cas d'adaptateur. Par exemple dans la *Visée pédiculaire* [Merloz 97], un foret est utilisé pour réaliser un trou dans une structure osseuse en vue de positionner une vis. Lors de la réalisation du trou, le foret constitue un outil pour la réalisation de la tâche (composant  $R_{tool}$ ). Cependant, pour connaître sa position dans l'espace, la perceuse portant le foret doit être équipée de marqueurs, repérés par une caméra par exemple. La relation entre la position du marqueur et celle du foret nécessite une phase de calibrage, qui consiste à faire décrire un cône à la perceuse munie du foret et du marqueur. A partir des positions du marqueur enregistrées au cours du calibrage, la position du point resté fixe par rapport au marqueur, c'est-à-dire de l'extrémité du foret, peut être calculée répondant ainsi aux exigences de la localisation de l'extrémité du foret lors de la réalisation du trou. Lors du calibrage, le foret joue donc le rôle d'objet de la tâche, et il constitue un composant  $R_{task}$ . Par conséquent le type d'une entité est lié au rôle rempli par cette entité dans le déroulement de la tâche considérée.

Notons également qu'un système interactif classique se limite au couple (homme, machine). Dans une description avec ASUR, ceci se traduit par l'absence de composant de type  $R_{task}$  et  $R_{tool}$ . Par conséquent, dans la description d'un système avec ASUR, les types de composant  $R_{task}$  et  $R_{tool}$  ne sont pas obligatoirement instanciés.

Dans le cas où seul le composant  $R_{task}$  n'est pas représenté, l'objet de la tâche réalisée avec le système est une entité virtuelle que l'on pourrait désigner par  $S_{task}$  à l'image du composant  $R_{task}$ . Il s'agit alors d'une interaction homme-machine qui utilise des entités du monde réel : c'est la virtualité augmentée de notre taxonomie présentée au chapitre III. Outre les capacités de calcul, de stockage et de manipulation de données, le système contient alors également l'objet de la tâche, qui peut par exemple être un fichier de données. Nous avons toutefois choisi de ne pas représenter ce composant dans une description ASUR. En effet, dans le cas où un composant  $S_{task}$  devrait apparaître, l'utilisateur interagit avec un objet du système informatique, en utilisant des outils du

système informatique et éventuellement des outils du monde réel. L'important étant de modéliser la passerelle entre les deux mondes informatique et réel, il ne nous a pas semblé nécessaire de distinguer le système informatique (composant  $S$ ) de l'objet d'une tâche informatique.

Enfin, dans le cas où la description ASUR d'un système ne fait pas intervenir de composant  $R_{\text{tool}}$  mais uniquement un composant  $R_{\text{task}}$ , cela signifie que l'objet de la tâche de l'utilisateur est une entité réelle, et l'interaction avec ce dernier ne s'appuie sur aucune autre entité réelle.

Illustrons ce premier principe d'ASUR en identifiant les composants du *Musée Augmenté*. La tâche de l'utilisateur consiste à observer une œuvre d'art. L'œuvre, entité du monde réel, constitue donc l'objet de la tâche, composant  $R_{\text{task}}$ . Le visiteur est la seule personne utilisant le système et est représenté dans la description ASUR du système par le composant  $U$ . Le visiteur dispose également d'un lecteur de code-barre permettant d'identifier l'œuvre devant laquelle il se trouve. Ce lecteur de code-barre permet de scanner une information du monde réel pour ensuite interroger une base de données. Il constitue donc un adaptateur d'entrée, composant  $A_{\text{in}}$  de notre notation. Cette base de données ainsi que les capacités d'interrogation de la base sont des composantes du système informatique, composant  $S$ . Une fois les informations retrouvées dans la base de données, le système informatique les rend perceptibles à l'utilisateur en les affichant dans un casque semi-transparent. Ce casque est donc un dispositif de sortie, c'est-à-dire un adaptateur de sortie, composant  $A_{\text{out}}$ , du système mixte. Ainsi, ce système mixte ne fait donc pas apparaître de composant  $R_{\text{tool}}$ , c'est-à-dire d'objet réel constituant un outil pour l'exécution de la tâche.

## 2 Relations

Le second principe repose sur l'existence de liens entre les composants ASUR, liens que nous appelons relations ASUR. Nous avons identifié deux types de relation :

- Des relations traduisant des échanges de données ;
- Des relations traduisant des liaisons physiques.

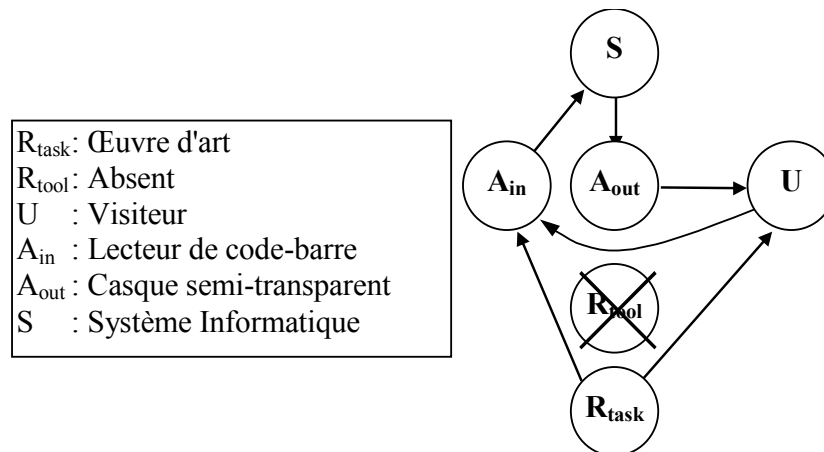
Les relations ASUR représentant un échange de données entre deux composants ASUR sont unidirectionnelles et sont représentées par une flèche, orientée du composant ASUR émetteur vers le composant ASUR récepteur. Ainsi la relation  $A_{\text{out}} \rightarrow U$  représente le fait qu'un ensemble de données pertinentes pour la tâche pourront être perçues par l'utilisateur grâce à un adaptateur de sortie (un écran par exemple). De la même manière, la relation  $A_{\text{in}} \rightarrow S$  caractérise le fait qu'un adaptateur d'entrée, par exemple une souris, transmet des données au système informatique. Dans le cas de la souris, il s'agit par exemple des coordonnées  $(x,y)$  du pointeur.

Le second type de relation ASUR, représentant des liaisons physiques entre composants ASUR, est mis en évidence dans la description avec ASUR d'un système mixte par une simple ligne entre deux composants: il s'agit de relations mécaniques. Ces relations, figurant un échange d'énergie, existent donc entre deux entités du monde réel, c'est-à-dire des composants de type objet réel,  $R_{\text{tool}}$  ou outil  $R_{\text{task}}$ , impliqués dans l'interaction Utilisateur-Système mixte. A la différence des du premier type de relation ASUR, les composants ASUR impliqués dans une relation mécanique ne réalisent

aucune transformation des données échangées. Dans la suite de ces travaux nous n'avons pas approfondi la caractérisation de ces relations.

Notons enfin qu'un adaptateur d'entrée et un adaptateur de sortie peuvent être regroupés à l'intérieur d'un même dispositif. C'est par exemple le cas de l'I/O-Bulb qui combine un projecteur et une caméra. Mais dans ce cas, il n'existe aucun transfert d'informations entre ces deux composants ASUR. Le lien existant entre ces deux composants ne peut donc être considéré comme une relation ASUR, telle que nous l'avons définie précédemment. Nous considérons qu'il s'agit d'une solution technique de réalisation du système qui ne modifie pas l'interaction Utilisateur-Système mixte. Nous reviendrons sur la prise en compte de la fusion de deux adaptateurs au sein d'un même dispositif dans le paragraphe 3.1, lors de la caractérisation des composants ASUR. Nous les représenterons dans les descriptions avec ASUR de systèmes mixtes, par une double ligne.

Poursuivons la caractérisation du *Musée Augmenté* en considérant les relations ASUR présentes. Dans le *Musée Augmenté*, le visiteur manipule un lecteur de code-barre, ce qui se traduit par les relations  $U \rightarrow A_{in}$  et  $A_{in} \rightarrow U$  (perception tactile de la préhension de l'objet). Le lecteur de code-barre, correctement positionné, peut lire le code-barre de l'œuvre d'art,  $R_{task} \rightarrow A_{in}$ , puis transmettre cette donnée au système informatique, d'où la relation  $A_{in} \rightarrow S$ . Après analyse du code-barre et recherche dans la base de données, le système informatique communique au casque semi-transparent les données à afficher :  $S \rightarrow A_{out}$ . Le visiteur peut alors percevoir ces données dans le casque,  $A_{out} \rightarrow U$ , tout en regardant l'œuvre d'art réelle,  $R_{task} \rightarrow U$ . La Figure 34 ci-après reprend la description ASUR du système.



**Figure 34** Description avec ASUR du Musée Augmenté.

En supposant que le visiteur puisse désigner une partie du tableau à l'aide d'une baguette, et que les informations ajoutées soient alors en rapport avec la zone du tableau désignée par le visiteur, la baguette serait un composant de type  $R_{tool}$  et une relation mécanique serait identifiée entre la baguette et l'œuvre d'art ( $R_{tool} - R_{task}$ ).

### 3 Caractéristiques de l'interaction Utilisateur-Système Mixte

Les différents aspects de l'interaction Utilisateur-Système Mixte apparaissent dans ASUR comme l'ensemble des relations connectées à l'utilisateur (composant **U**). Notre objectif étant de centrer notre approche d'analyse sur l'interaction, il est nécessaire d'en préciser les différentes facettes. Les relations liées à l'interaction sont connectées soit à un adaptateur, soit à un objet du monde réel, outil ou objet de la tâche. Aussi, dans les paragraphes suivants nous caractérisons les adaptateurs, les objets du monde réel et les relations connectées à l'utilisateur.

#### 3.1 Caractéristiques des composants

La première caractéristique liée à l'utilisation d'un objet réel ou d'un adaptateur est le **sens perceptif ou actionnel requis** pour percevoir les données du composant ou pour transmettre des données à destination de ce composant. Il existe cinq sens de perception et deux types d'actions, physiques et langagières.

L'être humain perçoit son environnement par l'intermédiaire des cinq sens **visuel, auditif, tactile, olfactif et gustatif**. Parmi ces cinq sens, les sens visuel et auditif sont les plus sollicités dans les interfaces actuelles. Ce constat n'est pas surprenant puisque l'humain perçoit plus de 70% des informations sensorielles par le canal visuel [Ware 96]. Enfin les interfaces impliquant les sens olfactif et gustatif en sont à leurs balbutiements. Il existe par exemple le dispositif *OLFACOM* [Olfacom] qui est un générateur de senteurs piloté par un ordinateur. Ce type de dispositif peut trouver des domaines d'application variés comme les jeux ou le commerce électronique. Le sens tactile est un peu moins exploité, néanmoins nous avons présenté plusieurs systèmes de GMCAO de ce type dans le chapitre I, comme par exemple le système *ACRobot* [Ho 95]. Le sens tactile est à prendre ici au sens large, c'est-à-dire de façon à inclure les retours d'effort. Nous utiliserons donc le terme haptique.

Inversement, l'utilisateur peut agir sur son environnement, et a fortiori sur les entités d'un système mixte, par l'intermédiaire de son système moteur (**action physique**) et par l'intermédiaire du **langage** parlé, écrit ou gestuel. Par exemple, l'utilisateur du système *Charade* [Baudel 93] porte un gant numérique avec lequel il contrôle l'équipement audiovisuel lors d'une présentation devant un auditoire. *Charade* distingue les gestes naturels de ceux qui donnent des directives au système : projeter la diapositive suivante, lancer la vidéo, etc. Enfin, la plupart des localisateurs exploitent les actions physiques de l'utilisateur. En effet, qu'il s'agisse de localisateurs optiques ou de systèmes de reconnaissance basés sur l'utilisation de caméras, les mouvements de la tête du chirurgien ou de ses mains prennent alors une signification pour le système.

L'utilisateur repose donc sur ses cinq sens de perception et ses deux types d'actions pour interagir avec les environnements réel et virtuel. Cependant la juxtaposition de plusieurs sens perceptifs n'est pas toujours concevable. Il convient donc d'identifier ceux requis par l'interaction dans le monde réel et le monde virtuel, pour pouvoir par la suite étudier l'adéquation de leur combinaison.

La seconde caractéristique est le **lieu de perception ou d'action**, c'est-à-dire le lieu géographique où l'utilisateur doit porter son attention pour pouvoir émettre des données à

destination du composant ou en recevoir. Dans le cas d'une interaction avec un adaptateur d'entrée, ce lieu représente l'endroit où l'utilisateur doit utiliser ce dispositif. Notons que dans certains cas, le lieu de perception peut être libre, c'est-à-dire non contraint. L'échange de données avec ce composant ne nécessite aucun lieu particulier. Ce point participe à la transparence de l'interaction. C'est notamment le cas lors de la diffusion non directionnelle de son ou lors de l'affichage de données dans un casque semi-transparent ne dépendant pas de l'orientation du regard de l'utilisateur.

Cette caractéristique est capitale devant l'accroissement du nombre de sources de données potentielles ainsi que du nombre de dispositifs d'interaction. En effet, si dans un système interactif traditionnel, les principaux sites d'interaction sont limités à l'écran et au bureau sur lequel sont posés clavier et souris, un système mixte met en œuvre un ensemble d'objets du monde réel qui sont potentiellement éloignés les uns des autres, mais surtout qui peuvent être distants des adaptateurs de sortie procurant des données virtuelles à l'utilisateur.

Enfin, cette caractéristique est complétée, uniquement pour les adaptateurs d'entrée et de sortie, par une autre caractéristique : leur localisation physique dans l'espace. Cette caractéristique permet d'exprimer le fait que deux adaptateurs sont accolés.

La combinaison du sens perceptif et du lieu de perception forme ce que nous avons nommé **l'environnement perceptif** d'un composant ASUR participant à l'interaction Utilisateur-Système Mixte. De la même façon, la combinaison du sens actionnel et du lieu d'action forme ce que nous avons nommé **l'environnement actionnel** d'un composant.

Les environnements perceptifs et actionnels s'appliquent donc aux adaptateurs d'entrée et de sortie, ainsi qu'aux entités du monde réel utilisées comme outils ou objets de la tâche. En fonction du type d'action que l'utilisateur entreprend avec ces entités, le sens perceptif peut prendre plusieurs valeurs. En effet lors du déplacement d'un objet réel par exemple, des actions physiques sont effectuées et un retour tactile est perçu par l'utilisateur, ce qui induit une perception haptique.

Une troisième caractéristique, moins générique que les deux précédentes puisqu'elle concerne les systèmes mixtes collaboratifs, dépeint l'aptitude d'un composant ASUR à transmettre des données à plusieurs utilisateurs simultanément, ou au contraire à en recevoir simultanément de plusieurs utilisateurs. Cette caractéristique que nous appelons **niveau de partage** du composant concerne aussi bien les adaptateurs d'entrée ou de sortie que les objets réels utilisés dans le système mixte. Cette caractéristique est particulièrement importante pour les dispositifs d'affichage. En effet, devant la multiplication de ces derniers, il nous est apparu que certains ne garantissent pas la même visibilité ; par exemple un casque opaque ou semi-transparent ne permet qu'à une seule personne de percevoir des données. Une projection sur un grand écran autorise par contre un plus grand nombre de participants. L'intérêt de cette caractéristique est donc d'analyser le nombre d'utilisateurs qui peuvent percevoir simultanément une donnée ou agir simultanément sur un même composant. Ce point est particulièrement important si la situation d'interaction doit être collaborative, ou si la situation est critique et qu'elle nécessite, en plus de l'opérateur, un ou plusieurs contrôleurs.



Concluons la présentation des caractéristiques des composants intervenant dans l'interaction en les illustrant dans le cas du *Musée Augmenté*. Les composants concernés sont :

- l'œuvre d'art, composant  $R_{task}$
- le casque semi-transparent, composant  $A_{out}$
- le lecteur de code-barre, composant  $A_{in}$

Pour être perçue, l'œuvre d'art sollicite le sens visuel du visiteur, au niveau de l'œuvre. Comme plusieurs personnes peuvent percevoir cette œuvre simultanément, le niveau de partage des données émises par ce composant est symbolisé par N.

Concernant le casque, le sens visuel de l'utilisateur est également requis. Etant semi-transparent, le lieu de perception des données émises par le casque est totalement libre. Par contre, seul un utilisateur peut percevoir ces données, donc leur niveau de partage est limité à 1.

Enfin, le lecteur de code-barre est manipulé par l'utilisateur. Il nécessite donc des actions physiques de la part de l'utilisateur qui perçoit nécessairement une information tactile au contact du lecteur. Le lecteur de code-barre doit être positionné sur le code-barre situé à proximité de l'œuvre. C'est donc à ce niveau que doit se focaliser le visiteur pour coordonner ses actions physiques. Le lecteur de code-barre sollicite donc un environnement perceptif (haptique, code-barre) et un environnement actionnel (action physique, code-barre).

Le tableau de la Figure 58 ci-dessous résume les valeurs prises par les trois caractéristiques identifiées pour chacun de ces composants.

	Sens/Action	Lieu	Niveau de partage
<b>Œuvre d'art</b>	Visuel	Œuvre d'art	N
<b>Casque</b>	Visuel	Libre	1
<b>Lecteur de code-barre</b>	Actions physiques	Code-barre	1
	Haptiques	Code-barre	1

**Figure 35** Caractérisation des composants participant à l'interaction dans le Musée Augmenté.

### 3.2 Caractéristiques des relations

Le **langage** utilisé par une relation pour véhiculer une donnée constitue la première caractéristique d'une relation formant une partie de l'interaction entre l'utilisateur et un système mixte. Plus précisément, nous définissons le langage utilisé par une relation comme étant la forme syntaxique, au sens de Bernsen [Bernsen 94], de ce qui est véhiculé, c'est-à-dire le vocabulaire et la forme utilisée pour le rendre perceptible. [Nigay 96] établit qu'une modalité est définie comme un couple  $\langle d, l \rangle$  où  $d$  désigne un dispositif physique et  $l$  est un langage d'interaction. Le langage d'interaction se définit par un vocabulaire d'éléments terminaux et une grammaire tandis que les éléments terminaux sont produits ou captés par les dispositifs d'entrée/sortie. Dans ce cadre, le langage d'une relation ASUR est le même que le langage d'une modalité. Le langage ne considère donc pas le média mis en œuvre, au sens défini dans [Nigay 96] où un média représente le dispositif physique d'entrée/sortie.

Le développement de systèmes mixtes met en œuvre plusieurs technologies permettant d'exploiter au mieux la richesse de l'environnement réel de l'utilisateur. Il en résulte une variété des langages mis à la disposition de l'utilisateur pour communiquer avec toutes les entités constituant le système mixte. Il est donc nécessaire de mieux décrire le langage mis en œuvre par les relations ASUR. [Bernsen 94] a identifié quatre caractéristiques d'un langage d'une modalité :

- **statique / dynamique** : cette caractéristique traduit la présence ou non d'une dimension temporelle dans la forme des données véhiculées ;
- **linguistique / non linguistique** : cette caractéristique représente l'existence ou non d'un système structuré de symboles remplissant une fonction de communication ;
- **analogique / non analogique** : cette caractéristique exprime la notion de ressemblance avec la réalité. Elle représente le fait que le langage utilisé s'appuie ou non sur une métaphore du monde ;
- **arbitraire / non arbitraire** : cette caractéristique décrit l'absence ou la présence de liens entre le contenu et le contenant. Un langage non-arbitraire repose sur une fonction de correspondance qui n'a pas été définie par le concepteur mais qui est conventionnelle et connue de l'utilisateur.

L'expression et la sémantique des données échangées reposent sur le langage. L'aptitude d'un utilisateur à maîtriser un langage constitue donc la principale caractéristique du langage. Cependant, cela est fortement dépendant de l'utilisateur, et par conséquent, la caractérisation d'un langage est périlleuse. Néanmoins, les quatre propriétés ci-dessus constituent un point de départ [Bernsen 94]. Nous avons donc retenu ces quatre caractéristiques, auxquelles nous jugeons nécessaire d'ajouter la prise en compte de la **dimension spatiale**. En effet les langages mis en œuvre peuvent être unidimensionnels, bidimensionnels ou tridimensionnels. Avec le développement des mécanismes de spatialisation des sons et de la visualisation stéréoscopique, un langage peut être tridimensionnel. D'autres études, concernant la visualisation d'information, prennent en compte la dimension de la donnée manipulée par le système [Schneiderman 96]. Toutefois, cet aspect de la dimension ne correspond pas à nos objectifs, dans la mesure où une donnée peut être intrinsèquement 3D, tout en étant présentée à l'utilisateur sous forme 2D par exemple. La dimension de la représentation intervient pour distinguer différentes catégories de points de vue, présentées dans [Bruley 99]. Selon cette approche, un point de vue regroupe différents aspects des données manipulées :

- Les attributs des données devant être perçus via la représentation graphique ;
- Les attributs des données pertinents pour l'accomplissement de la tâche ;
- La structure à adopter pour organiser les données.

Or dans notre approche, une relation ASUR représente un transfert d'informations qui ne donne pas nécessairement lieu à une représentation graphique. Aussi, le premier aspect constituant le point de vue selon [Bruley 99] ne correspond pas à notre situation. La dimension spatiale que nous retenons pour caractériser le langage mis en œuvre par une relation n'est donc pas la dimension de la représentation graphique des données. Pour la même raison, il ne s'agit pas non plus de la dimension de la structure adoptée pour organiser les données.

Il en résulte donc que notre dimension spatiale ne prend en compte que le second aspect du point de vue défini précédemment. La dimension spatiale du langage mis en

œuvre par une relation ASUR entre deux composants représente donc le nombre de dimensions de la présentation (graphique, textuelle, haptique, etc.) porteuses d'informations pertinentes pour la réalisation de la tâche. Par exemple, un texte servant de légende à une carte géographique est représenté graphiquement en 2D, ou en 3D si l'on ajoute des ombres aux lettres, mais seule une dimension est pertinente puisque ce texte doit être lu et que la lecture d'un texte se fait de manière linéaire.

Caractériser le langage d'une relation ASUR revient donc à exprimer le quintuplet (dimension, arb. ou non, anal. ou non, ling. ou non, stat. ou dyn.).

La seconde caractéristique que nous avons retenue pour une relation ASUR décrivant l'interaction Utilisateur-Système Mixte est son **poids**. Introduit dans [Graham 00] comme étant l'attention portée sur les données par l'utilisateur au cours de la réalisation de la tâche, le poids d'une relation est exprimé suivant un axe continu. A l'extrémité inférieure se trouvent les données ne recevant aucune attention, par exemple le thermomètre situé sur le bureau de l'utilisateur lorsqu'il est en train de rédiger une lettre. A l'autre extrémité se situent les données ayant une importance capitale lors de la réalisation de la tâche. En chirurgie augmentée, il s'agit par exemple des données de guidage fournies par le système. Le long de cet axe nous n'avons conservé que quatre valeurs pour exprimer le poids d'une relation : aucun, périphérique, faible et élevée. Le poids exprime donc la pertinence et l'importance des données portées par la relation pour la réalisation de la tâche. Pour pondérer une relation, des études ergonomiques semblables à celles menées par [Rau 96] permettent d'établir l'importance des données affichées en comptabilisant la fréquence de consultation de ces données par le clinicien au cours de la réalisation de la tâche.

L'introduction et l'utilisation d'objets réels entraîne nécessairement l'ajout d'échanges de données, ne serait-ce que le fait de palper l'objet manipulé. Or, selon le cas cette donnée peut ne pas être utile à la réalisation de la tâche, tout en étant requise pour garantir à l'utilisateur une analogie à la manipulation du monde réel. Prenons l'exemple du clavier. Son utilisation implique que l'utilisateur perçoive une sensation tactile lors du contact avec les touches. Toutefois, cette sensation n'a a priori pas d'impact sur la réalisation de sa tâche. Par contre, si la pression appliquée sur les touches doit avoir un impact sur le traitement des données, par exemple écrire en gras les caractères sur lesquels l'utilisateur a appuyé plus fortement, cette relation devient importante en regard de la tâche à effectuer.

Définir le poids des relations permet d'identifier les relations les plus importantes pour l'utilisateur au cours de la réalisation de sa tâche. Le poids des relations implique donc un ordre entre les différentes relations composant l'interaction. Les efforts de conception devront alors se porter en priorité sur les relations de poids important.

La dernière caractéristique d'une relation ASUR, très similaire au point de vue retenu dans l'approche de caractérisation du domaine de la synthèse d'images [Milgram 94], est le **référentiel de présentation** adopté pour rendre perceptibles les données. Considérons par exemple un cockpit d'avion représenté graphiquement en 3D. Il est possible de placer une caméra à l'arrière de l'avion et d'afficher à l'écran ce qui est vu par la caméra de sorte que les orientations de l'avion par rapport à l'horizon soient matérialisées à l'écran. Le référentiel de présentation, ou le point de vue, est alors extérieur au cockpit. [Milgram 94]

associe deux valeurs au point de vue. Ces deux valeurs relatent une approche centrée sur l'utilisateur du système : il s'agit des points de vue égocentrique et exocentrique. Le point de vue exocentrique sous-entend que le référentiel de présentation est associé à une des entités physiques composant le système mixte. Nous avons donc choisi d'affiner cette caractéristique en considérant qu'un référentiel de présentation pouvait être associé à chacun des composants physiques de la notation ASUR. Nous identifions par conséquent six types de référentiels de présentation : celui d'un adaptateur de sortie ( $A_{out}$ ) ou d'entrée ( $A_{in}$ ), celui d'un objet réel utilisé comme outil ( $R_{tool}$ ) ou constituant l'objet de la tâche ( $R_{task}$ ), celui associé à un utilisateur ( $U$ ) et enfin un référentiel associé à une donnée virtuelle ( $S$ ). Ainsi, dans le référentiel utilisateur les données sont présentées par le système informatique de la même manière que si elles existaient réellement dans l'environnement de l'utilisateur. Les données se juxtaposent à la réalité perçue par l'utilisateur. Si l'on reprend l'exemple du cockpit d'avion, le référentiel utilisateur présenterait des données identiques à celles que le pilote perçoit : une représentation des différents cadrans de navigation ainsi que les fenêtres du cockpit, au travers desquelles serait représenté le ciel. Considérons maintenant l'aileron de direction situé à l'arrière de l'avion. Il constitue un outil permettant au pilote de diriger l'avion. Visualiser l'avion dans ce référentiel outil se traduit alors par une vision complète de l'appareil depuis l'arrière de l'avion. Le système présente les données sous un angle différent de celui avec lequel l'utilisateur pourrait percevoir ces données. Lorsque le référentiel de présentation utilisé est associé à un adaptateur ASUR ou un objet réel ASUR, la représentation des données fournies dépend alors de la position et de l'orientation de cette entité.

Lors de l'utilisation d'un système mixte, l'utilisateur est amené à observer et interpréter les données fournies par plusieurs entités. Chacune de ces entités adopte un référentiel de présentation des données qu'elle fournit. L'utilisateur doit alors pour chaque entité comprendre le référentiel de présentation choisi, puis les données affichées. Il est donc opportun d'identifier cette caractéristique pour chaque relation afin d'aider le concepteur à choisir des référentiels de présentation cohérents pour chaque entité et réduire les efforts d'interprétation de l'utilisateur. Cette observation est également valable dans le sens inverse de l'interaction, c'est-à-dire lorsque l'utilisateur doit exprimer une donnée à destination d'une entité du système mixte.

Illustrons ces trois caractéristiques avec les relations constituant l'interaction dans le *Musée Augmenté*. Les relations sont :

- la perception de données affichées dans le casque,  $A_{out} \rightarrow U$ ,
- la perception de l'œuvre d'art,  $R_{task} \rightarrow U$ ,
- l'utilisation de lecteur de code-barre,  $U \rightarrow A_{in}$ ,
- la sensation de préhension du lecteur de code-barre,  $A_{in} \rightarrow U$ .

Les données affichées dans le casque ( $A_{out} \rightarrow U$ ) sont relatives à l'auteur et sont textuelles. Le langage associé à ces données est par définition non arbitraire puisque la langue utilisée pour afficher ces données est a priori connue. De plus le langage est non analogique, statique et linguistique. Le référentiel de présentation adopté est le référentiel utilisateur puisque l'utilisateur peut lire les caractères normalement, et le poids de cette relation est élevé puisque ces données sont pertinentes pour la tâche, qui consiste à découvrir une œuvre d'art.

La relation entre l'œuvre d'art et l'utilisateur ( $R_{task} \rightarrow U$ ) est plus délicate à considérer. Supposons qu'il s'agisse d'un musée de peintures, le langage utilisé est bidimensionnel. Il est statique, non linguistique. Le fait de regarder un tableau signifie que l'utilisateur perçoit une donnée réelle. Le langage est donc analogique et non arbitraire. Par contre si l'on cherche à caractériser le contenu de l'œuvre, il est évident que l'identification des caractères arbitraire et analogique posent un problème : ils dépendent entièrement de l'œuvre que l'utilisateur regarde. Enfin, le poids de cette donnée est élevé et le référentiel utilisateur est celui utilisé par cette relation, puisqu'elle représente la perception directe de l'œuvre par l'utilisateur.

L'utilisateur agit également sur le lecteur de code-barre ( $U \rightarrow A_{in}$ ). Il s'agit alors d'une action physique 3D sur cette entité, qui est totalement réelle donc non arbitraire, analogique, non linguistique et dynamique. Le référentiel de présentation est toujours celui associé à l'utilisateur et le poids de cette relation est élevé. En effet la lecture du code-barre, possible grâce à la manipulation du lecteur, est cruciale pour que le système soit en mesure de fournir des données.

En retour, ( $A_{in} \rightarrow U$ ), l'utilisateur perçoit une information tactile exprimée dans un langage tridimensionnel. Cette perception tactile est totalement réelle également. Les caractéristiques complémentaires sont donc les mêmes que pour la relation de l'utilisateur vers le code-barre. Le référentiel de présentation est là aussi celui de l'utilisateur. Par contre le poids de la relation est faible : si l'utilisateur ne perçoit pas de manière tactile le lecteur de code-barre, cela ne gêne en rien la réalisation de la tâche. On peut en effet imaginer une solution dans laquelle le lecteur serait manipulé par l'intermédiaire d'un bras robotisé.

Le tableau de la Figure 36 ci-dessous résume les valeurs prises par les trois caractéristiques identifiées pour chacune des relations constituant l'interaction induite par l'utilisation du *Musée Augmenté*.

	<b>Langage</b>	<b>Poids</b>	<b>Référentiel de présentation</b>
$A_{out} \rightarrow U$	Textuel Non arb., non anal., ling., stat.	Elevé	Utilisateur
$R_{task} \rightarrow U$	2D Non arb., anal., non ling., stat.	Elevé	Utilisateur
$U \rightarrow A_{in}$	Réel 3D Non arb., anal., non ling., stat.	Elevé	Utilisateur
$A_{in} \rightarrow U$	Réel 3D Non arb., anal., non ling., stat.	Faible	Utilisateur

**Figure 36** Caractérisation des relations participant à l'interaction dans le Musée Augmenté.

Nous venons d'exposer la notation ASUR : ses composants, ses relations et les caractéristiques des composants et relations impliqués dans l'interaction Utilisateur-Système mixte. Dans le paragraphe suivant, nous illustrons son application en décrivant plusieurs systèmes présentés dans les chapitres I et II.

## 4 Exemples de descriptions ASUR

### 4.1 Visée Pédiculaire, ACRobot et Craniotomie Assistée par Ordinateur

#### 4.1.1 Composants et Relations ASUR

Nous décrivons avec ASUR le système assistant le chirurgien lors de la réalisation d'une *visée pédiculaire* [Merloz 97], ainsi que le système *ACRobot* [Ho 95], présentés au chapitre I. L'affichage des données de guidage se fait par le biais d'un écran pour la *visée pédiculaire* alors qu'avec *ACRobot* la présentation des données sur l'écran est complétée par un retour d'effort transmis au chirurgien par l'intermédiaire d'un bras articulé et motorisé. Dans ces deux systèmes de GMCAO, la position et l'orientation de l'outil chirurgical, composant  $\mathbf{R}_{\text{tool}}$ , sont suivies en temps réel par un adaptateur d'entrée, composant  $\mathbf{A}_{\text{in}}$ , ( $\mathbf{R}_{\text{tool}} \rightarrow \mathbf{A}_{\text{in}}$ ) et l'outil est en contact avec le patient, composant  $\mathbf{R}_{\text{task}}$ , d'où la relation mécanique  $\mathbf{R}_{\text{tool}} \text{---} \mathbf{R}_{\text{task}}$ . Dans la *visée pédiculaire*, il s'agit d'un localisateur optique et dans *ACRobot*, des encodeurs sont fixés sur le bras articulé. L'adaptateur d'entrée transmet les données relatives à la position des outils au système informatique, composant  $\mathbf{S}$ , ( $\mathbf{A}_{\text{in}} \rightarrow \mathbf{S}$ ).

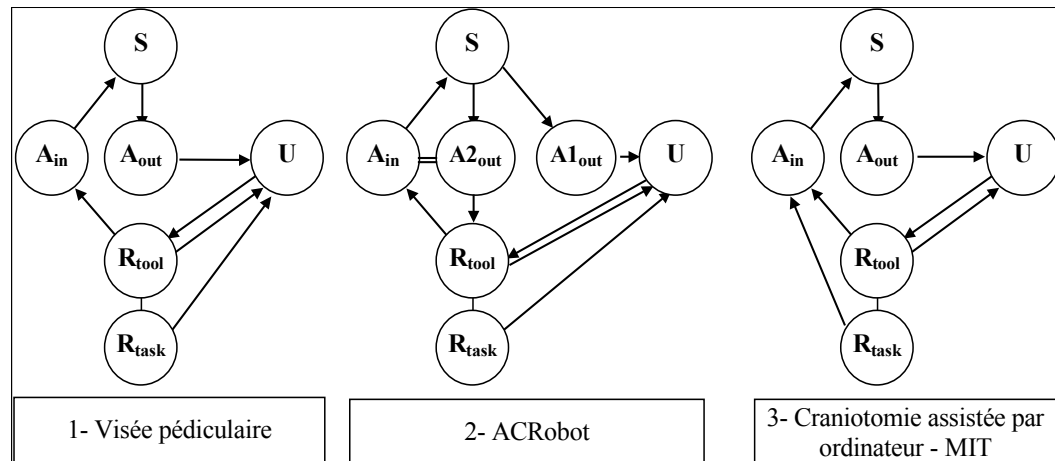
Pour la *visée pédiculaire*, le système calcule les données à rendre perceptibles par le chirurgien via l'adaptateur de sortie ( $\mathbf{S} \rightarrow \mathbf{A}_{\text{out}}$ ). Le retour d'information étant uniquement visuel, l'adaptateur de sortie est un écran et, les données sont donc perçues par le clinicien, composant  $\mathbf{U}$ , via cet écran ( $\mathbf{A}_{\text{out}} \rightarrow \mathbf{U}$ ), comme l'illustre la Figure 37.1.

Pour *ACRobot* le système met en œuvre deux adaptateurs : un écran ( $\mathbf{A1}_{\text{out}}$ ) et le bras robotisé ( $\mathbf{A2}_{\text{out}}$ ), qui constitue également un adaptateur d'entrée ( $\mathbf{A}_{\text{in}}$ ) comme expliqué précédemment. Ce double rôle est traduit par la lien  $\mathbf{A}_{\text{in}} = \mathbf{A2}_{\text{out}}$ . Par conséquent le système informatique produit des données à destination de deux adaptateurs de sortie différents ( $\mathbf{S} \rightarrow \mathbf{A1}_{\text{out}}$  et  $\mathbf{S} \rightarrow \mathbf{A2}_{\text{out}}$ ). En ce qui concerne les données visuelles produites par le système, le clinicien peut les percevoir directement sur l'écran ( $\mathbf{A1}_{\text{out}} \rightarrow \mathbf{U}$ ). Quant au retour d'information produit à destination du bras robotisé, il s'agit d'un retour d'effort. Le bras articulé motorisé agit directement sur l'outil chirurgical. On a donc la relation  $\mathbf{A2}_{\text{out}} \rightarrow \mathbf{R}_{\text{tool}}$ , comme schématisé dans la Figure 37.2. Le clinicien percevra ces données via son outil.

Enfin, dans les deux systèmes, le chirurgien perçoit le patient qui est l'objet réel de la tâche, composant  $\mathbf{R}_{\text{task}}$ , ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{U}$ ). D'autre part le clinicien perçoit l'outil, composant  $\mathbf{R}_{\text{tool}}$ , et le déplace, ce qui justifie la double relation :  $\mathbf{R}_{\text{tool}} \leftrightarrow \mathbf{U}$ .

Si l'on considère maintenant le système du MIT, présenté au chapitre 1, qui assiste le chirurgien lors de la définition du contour d'une craniotomie [Grimson 96], le crayon, composant  $\mathbf{R}_{\text{tool}}$ , permettant de tracer le contour de la craniotomie, et le patient, composant  $\mathbf{R}_{\text{task}}$ , sont tous les deux filmés par une caméra, composant  $\mathbf{A}_{\text{in}}$ , ( $\mathbf{R}_{\text{tool}} \rightarrow \mathbf{A}_{\text{in}}$ ,  $\mathbf{R}_{\text{task}} \rightarrow \mathbf{A}_{\text{in}}$ ). Comme précédemment, il existe une relation mécanique entre le crayon et le patient ( $\mathbf{R}_{\text{tool}} \text{---} \mathbf{R}_{\text{task}}$ ). Le flux vidéo est transmis au système informatique, composant  $\mathbf{S}$  ( $\mathbf{A}_{\text{in}} \rightarrow \mathbf{S}$ ). Le système réalise ensuite la segmentation du visage sur le flux vidéo puis son recalage avec le modèle 3D du visage, stocké en mémoire. Par rapport à ce modèle, la position du modèle 3D de la tumeur est connue. La fusion du modèle 3D de la tumeur et de la vidéo est alors affichée sur un écran, composant  $\mathbf{A}_{\text{out}}$ , ( $\mathbf{S} \rightarrow \mathbf{A}_{\text{out}}$ ) que le clinicien peut

observer ( $A_{out} \rightarrow U$ ). Enfin, le clinicien déplace le crayon et le perçoit d'un point de vue tactile ( $U \leftrightarrow R_{tool}$ ). La Figure 37.3 résume cette description.



**Figure 37** Descriptions avec ASUR de la visée pédiculaire, d'ACRobot, ainsi que du système de craniotomie assistée par ordinateur du MIT.

Approfondissons maintenant la description de ces trois systèmes en caractérisant les composants puis les relations liées à l'interaction Utilisateur-Système Mixte.

#### 4.1.2 Caractéristiques des composants

Parmi les caractéristiques d'un composant ASUR participant à l'interaction Utilisateur-Système Mixte, c'est-à-dire les composants adaptateurs et objets réels, nous étudions l'environnement perceptif requis. En effet les systèmes étant mono-utilisateur, nous n'étudierons pas le niveau de partage, caractéristique propre aux systèmes collaboratifs.

La description ASUR du système de *visée pédiculaire* (cas 1) fait apparaître trois composants participant à l'interaction du clinicien avec le système mixte : l'écran, composant  $A_{out}$ , le foret, composant  $R_{tool}$  et le patient, composant  $R_{task}$ . L'environnement perceptif lié à l'écran est représenté par le couple (vision, écran) puisque le clinicien perçoit les données graphiques de guidage au niveau de l'écran. Le clinicien perçoit le foret visuellement mais aussi de manière haptique, du fait de la résistance et de la préhension de l'outil, au niveau du champ opératoire. Il produit aussi des actions physiques au niveau de l'outil pour le déplacer, c'est-à-dire au niveau du champ opératoire. Enfin le patient est perçu visuellement et au niveau du champ opératoire, seule partie intéressante pour le clinicien. Le système de la *visée pédiculaire* (cas 1) requiert donc deux environnements perceptifs distincts représentés par les couples (visuel, écran) et (visuel/haptique, champ opératoire). Il requiert également un environnement actionnel représenté par le couple (action physique, champ opératoire).

Dans l'exemple 2, *ACRobot*, l'environnement perceptif lié à l'outil requiert également les sens visuel et haptique au niveau du champ opératoire. La composante haptique des données renferme les données de guidage à la différence de l'exemple 1 précédemment présenté. C'est également au niveau de l'outil, donc du champ opératoire, qu'apparaît un environnement actionnel (action physique, champ opératoire). L'environnement perceptif requis par le patient est identique au cas précédent : il est représenté par le couple (vision, champ opératoire). L'environnement perceptif lié à l'écran requiert la vision au niveau de

l'écran. L'exemple 2 nécessite donc deux environnements perceptifs, les mêmes que dans l'exemple 1. Pourtant, les données perçues au niveau de l'écran sont plus simples à interpréter, mais sont redondantes par rapport à celles perçues au niveau du champ opératoire.

Par conséquent toutes les données nécessaires à la réalisation de la tâche ne requièrent qu'un seul environnement perceptif représenté par le couple (visuel/haptique, champ opératoire) et un environnement actionnel représenté par le couple (action physique, champ opératoire). La multiplicité des environnements perceptifs requis par l'exemple 1 est ainsi éludée dans l'exemple 2. Par contre, la description met en évidence la relation d'un robot vers un outil du champ opératoire et donc à proximité du patient. Cet aspect, en milieu chirurgical, doit attirer l'attention du concepteur.

Enfin, dans l'exemple 3, la *Craniotomie*, seuls deux composants ASUR sont impliqués dans l'interaction du clinicien avec le système mixte : l'écran et le crayon, outil de réalisation de la tâche. Les environnements perceptifs requis par chacun de ces composants sont respectivement le couple (vision, écran) et le couple (haptique, champ opératoire). L'environnement actionnel, relatif au déplacement du crayon est représenté par le couple (action physique, champ opératoire). L'interaction repose donc sur deux environnements perceptifs distincts et un environnement actionnel, comme dans l'exemple 1.

Toutefois, à l'opposé de l'exemple 1, le sens perceptif requis par ces deux environnements perceptifs n'est pas le même : en effet la vision n'est pas requise au niveau du champ opératoire. Aussi, dans l'exemple 1, le chirurgien est contraint de regarder le champ opératoire pour percevoir une donnée nécessaire à la réalisation de sa tâche et, implicitement, ceci l'empêche de percevoir les données exprimées dans l'autre environnement perceptif, l'écran. Dans l'exemple 3 par contre, le retour haptique est perçu sans que l'utilisateur ne cesse de percevoir les données visuelles produites dans l'autre environnement perceptif, l'écran. Il peut donc ici percevoir simultanément les données exprimées dans ces deux environnements perceptifs. La présence simultanée de deux environnements perceptifs dans l'exemple 3 est donc moins problématique que dans le premier cas car la perception des données produites dans l'un n'exclut pas la perception simultanée des données produites dans l'autre.

Par contre, il ressort de l'analyse des lieux de perception de l'exemple 3 que le chirurgien n'est jamais amené à observer directement son patient, cas inhabituel en chirurgie et qu'il convient d'identifier clairement lors de la conception.

Terminons maintenant la description des trois systèmes en caractérisant les relations intervenant dans l'interaction Utilisateur-Système Mixte.

#### 4.1.3 Caractéristiques des relations

Parmi ces trois systèmes, seul *ACRobot*, exploitant un retour d'effort (exemple 2), présente toutes les données nécessaires à la réalisation de la tâche dans un référentiel de présentation utilisateur.

Dans la *visée pédiculaire* (exemple 1), les données de guidage sont présentées dans le référentiel de la trajectoire à suivre, alors que le chirurgien observe son patient dans un



référentiel utilisateur. L'exemple 1 présente donc des données en adoptant deux référentiels de présentation différents.

Quant au troisième système, la *craniotomie* assistée par ordinateur, toutes les données sont présentées dans le référentiel de la caméra braquée sur le champ opératoire. Il s'agit donc ici d'un référentiel adaptateur d'entrée.

Enfin l'analyse de la description ASUR de ces trois systèmes nous amène à conclure que les trois combinent des langages basés sur des dimensions différentes. La *visée pédiculaire* propose des données de guidage 2D alors que la visualisation de l'aiguille réelle se base sur un langage 3D. Le retour d'effort d'*ACRobot* combine une information haptique avec une visualisation 3D du patient, et le système du MIT utilisé pour définir le contour d'une *craniotomie* combine un flux vidéo 2D et une information haptique correspondant à la pression du feutre utilisé par le chirurgien pour tracer le contour à la surface du crâne du patient.

#### 4.1.4 Apports de la description ASUR

Dans ce paragraphe nous dressons un bilan de la caractérisation des composants et des relations établies dans les deux paragraphes précédents. Ce bilan illustre les points particuliers mis en évidence grâce à une analyse ASUR sur des systèmes mixtes.

Les systèmes 1 et 3, la *Visée pédiculaire* et la *Craniotomie*, utilisent des technologies identiques pour rendre perceptibles les données : il s'agit dans les deux cas d'un écran. D'autre part, ces deux systèmes appartiennent à la même classe de la taxonomie des systèmes de GMCAO : ce sont des systèmes passifs. L'analyse de la description ASUR de ces deux systèmes identifie dans les deux cas deux environnements perceptifs. Cependant, dans l'exemple 1, la perception des données émises dans chacun des deux environnements perceptifs est mutuellement exclusive alors qu'il est possible de percevoir simultanément les données émises dans les deux environnements perceptifs de l'exemple 3.

La comparaison des systèmes 1 et 2, la *Visée pédiculaire* et *ACRobot*, conduit à l'identification de deux environnements perceptifs également, mais le poids des données produites dans ces deux environnements perceptifs n'est pas le même pour les deux systèmes. Dans le cas d'*ACRobot*, le poids faible de l'une des deux relations traduit le fait que les données portées par celle-ci sont redondantes par rapport à celles portées par la seconde relation. Etant redondantes, la perception simultanée de ces deux données n'est pas nécessaire. Ainsi, la combinaison des deux environnements perceptifs est rendue acceptable dans le cas d'*ACRobot*.

L'analyse des environnements perceptifs basée sur ASUR permet donc de tous les identifier et les étudier, mais surtout d'envisager les conditions dans lesquelles l'utilisateur pourra combiner les données fournies par chacun. Ces conditions regroupent aussi bien la simultanéité de la perception de toutes les données, que le contexte global d'interaction, comme par exemple la perception indirecte du patient, donc de l'objet de la tâche, dans l'exemple 3.

Les systèmes 1 et 2, la *Visée pédiculaire* et *ACRobot*, permettent la réalisation d'une tâche similaire : la combinaison d'informations de stratégie et d'informations anatomiques

aide le clinicien à réaliser une intervention. Toutefois, l'analyse des langages utilisés par les relations ASUR constituant l'interaction du système identifie deux langages différents : il s'agit dans l'exemple 1 d'un langage graphique 2D alors que l'exemple 2 se base sur des actions physiques. Cette différence met alors clairement en évidence l'action d'un robot à proximité du champ opératoire, donc du patient. Cet aspect d'analyse n'est pas apparent dans les approches de la RA présentées au chapitre II. De plus, par rapport à la taxonomie de GMCAO, il ressort que les systèmes actifs sont basés sur la présence d'un robot à proximité du patient. Mais *ACRobot* est un système semi-actif et il existe des systèmes semi-actifs qui ne sont pas basés sur la présence d'un bras robotisé à proximité du patient, comme par exemple les *Templates* de [Radermacher 97]. Ainsi, la présence d'un robot agissant directement sur le patient n'est pas plus clairement mise en évidence par la taxonomie des systèmes de GMCAO que par celles de RA. ASUR au contraire le matérialise très concrètement.

Enfin, l'analyse des représentations ASUR des systèmes met nettement en évidence l'hétérogénéité des langages utilisés par les différentes facettes de l'interaction. Les trois exemples ci-dessus se basent chacun sur une combinaison différente de plusieurs langages : 2D et 3D pour l'exemple 1, haptique et 3D pour l'exemple 2, et haptique et 2D pour l'exemple 3. Cette richesse de combinaison des langages employés est donc identifiable avec ASUR.

Il en est de même des différents référentiels de présentation adoptés pour présenter les données. Nos trois exemples illustrent trois combinaisons de référentiel de présentation : combinaison de référentiels utilisateur et entité virtuelle (exemple 1), référentiel utilisateur seul (exemple 2) et référentiel adaptateur uniquement (exemple 3). L'analyse des descriptions ASUR des systèmes permet d'identifier les combinaisons de référentiels et donc de pouvoir comparer des systèmes par rapport aux référentiels de présentation choisis.

## 4.2 KARMA et le Musée Augmenté

Poursuivons l'illustration d'ASUR avec deux systèmes présentés au chapitre II : *KARMA* [Feiner 93] et le *Musée Augmenté* [Rekimoto 95]. Pour cela nous reprenons la description des composants et relations ASUR de ces deux systèmes, puis caractérisons les composants et les relations de ces deux descriptions.

### 4.2.1 Composants et relations ASUR

Dans *KARMA*, l'objectif est de réparer une imprimante qui constitue donc l'objet de la tâche, composant  $\mathbf{R}_{\text{task}}$ . Cette imprimante est filmée par une caméra, composant  $\mathbf{A}_{\text{in}}$  ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{A}_{\text{in}}$ ). Le flux vidéo est transmis au système informatique, qui identifie la position de marqueurs sur l'imprimante et en déduit la position des données 3D à afficher en fonction de la réparation en cours ( $\mathbf{A}_{\text{in}} \rightarrow \mathbf{S}$ ). Ces données sont affichées dans un casque semi-transparent, composant  $\mathbf{A}_{\text{out}}$ , rigidement fixé à la caméra filmant l'imprimante ( $\mathbf{S} \rightarrow \mathbf{A}_{\text{out}}$ ,  $\mathbf{A}_{\text{out}} = \mathbf{A}_{\text{in}}$  -contrainte de conception) et porté par l'utilisateur. L'utilisateur perçoit donc dans le casque les données de guidage ( $\mathbf{A}_{\text{out}} \rightarrow \mathbf{U}$ ) et au travers du casque il perçoit l'imprimante sur laquelle il agit ( $\mathbf{R}_{\text{task}} \leftrightarrow \mathbf{U}$ ). Cette description est présentée à la Figure 38.1. La Figure 38.2 permet d'établir une comparaison entre la description ASUR de *KARMA* et celle du *Musée Augmenté*, présentée dans le paragraphe 2.

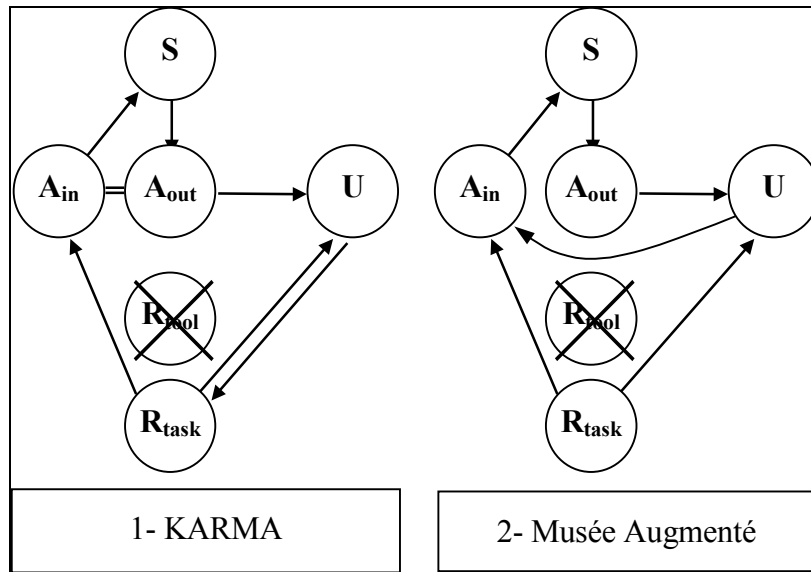


Figure 38 Descriptions avec ASUR de KARMA et du Musée Augmenté.

#### 4.2.2 Caractéristiques des composants

Comme dans le paragraphe 4.1.2, nous ne caractérisons ici que les environnements perceptifs et actionnels, puisque ces deux systèmes sont mono-utilisateur.

Dans *KARMA*, l'imprimante est observée par le technicien et des éléments de l'imprimante sont manipulés. Ceci implique des actions physiques ainsi que la perception tactile de ces éléments au niveau de l'imprimante. Un premier environnement perceptif requis est donc représenté par le couple (vision/haptique, imprimante) et un environnement actionnel est représenté par le couple (action physique, imprimante). Le casque utilisé dans *KARMA* ajoute au niveau de l'imprimante les données graphiques 3D nécessaires au dépannage de l'imprimante. Par conséquent, l'environnement perceptif requis par le casque est représenté par le couple (vision, imprimante).

Les caractéristiques des composants du *Musée Augmenté* sont complètement décrites dans la Figure 35. Les environnements perceptifs requis sont d'une part le couple (visuel, œuvre d'art) et d'autre part (haptique, code-barre). L'environnement actionnel requis est représenté par le couple (action physique, code-barre).

#### 4.2.3 Caractéristiques des relations

Dans *KARMA*, la relation fournissant à l'utilisateur les données produites par le système utilise un langage graphique 3D similaire aux données que l'utilisateur perçoit visuellement de l'objet de la tâche. Le langage nécessairement requis pour les actions physiques et la perception tactile des éléments de l'imprimante que l'utilisateur déplace est le retour d'effort que nous identifions à un langage 3D. Dans tous les cas, le référentiel de présentation est celui de l'utilisateur.

Les caractéristiques des relations du *Musée Augmenté* sont complètement décrites dans la Figure 36. Cette caractérisation met en évidence l'existence de langages différents du point de vue de la dimension : 1D pour les données fournies dans le casque, 2D ou 3D selon l'œuvre observée, et un langage 3D pour le retour tactile ainsi que les actions physiques. Pour toutes ces données, le référentiel de présentation est celui de l'utilisateur. Enfin, le poids de chaque relation est élevé, puisque toutes concourent à l'accomplissement de la tâche.

Dans le *Musée Augmenté*, une relation de l'utilisateur vers l'adaptateur d'entrée apparaît alors qu'elle n'existe pas dans *KARMA*. Le poids de cette relation est élevé, puisque sans intervention de l'utilisateur il est impossible pour le système de savoir en face de quelle œuvre il se trouve. Malgré des technologies d'affichage identiques, un casque semi-transparent, la description avec ASUR met donc en évidence un besoin supplémentaire en terme d'interaction dans le cas du *Musée Augmenté* pour la réalisation d'une tâche similaire : accéder au niveau d'un objet réel, à des données virtuelles, complémentaires et relatives à cet objet. Si l'on considère que dans *KARMA* l'objet réel est augmenté par les marqueurs permettant au système d'aligner les données, dans le *Musée Augmenté* l'utilisation du lecteur de code-barre traduit "l'augmentation" de l'utilisateur. Il semble donc que la distinction que permet d'établir ASUR est en accord avec l'identification des entités augmentées dans un système mixte [Mackay 96].

L'analyse du langage utilisé montre également une différence entre ces deux systèmes, tous deux classés par la taxonomie des systèmes de GMCAO parmi les systèmes passifs. Dans chacun des cas, les données qui sont utiles à la réalisation de la tâche sont présentées sur un même support, un casque semi-transparent, mais selon des langages différents. Dans *KARMA*, il s'agit de données graphiques 3D, alors que dans le *Musée Augmenté* il s'agit de données textuelles. Il est intéressant de constater que notre approche est en accord avec celle basée sur l'analyse des types de données [Azuma 97], [Feiner 93], [Webster 96].

#### 4.2.4 Apports de la description ASUR

Nous identifions dans ce paragraphe les apports de la description ASUR des deux systèmes *KARMA* et *Musée Augmenté*. L'analyse des relations de la description ASUR jette un pont avec les approches de la littérature et ASUR. En effet, l'approche centrée sur les entités augmentées [Mackay 96] ainsi que celles basées sur les différents types de données [Azuma 97], [Feiner 93], [Webster 96] établissent chacune des distinctions que notre notation permet aussi de relater (cf. paragraphe 4.2.3). Notre approche permet donc de fédérer les différents aspects pris en compte par les approches existantes.

De plus, ASUR favorise l'identification de caractéristiques du langage non mis en évidence par les approches basées sur l'analyse du type de données. En ajoutant des caractéristiques et en favorisant leur identification, notre approche complète donc les approches initiales [Azuma 97], [Feiner 93], [Webster 96].

### 4.3 Tangible Geospace et Digital Desk

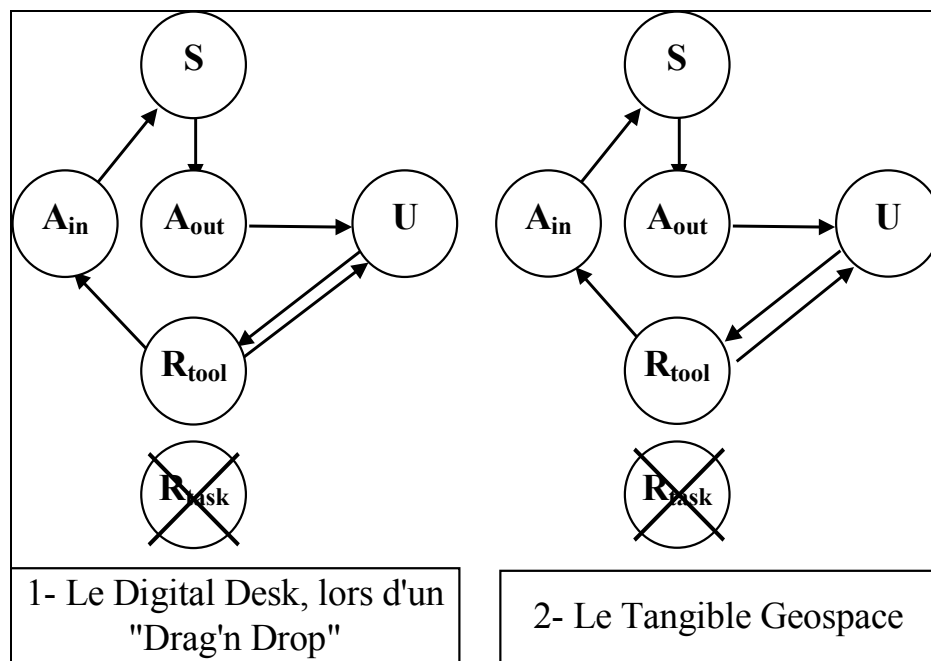
Terminons ces illustrations d'ASUR avec deux systèmes totalement différents des précédents : le *Digital Desk* [Wellner 93] et le *Tangible Geospace* [Ishii 97], présentés au chapitre II.

#### 4.3.1 Composants et relations ASUR

Considérons tout d'abord la tâche de glisser/déplacer ("*drag'n drop*") du *Digital Desk*. L'utilisateur, composant  $U$ , voit et agit sur une feuille de papier réelle dans le but de créer un fichier informatique. Le fichier informatique est donc l'objet de la tâche et il n'y a pas de composant  $R_{task}$  dans la description ASUR de ce système. La feuille papier constitue un objet réel utilisé comme support à la saisie de données dans le fichier informatique : c'est donc un composant  $R_{tool}$  perçu et modifié par l'utilisateur ( $U \leftrightarrow R_{tool}$ ). Cette feuille

est filmée par une caméra, composant  $A_{in}$ , ( $R_{tool} \rightarrow A_{in}$ ) qui transmet le flux vidéo au système informatique ( $A_{in} \rightarrow S$ ). Lors de la tâche de glisser/déplacer, le système calcule la version électronique du texte en cours de déplacement et le rend perceptible à l'utilisateur par l'intermédiaire d'un projecteur, composant  $A_{out}$ , ( $S \rightarrow A_{out}$ ,  $A_{out} \rightarrow U$ ). La Figure 39.1 présente cette description.

Quant à une tâche de manipulation d'un plan avec les *Tangible Geospace*, l'utilisateur, composant  $U$ , observe et manipule des briques physiques représentant des bâtiments du campus du MIT, composant  $R_{tool}$  ( $U \leftrightarrow R_{tool}$ ), dont la position est détectée par des capteurs, composant  $A_{in}$ , ( $R_{tool} \rightarrow A_{in}$ ). Une fois la position détectée, elle est communiquée au système informatique, ( $A_{in} \rightarrow S$ ), qui en déduit la taille et l'orientation du plan à projeter ( $S \rightarrow A_{out}$ ). L'utilisateur observe ces modifications ( $A_{out} \rightarrow U$ ) et poursuit la réalisation de sa tâche.



**Figure 39** Description avec ASUR de la tâche de "Drag'n Drop" avec le Digital Desk, et de la manipulation d'un plan avec le Tangible Geospace.

### 4.3.2 Caractéristiques des composants

Les données projetées par le *Digital Desk* requièrent le sens visuel de l'utilisateur pour être perçues au niveau de la feuille de papier posée sur le bureau. L'utilisateur peut en outre déplacer, en mettant en œuvre des actions physiques, et percevoir visuellement la feuille de papier. Le déplacement de la feuille de papier est nécessaire si l'utilisateur souhaite utiliser momentanément un autre document et donc le positionner à la place de la feuille de papier initialement utilisée. Ce déplacement entraîne nécessairement la perception par l'utilisateur d'informations tactiles. Les environnements perceptifs requis sont donc représentés par les couples (visuel, feuille de papier) en ce qui concerne les données fournies par l'adaptateur et (visuel/haptique, feuille de papier) qui représentent la perception visuelle ou tactile des données réelles. Enfin, un environnement actionnel est défini par le couple (action physique, feuille de papier).

Pour le *Tangible Geospace*, l'utilisateur peut déplacer par des actions physiques les briques représentant les bâtiments du campus et implicitement percevoir un retour tactile

lors du déplacement de ces briques. Tout ceci se passe au niveau du bureau sur lequel est projeté le plan du campus. Ces données fournies par le système via le projecteur requièrent le sens visuel. Ainsi, les deux environnements perceptifs requis sont représentés par les couples (visuel/haptique, bureau) concernant la perception des briques réelles et le couple (visuel, bureau) concernant la perception des données fournies par l'adaptateur. Enfin, comme précédemment, un environnement actionnel est représenté par le couple (action physique, bureau) puisque les briques déplacées se situent sur le bureau.

### 4.3.3 Caractéristiques des relations

Dans les deux systèmes, le poids de toutes les relations est élevé sauf en ce qui concerne la perception tactile liée au déplacement de la feuille ou des briques : pour ces relations le poids peut être qualifié de faible. Le référentiel de présentation utilisé par chacune des relations des deux systèmes est un référentiel utilisateur. Enfin les données projetées par les adaptateurs des deux systèmes sont basées sur un langage 2D. Les langages des relations représentant la perception directe d'entités réelles par l'utilisateur sont intrinsèquement 3D ainsi que les langages utilisés pour traduire le sens tactile et les actions physiques.

### 4.3.4 Apports de la description ASUR

L'analyse des composants, des relations et de leurs caractéristiques respectives ne procure aucune information majeure permettant de distinguer ces systèmes.

Une analyse plus générale montre que les deux systèmes offrent à l'utilisateur une façon nouvelle de réaliser sa tâche : manipuler des objets du monde réel pour interagir avec l'ordinateur. Sans le *Digital Desk*, l'utilisateur aurait dû reproduire à la main les mêmes tracés sur une autre partie de la feuille (pas de "*Drag'n Drop*" possible sans le système), et sans le *Tangible Geospace*, l'utilisateur aurait dû utiliser une souris ou un clavier pour modifier les paramètres de rendu du plan.

Une différence majeure entre ces deux systèmes et ceux précédemment présentés est que dans aucun des deux cas, il n'existe un composant  $R_{task}$ . Ceci signifie, comme expliqué dans le paragraphe 1.4, que l'objet de la tâche n'est pas une entité réelle et qu'il est inclus dans le système informatique (composant  $S$ ) : les systèmes sont donc deux cas de Virtualité Augmentée selon notre taxonomie présentée au chapitre III.

La description ASUR de ces deux derniers systèmes mixtes souligne que notre notation ASUR permet de clairement distinguer les situations dans lesquelles l'objet de la tâche est réel de celles dans lesquelles l'objet de la tâche est virtuel. Ainsi, les deux grands types de systèmes mixtes que nous avons identifiés au chapitre III, la Réalité Augmentée (RA) et la Virtualité Augmentée (VA) font intervenir des caractéristiques clairement identifiables dans notre notation de description. Néanmoins, il est intéressant de noter que le deuxième axe de notre classification, le type d'augmentation, ne peut être traduit dans une description ASUR. En effet les descriptions et analyses de ces deux systèmes ne présentent aucune différence majeure alors que nous avons expliqué au chapitre III que dans le *Tangible Geospace* la phase d'évaluation de la réalisation de la tâche n'est pas augmentée tandis qu'elle est augmentée dans le *Digital Desk*.

## 5 Intérêts et limitations

Nous concluons ce chapitre en dressant un bilan analytique de notre notation ASUR. Nous identifions d'abord deux limitations à ASUR puis nous soulignons ses intérêts.

La limitation de notre notation de description est double :

- le type d'augmentation n'est pas caractérisé.
- l'analyse est limitée à un utilisateur et son interaction.

La première limitation est illustrée par les exemples cités dans le paragraphe 4.3 : le *Tangible Geospace* et le *Digital Desk*. En effet ils illustrent un cas où la description ASUR est la même pour deux systèmes, alors que le type d'augmentation (évaluation ou exécution) diffère. La description d'un système mixte avec ASUR ne suffit donc pas toujours pour mettre en évidence les différents types d'augmentation. Néanmoins cette caractéristique du système est prise en compte par notre cadre de classification. Aussi, c'est en nous appuyant sur notre espace taxinomique que la description ASUR d'un système mixte pourra être complétée.

L'autre limitation d'ASUR est révélée par des systèmes multi-utilisateurs. En effet, dans ces systèmes, les utilisateurs sont tous intégrés à un seul et même composant : le composant **U**. Pourtant l'interaction entre ces utilisateurs ainsi que l'utilisation des adaptateurs et des entités réelles nécessite une réflexion particulière lors de la conception de systèmes mixtes multi-utilisateurs.

D'autre part, ASUR ne permet pas de décrire de manière aussi approfondie la totalité des entités intervenant dans le système. Souhaitant orienter notre analyse sur l'utilisateur, nous nous sommes particulièrement attachés à décrire les entités au contact de l'utilisateur. Il en résulte premièrement qu'un composant ASUR désigne expressément l'existence d'un objet de la tâche s'il fait partie du monde réel. Par contre, si l'objet de la tâche est une entité informatique, aucun composant ASUR n'apparaît dans la description du système pour représenter l'objet de la tâche. L'objet de la tâche, virtuel, est inclus dans le composant ASUR représentant le système informatique (composant **S**). Pourtant de tels systèmes peuvent néanmoins être des systèmes mixtes au sens où nous l'avons défini au chapitre II.

La deuxième conséquence d'une approche centrée sur l'utilisateur et son interaction réside dans le manque de caractérisation des composants ASUR, qui ne sont pas en contact direct avec l'utilisateur, en particulier le système informatique. Sa description peut amener entre autre l'identification de trois propriétés :

- **les fonctionnalités que le système offre à l'utilisateur.** Le spectre des fonctionnalités est d'autant plus large que les données rendues sont modélisées par le système. Ainsi, afficher une photographie du monde réel ne présume de la part du système aucune connaissance quant au contenu de l'image. Par contre afficher un modèle 3D graphique nécessite que le système connaisse les coordonnées des polygones le constituant. Par conséquent, agrandir un objet de la photographie indépendamment des autres objets présents dans la photographie n'est pas possible. Par contre agrandir un objet d'une scène 3D calculée par le système est indépendant de

l'agrandissement des autres objets composant cette même scène. Ces différences en termes de fonctionnalités sont à mettre en relation étroite avec la caractéristique "*Extent of World Knowledge*" de [Milgram 94].

- **le temps de réponse.** Le temps de réponse du système a un impact direct sur l'interaction [Bérard 99].
- **l'aspect réparti du système informatique.** Selon les fonctionnalités offertes, mais aussi les capacités de stockage et de traitement du système informatique et le contexte d'utilisation du système, celui-ci peut être réparti sur plusieurs plates-formes distantes. La conception du système nécessitera alors de prendre en compte les aspects de communication.

Ces trois points ne font qu'illustrer les besoins potentiels en termes de caractérisation du système informatique au sein d'un système mixte. Nous ne visons pas l'exhaustivité mais au contraire nous voulons illustrer l'apport que constituerait la caractérisation plus approfondie du système informatique.

Enfin, nous avons mentionné que la description d'un système avec ASUR ne le décrit que pour une tâche bien définie et offerte par le système. Si dans le cas des systèmes chirurgicaux, la définition de cette tâche est intimement liée au protocole chirurgical lui-même et donc à la méthodologie en trois étapes présentée au chapitre I, il est plus délicat de choisir le niveau de granularité idéal pour un système mixte quelconque. Néanmoins, analyser un système en considérant une tâche permet de centrer l'étude sur l'interaction de l'utilisateur avec un système interactif. Ainsi, une description des composants et relations ASUR d'un système, et l'identification des caractéristiques dépeignent les différentes facettes de l'interaction. Ce dernier point est particulièrement important car l'interaction avec un système mixte est très largement différente de celle traditionnellement offerte par un système interactif classique.

Un autre intérêt de notre approche repose sur son apport en terme de classification des systèmes. En effet, notre notation ASUR, par le biais de la caractérisation des entités du monde réel utilisées, permet de rapidement identifier à quelle classe de notre espace de classification le système appartient. Dans le cas où le concepteur identifie un composant  $R_{task}$  (objet de la tâche), le système appartient à la catégorie Réalité Augmentée. Il peut alors s'inspirer des systèmes déjà identifiés comme tels et répertoriés dans notre espace taxinomique, pour poursuivre la conception de son système. Dans le cas inverse, l'absence de composant  $R_{task}$  signifie que le système fait partie des systèmes de Virtualité Augmentée. Notre notation de description fait donc ressortir un des axes de caractérisation de notre espace de classification.

Nous avons montré que notre notation s'appuie sur les caractéristiques déjà mises en évidence par d'autres approches. Mais ASUR n'est pas uniquement un cadre fédérateur regroupant les approches précédentes. En effet, ASUR met en évidence plusieurs caractéristiques permettant de différencier des situations d'interactions proches. Ces différenciations portent principalement sur :

- le mode de présentation des données utilisé par les différentes facettes de l'interaction entre l'utilisateur et le système mixte. Ceci regroupe le langage



utilisé et le référentiel de présentation adopté pour la diffusion des données par le système mixte.

- les conditions d'utilisation des données émises dans différents environnements perceptifs, leur aspect mutuellement exclusif, redondant, équivalent ou la possibilité de leur exploitation simultanée.
- le contexte de l'interaction induit par le système.

Ces trois points viennent donc compléter les approches existantes de caractérisation des systèmes mixtes. L'intérêt clairement affiché est maintenant de pouvoir établir empiriquement un ensemble de principes garantissant l'utilisabilité des systèmes mixtes. ASUR constitue, au sein de cette démarche empirique, un outil permettant d'identifier des caractéristiques précises concernant les systèmes mixtes pour raisonner ensuite sur leur utilisabilité.

Dans cet objectif, nous présentons au chapitre suivant les points clé des systèmes mixtes et leur pré-requis en terme d'ergonomie. Sur ces bases et celles relatives à la description ASUR des systèmes, nous présentons les premiers éléments de conception pour systèmes mixtes.

## 6 Références

[Azuma 97]

Azuma, R. T., "*A survey of Augmented Reality*", dans la revue *Presence: Teleoperators and Virtual Environments* 6, 4, (1997), p.355-385

[Baudel 93]

Baudel, T. and Beaudouin-Lafon, M., "*CHARADE: Remote Control of Objects using Free-Hand Gestures*", *Communications de l'ACM*. 36(7): (1993), p. 28-35.

[Bérard 99]

Bérard, F., "*Vision par ordinateur pour l'interaction homme-machine fortement couplée*", Thèse de Doctorat de l'Université Joseph Fourier, Grenoble - France, (1999), 201 pages.

[Bernsen 94]

Bernsen, O., "*Foundations of multimodal representations. A taxonomy of representational modalities*", dans la revue *Interacting with Computers*, Vol. 6, 4, (1994), p.347-371.

[Bruley 99]

Bruley, C., "*Analyse des Représentations Graphiques de l'Information - Extension aux représentations tridimensionnelles*", Thèse de l'Université Joseph Fourier, Grenoble, (1999).

[Feiner 93]

Feiner, S., MacIntyre, B., Seligmann, D., "*Knowledge-Based Augmented Reality*", *Communication de l'ACM n°7*, (1993), p. 53-61.

[Graham 00]

Graham, T. C. N., Watts, L., Calvary, G., Coutaz, J., Dubois, E., Nigay, L., "*A Dimension Space for the Design of Interactive Systems within their Physical Environments*", Acte de la conférence DIS'2000, (2000), p. 406-416.

[Grimson 96]

Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells, W.M., Kikinis R., "*An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization*", *IEEE Transactions on Medical Imaging*, 15(2), (1996), p. 129-140.

[Ho 95]

Ho, S., C., Hibberd, R., D., Davies, B., L., "*Robot Assisted Knee Surgery*", dans *IEEE Engineering in Medicine and Biology Magazine*, Special Issue on Robotics in Surgery, (1995), p. 292-300.

[Ishii 97]

Ishii, H., Ullmer, B., "*Tangible Bits : Towards Seamless Interfaces between People, Bits and Atoms*", Acte de la conférence CHI'97, Atlanta, (1997), p. 234-241.

[Mackay 96]

Mackay, W.E., "*Réalité Augmentée : le Meilleur des Deux Mondes*", *La Recherche* n°285, (1996), p. 80-84.

[Merloz 97]

Merloz, P., Tonnetti, J., Eid, A., Faure, C., Lavalley, S., Troccaz, J., Sautot, P., Hammadeh, A., Cinquin, P., "*Computer Assisted Spine Surgery*", dans la revue *Clinical Orthopaedics and Related Research*, n°337, (1997), p. 86-96.

[Milgram 94]

Milgram, P., Kishino, F., "*A Taxonomy of Mixed Reality Visual Displays*", dans la revue *IEICE Transactions on Information Systems*, E77-D(12), (1994), p.1321-1329.

[Nigay 96]

Nigay L. et Coutaz J., "*Espaces conceptuels pour l'interaction multimédia et multimodale*", dans la revue TSI, spécial Multimédia et Collecticiel, AFCET &Hermès Publ. 15(9), (1996), pp. 1195-1225.

[Olfacom]

OLFACOM : Système de perception olfactif . Développé par la société A.C.2.I., URL : <http://www.olfacom.com/anglais/pages/presenta.html>.

[Radermacher 97]

Radermacher, K., Portheine, F., Zimolong, A., Eichhorn, C., Staudte, H., W., Rau, G., "*Image Guided Orthopedic Surgery Using Individual Template*", Acte de la conférence CVRMEd-MRCAS'97, Grenoble, France, (1997), p. 606-615.

[Rekimoto 95]

Rekimoto, J., Nagao, K., "*The World through the Computer : Computer Augmented Interaction with Real World Environments*", Acte du symposium UIST'95, (1995).

[Schneiderman 96]

Schneiderman, B., "*The Eyes Have It: A Task by Data Taxonomy for Information Visualizations*", Acte de la conférence Visual Languages'96, Boulder, USA, (1996), p. 336-343.

[Ware 96]

Ware, C., "*Multimedia output devices and techniques*", dans la revue ACM Computing Surveys (28, 1), (1996), pp. 181-183.

[Webster 96]

Webster, A., Feiner, S., MacIntyre B., Massie, W., Krueger, T., "*Augmented Reality in Architectural Construction, Inspection, and Renovation*", présenté au congrès ASCE Computers in Civil Engineering, (1996), p. 913-919.

[Wellner 93]

Wellner, P., "*Interacting with paper on DigitalDesk*", CACM, (36), 7, (1993).

# CHAPITRE V

## ELEMENTS DE CONCEPTION ET DE REALISATION DE SYSTEMES MIXTES

# Sommaire

## *Chapitre V : Eléments de Conception et de Réalisation de Systèmes Mixtes 129*

<b>1</b>	<b>Etapes de conception et de réalisation d'un système interactif</b>	<b>129</b>
<b>2</b>	<b>Propriétés ergonomiques de la dualité de l'interaction dans un système mixte</b>	<b>132</b>
<b>3</b>	<b>Démarche d'analyse prédictive de la compatibilité et de la continuité</b>	<b>137</b>
<b>4</b>	<b>Illustration de la démarche</b>	<b>139</b>
4.1	Analyse de KARMA	140
4.1.1	Mise en œuvre de la démarche	140
4.1.2	Conclusion et discussion	141
4.2	Analyse de la Visée Pédiculaire	142
4.2.1	Mise en œuvre de la démarche	142
4.2.2	Conclusion et discussion	143
<b>5</b>	<b>Éléments de réalisation logicielle</b>	<b>145</b>
5.1	Modèle d'architecture de référence	145
5.2	Adaptation du modèle aux systèmes mixtes	146
5.3	Flux d'informations dans le modèle adapté	149
<b>6</b>	<b>Conclusion</b>	<b>150</b>
<b>7</b>	<b>Références</b>	<b>152</b>

Nous avons présenté au chapitre précédent ASUR, une notation de description des entités constituant un système mixte ainsi que de l'interaction de l'utilisateur avec le système mixte. Cette notation permet d'identifier des situations d'interaction différentes. Pour chacune de ces situations, l'établissement d'éléments utiles à la conception passe par l'intégration des recommandations de l'ergonomie. A ces fins, nous définissons dans ce chapitre des propriétés ergonomiques liées aux spécificités des systèmes mixtes. Puis nous traduisons ces propriétés en terme d'ASUR. Ainsi la description et la caractérisation d'un système mixte selon ASUR sont complétées par une analyse ergonomique de l'interaction décrite. Description, caractérisation et analyse ergonomique d'un système mixte sont des étapes de conception ergonomique.

Il convient ensuite de traiter le problème de la conception logicielle et d'identifier des solutions qui soient en accord avec la description ASUR (composants, relations et leur caractéristiques) et les propriétés ergonomiques décrites lors de la conception. La conception logicielle doit reposer si possible, sur le cadre structurant d'un modèle d'architecture. La solution avancée dans ce chapitre est une adaptation d'un modèle d'architecture, PAC-Amodeus, pour la réalisation de systèmes mixtes. La mise en place du modèle d'architecture logicielle a conduit à établir une correspondance entre les composants logiciels du modèle et les entités et relations ASUR. Nous définissons ainsi un cadre conceptuel générique à la réalisation logicielle.

Ce chapitre est organisé ainsi : nous motivons d'abord nos résultats de conception et de réalisation en les situant dans un cycle de vie du logiciel. Le Chapitre est ensuite organisé selon les étapes du cycle de vie. Aussi nous développons d'abord les éléments de conception de l'interaction : dans le paragraphe 2 nous présentons des propriétés ergonomiques, spécifiques aux systèmes mixtes. Reposant sur ces propriétés, nous détaillons les étapes de notre démarche d'analyse ergonomique dans le paragraphe 3, démarche illustrée ensuite dans le paragraphe 4. Nous présentons enfin notre modèle d'architecture adaptée du modèle PAC-Amodeus pour la réalisation logicielle. Les liens avec nos résultats conceptuels en font une solution générique pour la réalisation logicielle de systèmes mixtes.

## **Chapitre V :**

# **Éléments de Conception et de Réalisation de Systèmes Mixtes**

## **1 Etapes de conception et de réalisation d'un système interactif**

Afin de situer notre contribution pour la conception et la réalisation de systèmes mixtes, nous commençons ce chapitre par un rappel des étapes de conception et de réalisation d'un système interactif. Pour cela, nous faisons d'abord appel aux modèles issus du Génie Logiciel. Le génie logiciel est concerné par la production et la maintenance de logiciels sous le contrôle d'un ensemble de contraintes. Le processus de développement nécessite un cadre structurel. De nombreux modèles existent comme souligné dans [Nigay 94, Chapitre V]. Le modèle en V, bien que simpliste, a l'avantage

de spécifier les étapes clés de tout processus de développement. Pour cette raison, nous l'utilisons comme support pour y situer ensuite les étapes de conception ergonomique propres aux systèmes interactifs. Comme le montre la Figure 40, le modèle en V distingue l'analyse des besoins, la conception, le codage et les tests :

- l'analyse des besoins a pour objet la définition des services requis du système et des contraintes de développement.
- l'activité de conception consiste à définir une solution matérielle et logicielle qui répond à l'analyse des besoins. Elle comprend deux volets essentiels : la conception proprement dite, et les conceptions globale et détaillée. La conception établit la solution à mettre en œuvre, tandis que les conceptions globale et détaillée sont résolument orientées vers la réalisation logicielle.
- le codage ou implémentation traduit les spécifications issues de l'étape de conception logicielle en un code exécutable.
- les tests permettent de vérifier les étapes précédentes par rapport au code produit. Ils se situent sur la pente ascendante du modèle en V. Chaque test correspond à une des étapes exposées ci-dessus. Bien que les tests soient effectués après le codage, leur planification est faite à la fin de chaque étape de la pente descendante du modèle en V.

Le modèle en V décrit les étapes de conception et de réalisation de tous systèmes informatiques. Il convient maintenant d'affiner ces étapes au cas de systèmes interactifs. A la Figure 40, nous situons les étapes de la conception ergonomique de l'interaction au sein du cycle en V. La conception ergonomique se décompose en cinq étapes essentielles : esquisse du problème à résoudre, identification de l'utilisateur, définition des tâches et des objets spécifiques au domaine d'application, définition des objets et fonctions informatiques, définition de la présentation de l'interface :

- les deux premières étapes concernent d'une part la définition du problème, et d'autre part la modélisation des utilisateurs cibles. Cette modélisation s'appuie en général sur des interviews et des observations des utilisateurs en situation de travail.

Ces deux étapes participent à la définition des besoins du cycle en V.

- la tâche que l'utilisateur doit réaliser avec le système est l'objet de la troisième étape : définition des tâches et des objets spécifiques au domaine d'application. Il s'agit dans un premier temps de définir les objets et opérateurs du domaine puis de les organiser dans un arbre de tâches. La technique usuelle pour construire cet arbre est d'adopter une approche descendante de la tâche la plus abstraite aux tâches élémentaires manipulant des objets du domaine. Chaque tâche de l'arbre est ensuite décorée : ces attributs ont un impact direct sur la conception de l'interaction. Des exemples d'attributs incluent la complexité ou encore la fréquence de la tâche. Un des attributs est l'importance de la tâche que nous retrouvons sous la forme d'une caractéristique dans ASUR : le poids d'une relation ASUR. L'arbre décoré est ensuite validé : des incohérences syntaxiques peuvent être détectées lors de cette analyse.

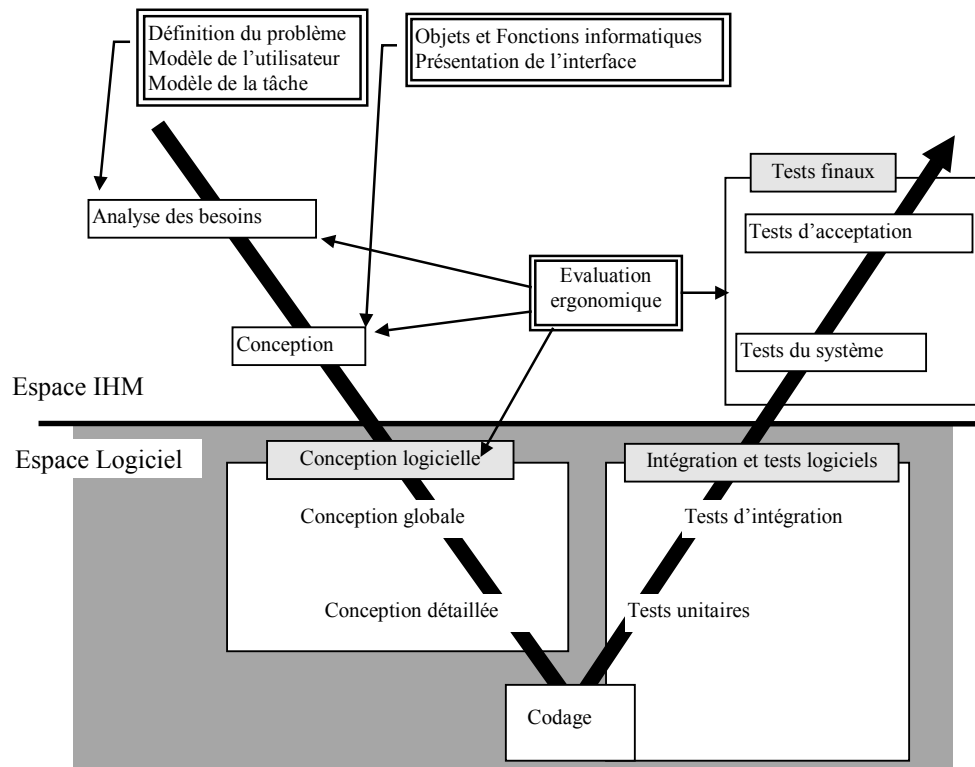
Le modèle de tâche résultant de cette étape appartient là encore à l'analyse des besoins du cycle en V, en contribuant à la définition des services requis du système.

- l'étape suivante, la définition des objets et fonctions informatiques, comporte l'identification de l'ensemble des concepts manipulés dans l'interaction. Cette étape s'appuie sur la précédente. En effet, chaque tâche élémentaire est traduite en termes de concepts informatiques et de fonctions les manipulant. Définissant l'ensemble des objets manipulés lors de l'interaction, cette étape appartient à la phase de conception du cycle en V.
- la dernière étape, la définition de la présentation de l'interface, consiste à définir le comportement perceptible des objets informatiques identifiés à l'étape précédente. Une cohérence globale de l'interaction doit être visée, grâce à l'application d'une métaphore d'interaction par exemple. Le résultat de cette étape est une description de l'interaction en sortie, du système vers l'utilisateur et de celle en entrée, de l'utilisateur vers le système. La description de l'interaction, résultat de cet étape, fait l'objet du document intitulé Spécifications Externes, un des documents à fournir à la fin de la phase de conception du cycle en V.

Ce schéma en cinq étapes est une vue simplifiée du processus de conception ergonomique de l'interaction. En pratique, les étapes sont chacune soumises à des révisions accompagnées de nombreux retours arrière, suite à l'évaluation de l'utilisabilité notée évaluation ergonomique. Aussi l'évaluation ergonomique est placée au centre de la Figure 40 pour traduire le fait que chaque étape doit faire l'objet d'évaluation ergonomique.

Outre les étapes de conception ergonomique de l'interaction, des modèles d'architecture logicielle et des outils comme des générateurs d'interface sont propres à la réalisation logicielle de systèmes interactifs. Tandis que les outils dédiés aux interfaces sont utiles lors de la phase de codage du cycle en V, les modèles d'architecture logicielle de systèmes interactifs sont utilisés pour la conception logicielle lors de la phase de conception globale du cycle en V. Rappelons qu'un modèle d'architecture de systèmes interactifs définit les éléments directeurs d'aide à l'organisation modulaire du logiciel de l'interface d'un système informatique. Parce que ces modèles sont propres aux systèmes interactifs, ils doivent véhiculer une forte modifiabilité du code. En effet la modifiabilité est nécessaire et incontournable dans le cadre d'un processus itératif de mise au point de l'interface. De plus, comme le souligne la Figure 40, ces modèles d'architecture occupent une situation charnière entre la conception ergonomique et la conception logicielle, situation tiraillée entre les propriétés ergonomiques et les propriétés du code issues du génie logiciel. Ainsi des propriétés ergonomiques utiles à la conception ergonomique ont des répercussions sur l'organisation modulaire du code : des liens entre les propriétés ergonomiques et les modèles d'architecture peuvent être établis [Nigay 97]. Ce constat justifie, à la Figure 40, l'évaluation ergonomique effectuée lors de la phase de conception globale.





**Figure 40** Mise en relation des étapes du cycle de vie en V du logiciel (cadre simple) avec les étapes de conception de l'interaction (cadre double).

Nous avons mentionné que des retours en arrière dans le processus de conception sont autorisés dans le cas où des problèmes seraient identifiés lors de l'évaluation ergonomique. Plus les problèmes sont identifiés tôt dans le cycle, moins leur rectification par retour arrière est coûteuse. Ce point justifie la nécessité d'une évaluation ergonomique à chaque étape. Lorsque cette évaluation concerne des étapes en amont du codage, on dit qu'il s'agit d'évaluation ergonomique prédictive.

Notre contribution à la conception de systèmes mixtes réside dans une démarche d'analyse ergonomique prédictive. Dans cet objectif, nous identifions dans le paragraphe 2 des propriétés ergonomiques qui sont propres aux systèmes mixtes. Nous montrons dans le paragraphe 3 comment notre notation ASUR permet l'étude des propriétés ergonomiques dans le cadre d'une analyse prédictive. Dans le paragraphe 4, nous illustrons notre démarche d'analyse ergonomique avec deux systèmes.

Notre contribution à la réalisation logicielle de systèmes mixtes consiste en un modèle d'architecture logicielle, qui est une adaptation du modèle PAC-Amodeus. Nous expliquons dans le paragraphe 5 ses principes et les points d'ancrage du modèle avec nos résultats conceptuels.

## 2 Propriétés ergonomiques de la dualité de l'interaction dans un système mixte

La particularité d'un système mixte par rapport à un système interactif normal réside dans la dualité de l'interaction induite par ces systèmes. En effet, l'interaction de

l'utilisateur avec le système se compose de deux facettes : l'une à destination des entités virtuelles, l'autre à destination d'entités réelles, telles que les outils réels manipulés par exemple. Par conséquent, l'étude de l'utilisabilité du système doit tenir compte de cette spécificité. L'utilisabilité d'un système mixte doit donc considérer les propriétés établies d'un système interactif, mais doit aussi étudier la combinaison des deux facettes de l'interaction. Garantir l'harmonie des deux facettes de l'interaction est donc un objectif supplémentaire à prendre en compte lors de l'étude de l'utilisabilité d'un système mixte.

L'approche que nous avons adoptée pour répondre à cet objectif s'appuie sur des propriétés ergonomiques des IHM ainsi que sur celles proposées pour la conception de systèmes de GMCAO. Ainsi, interagir avec un système interactif requiert une interprétation correcte par l'utilisateur des données produites par le système. Un point central de la conception d'une interface consiste à rendre perceptible des données pertinentes vis à vis de la réalisation de la tâche de l'utilisateur, et ce de manière correcte et facilement interprétable. Les travaux de [Radermacher 96] et [Rau 96] ont à ce titre identifié douze principes devant servir de base au développement de règles ergonomiques pour la conception de systèmes de GMCAO. Parmi ces douze principes, certains ont trait à la sécurité, à la tolérance du rythme et à des conditions plus générales liées à l'environnement de travail spécifique que rencontrent les cliniciens. Trois principes expriment des requis pour les données fournies par le système :

- adaptation à la tâche : présenter seulement les concepts pertinents pour la réalisation de la tâche et privilégier les métaphores du monde réel.
- conformité aux attentes de l'utilisateur : présenter les données de manière cohérente en s'appuyant sur les connaissances de l'utilisateur, ses habitudes, ses attentes, le sujet de sa tâche et les conventions.
- auto-description : simplifier l'interaction tout en la laissant explicite et simple à apprendre.

Ces trois principes expriment le fait qu'une interface de systèmes mixtes doit permettre à l'utilisateur de percevoir des données pertinentes (adaptation), de savoir s'en servir (conformité) et de la comprendre (auto-description). Ces principes rejoignent deux propriétés ergonomiques du domaine des IHM caractérisant les sorties d'un système interactif : **l'observabilité** et **l'honnêteté**. L'observabilité caractérise la capacité du système à permettre à l'utilisateur de percevoir l'état interne du système, c'est-à-dire un concept, au travers des données fournies par le système [Abowd 92] [Dix 98] [Gram 96]. L'honnêteté caractérise la capacité du système à fournir des données relatives à un concept manipulé, qui soient en accord avec l'état interne du système et de manière à favoriser une interprétation correcte de la part de l'utilisateur [Abowd 92] [Gram 96]. Ces deux propriétés sont utilisées pour caractériser la présentation d'un concept manipulé par le système, à un instant donné au cours de l'interaction.

Compte tenu de la dualité de l'interaction dans les systèmes mixtes, les utilisateurs se trouvent fréquemment exposés à plusieurs sources de données : un objet réel et sa représentation numérique (virtuelle), multiples procédés d'imagerie mettant en évidence différentes caractéristiques d'un même objet, représentation simultanée de plusieurs concepts, etc. La cohérence de ses différentes sources de données doit être garantie sans quoi l'utilisateur risque d'interpréter de manière erronée les données perçues. L'effet McGurk [McGurk 76] est une bonne illustration de ce phénomène : une bande vidéo

montrant un orateur est présentée à un sujet. L'orateur prononce la syllabe "ga" mais la bande sonore a été réenregistrée avec la syllabe "ba". Lors de l'expérimentation, 98% des sujets rapportent entendre la syllabe "da". Cet exemple très simple en comparaison des situations qu'un utilisateur de système mixte peut rencontrer, motive la nécessité de l'extension des propriétés ergonomiques précitées, l'observabilité et l'honnêteté, de manière à garantir l'harmonie des deux facettes de l'interaction.

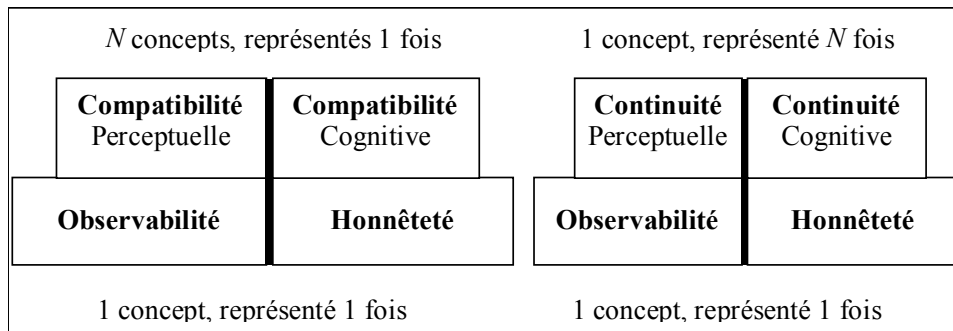
Partant de ce constat, nous complétons ces deux propriétés afin de considérer l'observabilité et l'honnêteté de multiples concepts à un instant donné et de garantir ainsi que la combinaison des données présentées simultanément fasse sens pour l'utilisateur. C'est par exemple le cas en bloc opératoire lorsque l'utilisateur est amené à percevoir plusieurs données médicales simultanément. Afin de compléter ces propriétés d'observabilité et d'honnêteté, nous avons ajouté deux nouvelles propriétés à celles existantes :

- la **compatibilité perceptuelle**, extension de la propriété d'observabilité au cas de N concepts, traduit la possibilité pour l'utilisateur de percevoir les données relatives à N concepts en prenant en compte leur dispersion géographique dans l'environnement et les sens perceptifs mis en jeu. La compatibilité perceptuelle implique donc l'observabilité de chacun des concepts considérés isolément.
- la **compatibilité cognitive**, extension de la propriété d'honnêteté au cas de N concepts, traduit la capacité de l'utilisateur à interpréter correctement les données relatives à N concepts en considérant les différents modes de représentation mis en œuvre par le système pour représenter ces concepts. Là encore, la compatibilité cognitive implique l'honnêteté des représentations de chacun des concepts considérés isolément.

Cas particulier de la perception et interprétation de plusieurs concepts, l'utilisateur est amené à percevoir plusieurs représentations relatives à un même concept. Nous avons dénoté ce cas particulier de la compatibilité par le terme continuité. Comme précédemment, cette propriété s'applique aux niveaux perceptuel et cognitif, ce qui donne lieu à deux nouvelles propriétés :

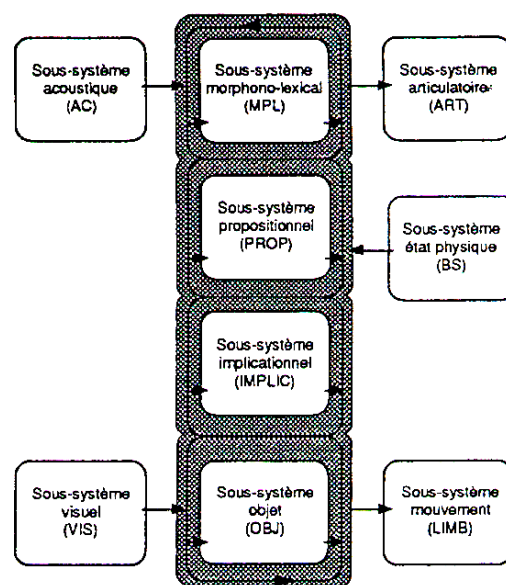
- la **continuité perceptuelle**, adaptation de la compatibilité perceptuelle, traduit la possibilité pour l'utilisateur de percevoir les données relatives à UN concept en prenant en compte leur dispersion géographique dans l'environnement de l'utilisateur et les sens perceptifs mis en jeu.
- la **continuité cognitive**, adaptation de la compatibilité cognitive, traduit la capacité de l'utilisateur à interpréter correctement les données relatives à UN concept en considérant les différents modes de représentation mis en œuvre par le système pour représenter ces données.

La Figure 41 présente les deux propriétés de compatibilité et de continuité par rapport à l'observabilité et l'honnêteté.



**Figure 41** Propriétés ergonomiques : observabilité, honnêteté, compatibilité et continuité.

Ces propriétés caractérisent des capacités du systèmes, mais elles ne peuvent être considérées indépendamment de l'utilisateur : elles doivent être en accord avec les capacités de l'utilisateurs. Pour expliquer ces propriétés du point de vue de l'utilisateur, nous nous tournons vers la psychologie cognitive qui a pour objet d'étude la perception et l'interprétation. Pour cela, nous expliquons les propriétés dans la théorie de psychologie cognitive ICS, Interacting Cognitive Subsystems [Barnard 85]. La théorie ICS structure le système de traitement de l'information humaine en un ensemble de neuf sous-systèmes comme le montre la Figure 42 ; ICS modélise l'activité mentale de l'utilisateur en termes de flux de données entre des sous-systèmes dédiés à la perception (sous-systèmes perceptifs, notés AC, VIS et BS), la cognition (sous-systèmes centraux, notés MPL, PROP, IMPLIC et OBJ) et l'action (sous-systèmes effecteurs, notés ART et LIMB). Chaque sous-système est composé d'entrées, de sorties, de capacités de traitements de l'information et d'une mémoire. Tout sous-système prend en entrée un seul type de signal, visuel ou sonore par exemple, et peut fournir en sortie un signal compréhensible par d'autres sous-systèmes. Les sous-systèmes perceptifs sont chargés d'acquérir l'information en provenance de l'environnement ou du corps. Les sous-systèmes centraux réalisent des transformations de l'information acquise. Ces sous-systèmes centraux sont au cœur du processus d'interprétation. Les sous-systèmes effecteurs sont responsables des actions de l'utilisateur.



**Figure 42** Schématisation du modèle ICS.

L'intérêt de cette théorie est qu'elle utilise des termes du domaine de l'informatique, tel que système ou encore processus, pour modéliser l'activité mentale de l'humain. Ainsi cette théorie est compréhensible par les concepteurs de systèmes interactifs. De plus, cette théorie a été utilisée pour étudier l'intégration et la combinaison de nouveaux médias [Salber 95] et constitue la base d'une modélisation de l'interaction multimodale [May 97]. Nous montrons ici comment la théorie ICS peut nous aider à comprendre qualitativement les difficultés perceptuelles et cognitives rencontrées par un utilisateur face à la multiplicité des dispositifs et des langages mis en œuvre pour lui fournir des informations. En effet, les sous-systèmes peuvent être considérés comme un ensemble de ressources perceptuelles et cognitives et ICS peut indiquer et justifier qu'une ressource risque d'être surchargée.

Ainsi, au niveau de la perception, et donc en relation avec le niveau perceptuel de la compatibilité et de la continuité, la théorie ICS explique que les sous-systèmes perceptifs peuvent acquérir de l'information en parallèle. De plus les limitations intrinsèques de nos sens ou certaines de leurs particularités permettent une perception globale d'informations discrètes : l'utilisateur peut par exemple percevoir plusieurs concepts affichés à l'écran. Ainsi sur un même canal sensoriel, plusieurs concepts peuvent être perçus en parallèle.

Au niveau cognitif, la théorie ICS permet de déterminer les conditions que les différentes sources d'information doivent respecter afin d'être interprétées comme un tout cohérent. ICS donne des critères pour que ces combinaisons soient possibles. ICS prédit que la combinaison d'informations provenant d'un ou plusieurs canaux sensoriels dans les niveaux supérieurs centraux ne peut avoir lieu que si les flots d'information aux entrées du sous-système considéré sont simultanés et sont cohérents. Or il est impossible de déterminer le temps de traitement de chaque information. Par conséquent, l'évaluation de la compatibilité et de la continuité au niveau cognitif ne peut reposer que sur des évaluations expérimentales auprès d'un ensemble de sujets représentatifs des utilisateurs du système mixte. Aussi dans le cadre d'une évaluation prédictive, nous exposons dans le paragraphe 3 des facteurs influençant la compatibilité et la continuité au niveau cognitif : ces facteurs sont exprimés dans les termes ASUR. Comme nous l'enseigne ICS, nous ne pouvons pas a priori affirmer que la propriété est vérifiée ou non, nous ne pouvons que fournir des facteurs en faveur ou à l'encontre de la compatibilité et de la continuité au niveau cognitif. Ces assertions lors de la conception doivent ensuite être vérifiées par évaluation expérimentale lorsque le système est développé.

Les quatre propriétés ergonomiques présentées dans ce paragraphe caractérisent l'utilisabilité d'un système mixte en prenant en compte la nécessité d'interprétation d'informations perçues dans deux mondes distincts. Il conviendrait d'approfondir cette approche en étudiant d'autres propriétés ergonomiques notamment en ce qui concerne les entrées du système. En effet, la cohérence des actions requises par l'utilisateur pour transmettre des données au système fait également face à la dualité de l'interaction : par exemple le clinicien doit être capable de manipuler un outil chirurgical tout en énonçant une commande au système. Aussi dans ce paragraphe sans viser l'exhaustivité, notre objectif a été d'identifier des propriétés répondant à des caractéristiques de l'interaction qui soient propres aux systèmes mixtes.

Toutefois, il est crucial de noter que la présence de facteurs d'incompatibilité ou de discontinuité, perceptuelles ou cognitives, n'est pas synonyme de non utilisabilité du système ou de défaut de conception. Par exemple, la superposition d'une légende textuelle, unidimensionnelle selon notre approche, à une carte géographique bidimensionnelle, constitue du fait de la différence entre les dimensions des langages utilisés, un facteur d'incompatibilité cognitive. Pourtant, l'usage de légende est habituel. Il conviendrait donc de mener des études utilisateurs, en collaboration avec des psychologues cognitivistes et des ergonomes, afin de quantifier l'impact des différentes formes d'incompatibilité ou de discontinuité (utilisation de langages aux dimensions différentes, référentiels de présentations multiples, etc.). Nous préconisons donc, pour compléter cette approche, d'identifier des facteurs d'incompatibilité et de discontinuité et d'en évaluer leurs conséquences sur l'utilisabilité du système. Inversement, la mise en œuvre par un système mixte de l'une de ces quatre propriétés ne présume pas de l'utilisabilité globale du système. En effet, l'utilisation de propriétés ergonomiques amène le concepteur à devoir faire des compromis. Il est donc possible que la non compatibilité d'une interaction soit justifiée par le choix du concepteur de faire respecter une autre propriété ergonomique plus importante pour le système et la tâche considérée, comme par exemple la cohérence entre systèmes.

Dans le paragraphe suivant, nous étudions les quatre propriétés au sein d'une description ASUR. Ainsi nous fournissons un outil de conception, qui aide le concepteur à raisonner de façon prédictive sur les propriétés ergonomiques du système à développer.

### 3 Démarche d'analyse prédictive de la compatibilité et de la continuité

Notre démarche d'analyse s'appuie sur la notation ASUR, présentée au Chapitre IV. Rappelons que les relations ASUR représentent un flux de données entre deux composants ASUR. Par conséquent toutes les relations reliées à l'utilisateur, composant U, représentent les différentes parties formant l'interaction de l'utilisateur avec le système. Les relations connectées à un composant R, objet réel, composent donc l'interaction avec le monde réel, tandis que les relations connectées à un adaptateur d'entrée ou de sortie ( $A_{in}$  ou  $A_{out}$ ) définissent l'interaction de l'utilisateur avec le monde virtuel. La compatibilité et la continuité ont donc trait à l'interaction de l'utilisateur avec le système mixte du point de vue des sorties proposées par le système. Etudier la compatibilité ou la continuité nécessite donc l'analyse des relations ASUR à destination de l'utilisateur d'une part, et d'autre part des composants à partir desquels l'utilisateur obtient les données. Les composants ASUR étudiés sont donc les adaptateurs de sortie ( $A_{out}$ ) ainsi que les objets réels utilisés comme outil ( $R_{tool}$ ) ou ayant le rôle d'objet de la tâche ( $R_{task}$ ).

Le niveau perceptuel de nos propriétés caractérise la capacité de l'utilisateur à percevoir toutes les données fournies en considérant notamment leur dispersion géographique dans l'environnement ainsi que les sens perceptifs requis par chacune des données. En terme d'ASUR, ceci se traduit par l'analyse de **l'environnement perceptif** de chaque composant ASUR participant à l'interaction de l'utilisateur avec le système. Si la superposition des environnements perceptifs requis n'est pas possible alors cela caractérisera une incompatibilité ou une discontinuité perceptuelle. Un environnement

perceptif, caractéristique d'un composant ASUR, est défini par un couple (sens perceptif, lieu de perception) au Chapitre IV. Le fait que les environnements perceptifs ne soient pas superposables peut donc être dû à l'un ou l'autre de ces deux éléments constitutifs.

Le niveau cognitif de la compatibilité et de la continuité traduit la capacité du système à fournir des données de manière à favoriser une interprétation correcte par l'utilisateur. Compte tenu des observations mentionnées dans le paragraphe précédent, l'évaluation de la compatibilité et de la continuité cognitive ne peut donner lieu qu'à l'identification de facteurs favorisant ou défavorisant l'interprétation des données par l'utilisateur. Dans la description ASUR d'un système, les données étant portées par les relations ASUR, ces facteurs d'incompatibilité et de discontinuité résident donc parmi les caractéristiques des relations ASUR de poids élevé et constituant l'interaction de l'utilisateur avec le système. Ainsi, les cinq caractéristiques des **langages** utilisés, leur dimension (1D, 2D ou 3D) et les quatre caractéristiques du langage selon [Bernsen 94] (analogique ou non, arbitraire ou non, linguistique ou non, statique ou non), ainsi que le **référentiel de présentation** adopté pour chaque relation ASUR constituent des pistes de réflexions quant à l'évaluation de la capacité de l'utilisateur à interpréter les données fournies simultanément au moyen de langages différents. Il conviendrait d'identifier précisément lesquelles ont une plus grande influence. Pour se faire, une collaboration avec des psychologues cognitivistes est requise, afin de mener une étude expérimentale visant à identifier les caractéristiques majeures du langage.

Quelque soit le niveau d'analyse, perceptuel ou cognitif, les relations et les composants ASUR concernés par l'étude de la compatibilité et de la continuité ne sont pas les mêmes. Dans le cas de la compatibilité, il convient de prendre en compte toutes les relations orientées vers l'utilisateur, ainsi que les composants situés à l'autre extrémité de ces relations. Par contre dans le cas de la continuité, seules les relations porteuses de données relatives à un même concept sont à considérer. Il en va de même pour les composants : seuls les composants connectés à ces relations doivent être pris en compte pour l'analyse de la continuité. La Figure 43 résume les outils d'analyse de nos propriétés ergonomiques dans ASUR en regroupant les composants et les relations ASUR, ainsi que leurs caractéristiques fournissant un support à l'analyse de la compatibilité et de la continuité dans l'interaction en sortie d'un système mixte.

<b>Niveau Perceptuel</b>	<b>Environnements Perceptifs</b> de <b>tous les composants</b> impliqués dans l'interaction de sortie	<b>Environnements Perceptifs</b> des <b>composants</b> impliqués dans l'interaction de sortie pour <b>un</b> concept en particulier
<b>Niveau Cognitif</b>	<b>Langages et</b> <b>référentiels de présentation</b> de <b>toutes les relations</b> impliquées dans l'interaction de sortie	<b>Langages et</b> <b>référentiels de présentation</b> des <b>relations</b> impliquées dans l'interaction de sortie pour <b>un</b> concept en particulier
	<b>Compatibilité</b>	<b>Continuité</b>

**Figure 43** *Caractéristiques ASUR pour l'analyse prédictive des compatibilités et des continuités perceptuelles et cognitives de l'interaction en sortie d'un système mixte.*

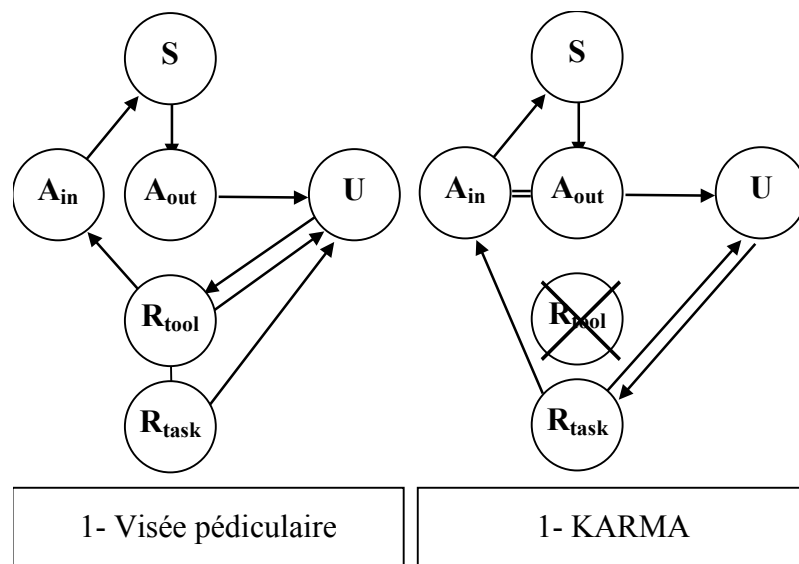
La démarche d'analyse ergonomique permettant l'étude de la compatibilité et de la continuité de l'interaction en sortie d'un système mixte se déroule en neuf étapes :

1. **Décrire** le système avec ASUR pour une tâche donnée
2. Identifier les **concepts** majeurs manipulés par le système
3. Identifier les **relations** ASUR, de poids élevé, composant l'interaction de sortie du système
4. Caractériser les **environnements perceptifs** des composants ASUR à l'origine des relations identifiées à l'étape 3
5. Caractériser les **langages** et les **référentiels de présentation** mis en œuvre par les relations identifiées à l'étape 3
6. Analyser la **compatibilité** de l'interaction aux niveaux perceptuel et cognitif en suivant les indications de la Figure 43
7. Dédurre parmi les **relations** identifiées à l'étape 3, celles relatives à un même concept
8. Analyser la **continuité** de l'interaction aux niveaux perceptuel et cognitif en suivant les indications de la Figure 43
9. **Recommencer** à l'étape 7 pour un autre concept.

Nous illustrons les propriétés ainsi que la démarche d'analyse dans le paragraphe suivant avec deux systèmes mixtes : *Karma* [Feiner 93], présenté au chapitre II et la *visée pédiculaire* [Merloz 97], présentée au chapitre I. Ces deux systèmes ont également été décrit avec notre notation ASUR dans le chapitre IV, dans le paragraphe 4.

#### 4 Illustration de la démarche

Nous rappelons tout d'abord les descriptions ASUR de ces deux systèmes à la Figure 44, puis nous analysons la compatibilité et la continuité de l'interaction induite par ces deux systèmes.



**Figure 44** Description avec ASUR de la Visée Pédiculaire et de KARMA.

Aussi dans les deux paragraphes suivants nous conduirons l'analyse ergonomique de ces deux systèmes à partir de l'étape 2 de notre démarche d'analyse.



## 4.1 Analyse de KARMA

### 4.1.1 Mise en œuvre de la démarche

#### 2. Identification des concepts principaux

La tâche considérée est la réparation de l'imprimante. Les concepts intervenant lors de l'utilisation de *KARMA* sont l'imprimante à réparer, ainsi que les données de maintenance. Or les données de maintenance sont constituées d'une représentation des éléments de l'imprimante à manipuler. Le seul concept manipulé est donc finalement l'imprimante.

#### 3. Identification des relations ASUR, de poids élevé, formant l'interaction de sortie du système

Deux relations, de poids élevé, composent l'interaction en sortie du système avec l'utilisateur :

- $A_{out} \rightarrow U$ , représentant la perception des données de maintenance par le technicien,
- $R_{task} \rightarrow U$ , matérialisant la perception directe de l'imprimante par le technicien. La perception d'informations tactiles liée à la manipulation d'éléments de l'imprimante est également représentée par cette relation, mais cette composante à un poids faible. Nous ignorons donc cette composante de la relation  $R_{task} \rightarrow U$  dans la suite de cette analyse.

#### 4. Caractérisation des environnements perceptifs des composants ASUR à l'origine des relations identifiées à l'étape 3

Pour l'adaptateur de sortie, le casque semi-transparent, le sens perceptif requis est la vue. Le lieu de perception est l'imprimante puisque les données affichées sont alignées avec l'imprimante réelle.

Pour l'objet de la tâche, l'imprimante, le sens perceptif requis est également la vue et le lieu de perception est bien entendu limité à l'imprimante elle-même.

#### 5. Caractérisation des langages et des référentiels de présentation mis en œuvre par les relations identifiées à l'étape 3

Les langages utilisés par les deux relations identifiées à l'étape 3 sont :

- pour  $A_{out} \rightarrow U$ , un langage graphique 3D, non arbitraire, analogique, non linguistique et statique puisqu'il s'agit d'une représentation 3D des éléments de l'imprimante à déplacer. Le référentiel de présentation est le référentiel utilisateur.
- pour  $R_{task} \rightarrow U$ , le langage correspondant à la visualisation naturelle que nous qualifions de 3D. Il est de plus non arbitraire, analogique, non linguistique et statique. Le référentiel de présentation est le référentiel utilisateur également.

#### 6. Analyse de la compatibilité de l'interaction

Les relations ASUR composant l'interaction en sortie, c'est-à-dire du système vers l'utilisateur, sont porteuses de données relatives au même concept : l'imprimante. Par conséquent l'analyse de la compatibilité n'a pas lieu d'être menée et seule l'analyse de la continuité est pertinente.

#### 7. Identification des relations relatives à un même concept

Il s'agit des mêmes relations que celles identifiées à l'étape 5.

#### 8. Analyse de la continuité de l'interaction

Du point de vue perceptuel, les deux environnements perceptifs requis sont les mêmes et l'interaction ainsi conçue garantit la continuité perceptuelle.

Du point de vue cognitif, les caractéristiques des langages des deux relations ainsi que les référentiels de présentation sont les mêmes. Aucune des conditions que nous avons identifiées et exprimées en termes des caractéristiques des relations, et qui contribuent à une discontinuité cognitive n'est illustrée par ce système. Par conséquent, la continuité cognitive de l'interaction semble particulièrement favorisée par ce système.

#### 9. Recommencer à l'étape 7 pour un autre concept

Le seul concept impliqué dans l'interaction en sortie du système est l'imprimante. L'analyse ergonomique de *KARMA* est donc terminée.

### 4.1.2 Conclusion et discussion

Le système Karma garantit donc la continuité perceptuelle en sortie, et présente des caractéristiques très favorables à la continuité cognitive. Notons également que l'étude de la compatibilité en sortie n'est pas utile dans Karma car les données fournies par le système ne concernent qu'un seul concept.

Revenons sur l'aspect perceptuel de la continuité et imaginons maintenant que l'environnement perceptif requis par l'adaptateur de sortie soit localisé dans la main de l'utilisateur. Ce serait le cas si les données étaient affichées sur un *Personal Digital Assistant*, un *Palm Pilot* par exemple. Dans ce cas, les environnements perceptifs ne seraient plus les mêmes. De plus, le technicien ne pourrait percevoir les données présentes dans les deux environnements perceptifs que de manière exclusive. Les environnements perceptifs ne seraient alors plus superposables et ce serait une illustration de discontinuité perceptuelle liée à la localisation des données.

Du point de vue cognitif, la continuité est favorisée grâce à l'expression de données dans des langages aux caractéristiques identiques comme un référentiel de présentation commun. Les processus d'interprétation des données portées par les relations sont donc les mêmes. En s'appuyant sur ICS, les conditions (temps de traitement des flux d'informations) sont réunies pour une fusion des informations au niveau cognitif. Considérons maintenant le cas où le technicien recevrait l'aide par le biais d'instructions textuelles. Les langages se baseraient alors sur la visualisation 3D analogique, non arbitraire, non linguistique et statique d'une part et, d'autre part sur un langage 1D textuel, c'est-à-dire un langage analogique, non arbitraire, linguistique et statique. Les différences entre les caractéristiques des deux langages (linguistique versus non linguistique ou encore les dimensions 3D versus 1D) constituent des facteurs de discontinuité cognitive. En regard de la théorie ICS, la discontinuité se traduit par des processus d'interprétation différents au sein des sous-systèmes centraux. Les temps de traitement ne sont donc pas les mêmes et la condition de simultanéité nécessaire à la fusion n'est plus garantie.

## 4.2 Analyse de la Visée Pédiculaire

### 4.2.1 Mise en œuvre de la démarche

Nous considérons ici la première version de la *visée pédiculaire*, dans laquelle les données de guidage sont présentées par le système sous la forme d'un viseur constitué de trois croix. Une description plus détaillée de ce système se trouve dans le chapitre I.

#### 2. Identification des concepts principaux

La tâche considérée est ici la phase de guidage de ce système de GMCAO. Elle consiste en la réalisation d'un trou dans une vertèbre à l'aide d'un foret et d'une perceuse. Les concepts intervenant dans l'interaction en sortie de ce système sont le patient, l'outil et la trajectoire à suivre avec l'outil.

#### 3. Identification des relations ASUR, de poids élevé, formant l'interaction de sortie du système

Trois relations composent l'interaction en sortie du système avec le clinicien :

- $A_{out} \rightarrow U$ , qui traduit la perception des données de guidage par le clinicien sur l'écran,
- $R_{tool} \rightarrow U$ , qui représente la perception qu'a le clinicien de l'outil chirurgical,
- $R_{task} \rightarrow U$ , qui exprime la perception du patient par le clinicien.

#### 4. Caractérisation des environnements perceptifs des composants ASUR à l'origine des relations identifiées à l'étape 3

Ce système est caractérisé au chapitre IV : deux environnements perceptifs sont identifiés dans la visée pédiculaire. Le premier est constitué du couple (vision, 3D) pour la perception des données de guidage par le clinicien, et (vue / haptique, champ opératoire) pour la perception de l'outil réel et du patient par le clinicien.

#### 5. Caractérisation des langages et référentiels de présentation mis en œuvre par les relations identifiées à l'étape 3

Les langages utilisés par les relations identifiées à l'étape 3 sont :

- pour  $A_{out} \rightarrow U$ , un graphique 2D, représentant un viseur. Cette représentation est arbitraire, non analogique, non linguistique et dynamique.
- pour  $R_{tool} \rightarrow U$ , la réalité 3D, non arbitraire, analogique, non linguistique et statique,
- pour  $R_{task} \rightarrow U$ , la réalité 3D, non arbitraire, analogique, non linguistique et statique.

Le référentiel de présentation utilisé pour la perception des entités réelles (foret, patient) est le référentiel utilisateur. Par contre, les données de guidage fournies par la relation  $A_{out} \rightarrow U$  sont exprimées selon un référentiel de présentation lié au foret, c'est-à-dire à l'outil chirurgical.

#### 6. Analyse de la compatibilité de l'interaction

L'analyse de tous les environnements perceptifs mis en jeu par les composants à l'origine de l'interaction en sortie nous conduit à identifier deux lieux différents : l'écran et le champ opératoire où se trouvent le patient

et les outils chirurgicaux. Il s'agit donc ici d'un exemple **d'incompatibilité perceptuelle**, liée à la localisation des données.

Il résulte de l'analyse des langages que des caractéristiques différentes sont mises en jeu par les différents langages utilisés : langage bidimensionnel et arbitraire pour les données de guidage et, tridimensionnel et non arbitraire pour la perception de la réalité. Ces différences constituent deux facteurs **d'incompatibilité cognitive**. L'utilisation de ces deux langages peut être à l'origine d'un effort d'adaptation de la part du clinicien lorsqu'il doit percevoir alternativement des données 2D puis des données 3D : l'utilisateur doit mettre en correspondance ces différentes données. De plus, l'une des données est basée sur une représentation arbitraire (le viseur et les trois croix) tandis que l'autre est non arbitraire. Cette différence entre les deux langages utilisés accroît la discontinuité cognitive.

#### 7. Identification des relations relatives à un même concept

Parmi les données perçues par le clinicien, seules des données concernant l'outil chirurgical sont portées par plusieurs relations ASUR. Il s'agit des relations  $A_{out} \rightarrow U$  et  $R_{tool} \rightarrow U$ .

#### 8. Analyse de la continuité de l'interaction en sortie

Les mêmes observations qu'à l'étape 6 peuvent être faites et conduisent à l'identification d'une **discontinuité perceptuelle** due à l'existence de deux lieux de perception des données qui sont mutuellement exclusifs.

Les observations de l'étape 6 nous conduisent également à identifier une propension à la **discontinuité cognitive** due à l'utilisation de deux langages différents de par leurs dimensions (3D et 2D) mais aussi de par leurs formes (non arbitraire et arbitraire).

#### 9. Recommencer à l'étape 7 pour un autre concept

Le seul concept impliqué dans plusieurs relations de l'interaction en sortie du système est l'outil manipulé, et plus précisément son orientation et sa position. L'analyse ergonomique de la *visée pédiculaire* selon notre démarche, est donc terminée.

### 4.2.2 Conclusion et discussion

La première version de la *Visée pédiculaire*, visualisant les données de guidage selon une représentation assimilable à un viseur, présente donc une **incompatibilité perceptuelle**. Celle-ci est liée à la présentation de données dans deux lieux distincts qui empêche l'utilisateur de percevoir simultanément les données. L'analyse ergonomique selon ASUR révèle également deux facteurs **d'incompatibilité cognitive** : d'une part la dimension des langages utilisée n'est pas la même (3D pour la perception de l'outil et du patient, 2D pour les données de guidage) et d'autre part, le langage utilisé pour les données de guidage est arbitraire, alors que la perception de l'outil et du patient est naturelle donc non arbitraire.

En conduisant l'étude pour les relations véhiculant des données relatives à l'outil (sa position et sa trajectoire), les conclusions sont les mêmes. Il y a donc **discontinuité perceptuelle** et deux facteurs de **discontinuité cognitive**, due à la dimension et à la forme (arbitraire ou non) des langages utilisés.

Considérons maintenant la seconde version de la *visée pédiculaire*, également présentée dans le chapitre I [Merloz 00]. Dans cette version, les données affichées à l'écran pour guider le clinicien comportent trois coupes de la structure anatomique dans laquelle une vis doit être insérée. Sur ces trois coupes apparaît une projection de la trajectoire idéale pré-planifiée ainsi qu'une projection de la représentation 3D du foret. Cette solution n'apporte aucune modification quant à **l'incompatibilité perceptuelle** puisque deux lieux de perception sont toujours requis. Au niveau cognitif par contre, la dimension des langages utilisés est toujours différente (coupe 2D sur l'écran versus réalité 3D) mais le langage utilisé pour présenter les données de guidage est cette fois-ci non arbitraire, puisqu'il s'agit d'une coupe réalisée dans la structure anatomique 3D. De plus, les coupes 2D correspondent à des informations médicales conventionnelles pour le clinicien, dont le codage n'est pas spécifique au système (au contraire du viseur basé sur les croix par exemple). Au niveau cognitif, le seul facteur restant qui favorise l'incompatibilité est donc la différence de dimensions des langages utilisés : le risque d'**incompatibilité cognitive** est donc **moindre**.

Tandis que dans la première version, la continuité ne concernait qu'un seul concept, l'outil, dans cette version, l'étude de la continuité concerne deux concepts : le patient et l'outil. Pour chacun de ces concepts, deux relations sont porteuses de données. Pour le patient, il s'agit de la perception directe du patient réel et de la perception de l'image anatomique du patient affichée sur l'écran. Pour l'outil, il s'agit de la perception directe de l'outil réel et de la perception sur l'écran de la projection dans trois plans du modèle 3D de l'outil chirurgical. Pour ces deux concepts, la **discontinuité perceptuelle** persiste. L'utilisation de langages dont les dimensions sont différentes constitue toujours un facteur de **discontinuité cognitive**. Le nombre de facteurs de discontinuité cognitive est réduit, mais s'applique à deux concepts au lieu d'un seul. Lors de la conception, le choix entre ces deux solutions fait l'objet d'un compromis : soit le concepteur privilégie la continuité des données relatives au patient, soit il privilégie l'utilisation d'un langage non arbitraire. Ce choix peut être influencé par des expérimentations visant à identifier la solution avec laquelle le clinicien obtient les meilleurs résultats lors de l'accomplissement de la tâche. Il peut aussi être influencé par les différences induites au niveau de la réalisation de la tâche par la quantité et le type des données fournies à l'utilisateur. Par exemple dans la deuxième version, la superposition de la représentation de l'outil chirurgical à la structure anatomique traitée sert à permettre au clinicien de dévier de la trajectoire planifiée tout en sachant où se trouve son outil par rapport à la structure anatomique. Il peut donc modifier légèrement la planification établie, alors que dans la première version le clinicien pouvait juste savoir la distance le séparant de la trajectoire idéale sans pouvoir évaluer clairement les impacts sur la structure anatomique.

Ces deux exemples illustrent notre démarche d'analyse prédictive de l'interaction selon nos quatre propriétés ergonomiques. Il apparaît extrêmement important de bien définir au départ du processus de conception, la tâche que l'on souhaite étudier. En effet de l'identification de la tâche dépendent les données qui vont être fournies à l'utilisateur et donc l'analyse de l'utilisabilité de l'interaction induite par le système. Par ailleurs, nous avons montré que la compatibilité et la continuité ont un impact sur la réalisation de la tâche. Cependant, incompatibilité et discontinuité ne riment pas obligatoirement avec

"inutilisabilité". Comme toute conception ergonomique, ces exemples ont illustré que des compromis devaient être établis.

Ces exemples concluent la présentation de notre démarche d'analyse prédictive de l'interaction lors de la conception d'un système mixte. L'étape suivante dans le cycle en V de la Figure 1 est la conception logicielle. Nous passons du monde de la conception ergonomique de l'interaction à celui de la conception logicielle de l'interaction.

## 5 Éléments de réalisation logicielle

Comme pour tout logiciel, la conception logicielle d'un système mixte doit s'appuyer sur un modèle d'architecture logicielle afin d'une part de favoriser sa maintenance et sa portabilité et d'autre part de mettre en œuvre les propriétés logicielles véhiculées par le modèle. Ces propriétés sont par exemple les propriétés internes du logiciel telles que la portabilité, la capacité à être évalué, la complétude des fonctions proposées, etc. [Gram 96]. Dans le cas des systèmes mixtes, il convient de dresser un cadre structurant pour une réalisation logicielle de ces systèmes, qui soit en accord avec les éléments de description fournis par ASUR ainsi que les propriétés ergonomiques décrites lors de la conception.

### 5.1 Modèle d'architecture de référence

Un principe commun à tous les modèles d'architecture logicielle pour systèmes interactifs réside dans la séparation de la partie logicielle gérant l'interaction Utilisateur-Système, du noyau fonctionnel, c'est-à-dire la partie logicielle maintenant les concepts et fonctions propres au domaine d'application. En partant de cette décomposition en deux briques essentielles, le modèle Seeheim [Pfaff 85] puis le modèle Arch [UIMS 92] proposent une décomposition plus fine du logiciel. Le modèle Arch préconise 5 composants logiciels, le Contrôleur de Dialogue étant le composant principal. Le modèle PAC-Amodeus [Nigay 94] de la Figure 45 repose sur la décomposition logicielle d'Arch et affine le Contrôleur de Dialogue en termes d'agents logiciels PAC.

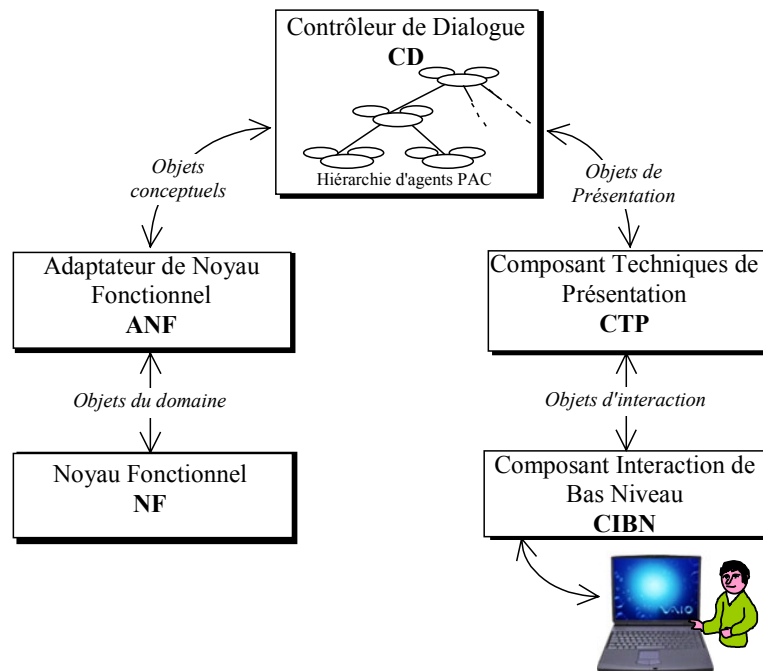


Figure 45 Composants du modèle d'architecture PAC-Amodeus [Nigay 94].

Le Noyau Fonctionnel (NF) comprend les routines et données informatiques qui sont propres au domaine d'application. Il regroupe les objets du domaine. L'Adaptateur de Noyau Fonctionnel (ANF) réalise l'interface logicielle entre le NF et le Contrôleur de Dialogue (CD) permettant ainsi au CD d'être indépendant du NF. C'est aussi dans l'ANF que le principe de réparation sémantique peut être implémentée permettant par exemple de regrouper en une seule entité qui sera présentée à l'utilisateur un ensemble de fonctionnalités que le noyau fonctionnel offre de manière indépendante.

Le CD est chargé de mettre en correspondance les concepts manipulés par l'utilisateur dans l'interface avec les concepts informatiques manipulés par le NF. Il est également chargé de gérer l'enchaînement des tâches.

Le deuxième pilier de l'arche, comprend en son sommet le Composant Techniques de Présentation (CTP). Il exprime de manière abstraite le comportement des objets interactionnels requis par l'interface. Le CTP permet donc de rendre le CD indépendant de la boîte à outils du Composant d'Interaction de bas Niveau (CIBN) implémentant les objets d'interaction de l'interface (fenêtres, boutons, etc.), de la même manière que le CD est rendu indépendant du noyau fonctionnel par le biais de l'ANF. Toujours par analogie avec l'ANF, le CTP permet la réalisation d'agrégats d'objets d'interaction dans le cas où ceux proposés par la boîte à outils ne satisfont pas les besoins du CD. Enfin, le dernier composant est le Composant d'Interaction de Bas Niveau (CIBN). Le CIBN regroupe les éléments logiciels et matériels de la plate-forme : la boîte à outils, les pilotes des dispositifs, etc. C'est donc au niveau du CIBN que les actions physiques de l'utilisateur sont captées : déplacement de la souris, touches de clavier appuyées, etc. Elles sont ensuite interprétées et transmises au CD via le CTP. Le CD vérifie alors la cohérence de l'action initiée par l'utilisateur par rapport au contexte courant d'interaction et, via l'ANF, fait appel aux fonctionnalités du NF pour effectuer la commande. Si l'action implique un retour d'information du niveau sémantique à l'utilisateur, celui-ci est généré par le NF, envoyé via l'ANF au CD qui l'adapte pour le rendre perceptible par l'utilisateur au niveau du CIBN via le CTP.

Ce modèle d'architecture constitue un cadre conceptuel adapté à l'interaction multimodale en entrée [Nigay 94] et en sortie [Vernier 01]. Nous montrons ci-après que par adaptation il est également adéquat à la conception logicielle de systèmes mixtes.

## 5.2 Adaptation du modèle aux systèmes mixtes

La présence de ces deux piliers s'accorde parfaitement avec la spécificité des systèmes mixtes. Le pilier de gauche, pilier informatique, est dédié au système informatique, c'est-à-dire aux entités du monde virtuel regroupées dans le composant S de la description avec ASUR d'un système mixte.

Classiquement, le pilier de droite, pilier réel, porte l'interaction de l'utilisateur avec le système au moyen des périphériques d'entrée/sortie du système. Dans un système mixte, ces périphériques ne sont pas limités à la souris et au clavier, ils incluent de façon générale tous les adaptateurs identifiés par la description avec ASUR du système. De plus, l'utilisateur n'est plus la seule entité réelle à être impliquée dans l'interaction. Le pilier réel est donc également responsable de l'intégration des entités réelles utilisées dans le système mixte. Le pilier réel est donc dédié à l'intégration des adaptateurs (composants

$A_{in}$  et  $A_{out}$ ) et des objets réels ( $R_{tool}$  et  $R_{task}$ ) identifiés dans la description avec ASUR du système mixte.

Un système mixte faisant intervenir plusieurs entités réelles, nous considérons que le pilier réel de l'arche du modèle PAC-Amodeus (composant CIBN et CTP) doit être composé de plusieurs branches. Chacune de ces branches est donc liée à une entité du monde réel utile à l'interaction et, la base de chaque branche est constituée d'un adaptateur permettant d'établir un lien entre l'entité réelle considérée et le reste du système. Cette adaptation aux systèmes mixtes est schématisée par la Figure 46. L'adaptation du modèle permet une indépendance entre les traitements liés aux différentes entités du monde réel intervenant dans l'interaction de l'utilisateur avec le système mixte. Cette adaptation apporte en particulier une indépendance entre les différentes boîtes à outils utilisées pour permettre l'intégration des entités réelles dans le système. Un CTP unique assure l'indépendance entre la mise en œuvre concrète des objets d'interaction (CIBN) et le CD. De multiples CTP garantissent en plus une indépendance horizontale entre les différentes boîtes à outil requises.

Cette indépendance entre les CTP de chaque branche favorise la modifiabilité du code. En effet, l'utilisation des localisateurs et autres adaptateurs est indépendante du reste du code. Un changement d'adaptateur est donc simplifié et ne nécessite pas de retrouver les concepteurs des différentes parties du code pour adapter le code au nouvel adaptateur. Réciproquement, le noyau fonctionnel et le code global de l'application sont indépendants des adaptateurs utilisés. Les adaptateurs, tels que les localisateurs utilisés dans les systèmes de GMCAO, étant parfois partagés entre plusieurs équipes de développeurs, le développement global de l'application peut se faire localement en utilisant un adaptateur de substitution, c'est-à-dire moins précis par exemple mais aussi moins onéreux.

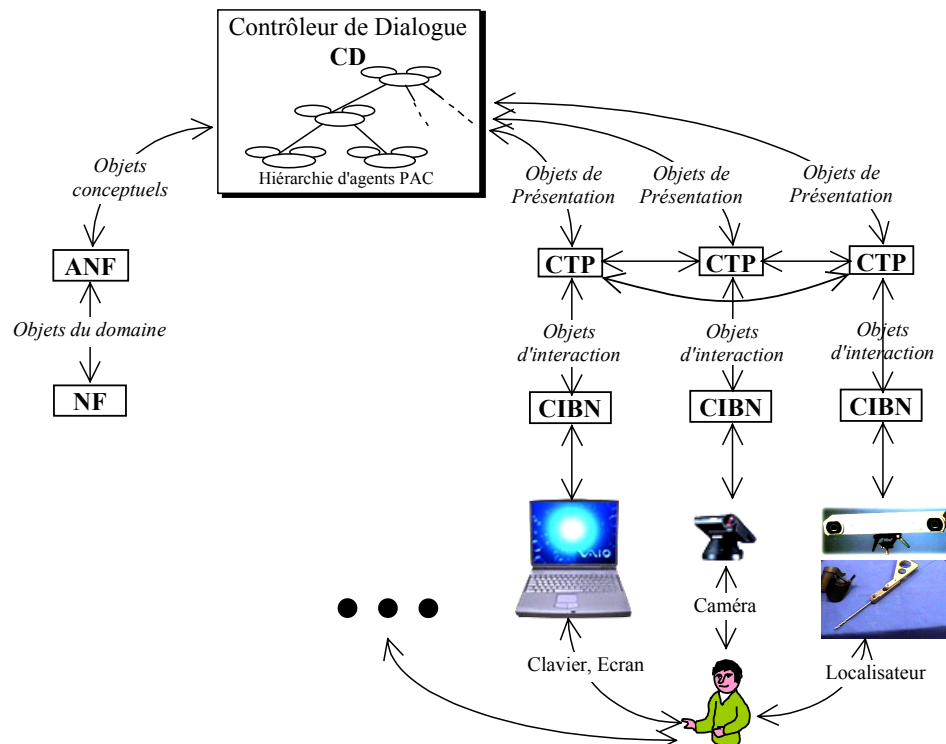


Figure 46 Adaptation du modèle d'architecture PAC-Amodeus aux systèmes mixtes.



Notre adaptation du modèle PAC-Amodeus repose sur le mécanisme appelé "*Branching Mechanism*" [Gram 96]. Ce mécanisme applicable au modèle Arch est donc applicable au modèle PAC-Amodeus.

Ce mécanisme a été conçu pour autoriser l'ajout de branches dans le pilier informatique. Deux raisons principales motivent ce mécanisme :

- variation des technologies constituant le NF,
- réutilisation de NF existants.

Une variation de la technologie constituant le NF se produit quand une application requiert les fonctionnalités offertes par deux NF distincts. Considérons par exemple, une application permettant la visualisation des résultats d'une requête spécifiée par l'utilisateur et soumise à une base de données. Le NF contient la base de données ainsi qu'un langage d'interrogation, par exemple une base de données relationnelle. Si l'application doit par la suite être capable d'interroger une base de données orientée objet, celle-ci ainsi que son langage d'interrogation représentent un second NF directement connectable sur le même ANF. En effet, les fonctionnalités du CD n'ont pas changé, il n'est donc pas utile de changer l'ANF. C'est la première mise en œuvre du "*branching mechanism*".

Une autre forme du "*Branching Mechanism*" apparaît lorsque le NF d'une application doit être complété par des fonctionnalités proposées par un autre NF. Par exemple dans un système de GMCAO, l'acquisition d'images, leur segmentation puis la visualisation de données de guidage constituent les composantes de base du NF des systèmes de GMCAO. Toutefois, le développement de techniques algorithmiques permettant la segmentation automatique d'images permet désormais de segmenter automatiquement un spectre de plus en plus large de types d'images. Aussi dans les applications précédemment conçues, il est envisageable d'ajouter ce module de segmentation automatique. Or les données nécessaires à la réalisation d'une segmentation automatique ne sont pas les mêmes que celles requises pour une segmentation manuelle. Au lieu de modifier l'intégralité du code du NF et de l'ANF et en appliquant le mécanisme "*Branching mechanism*", il convient de relier le CD à un nouvel ANF, permettant l'interface entre le CD existant et le NF de segmentation automatique existant par ailleurs. Le "*Branching Mechanism*" est donc un support à l'intégration de NF existants dans un système déjà développé.

L'application du "*Branching Mechanism*" au pilier réel de l'arche, telle que nous le proposons dans notre adaptation, n'est qu'évoquée dans [Gram 96]. La nécessité du "*Branching Mechanism*" appliqué au pilier réel de l'arche se justifie maintenant pleinement dans le cas de la réalisation de systèmes mixtes.

Il ressort donc, en adoptant un point de vue logiciel sur les systèmes mixtes, que le contrôleur de dialogue (CD) joue le rôle particulier de lien entre les entités du monde réel et celles du monde virtuel. Chaque branche du pilier réel, c'est-à-dire chaque entité du monde réel et chaque adaptateur, ne peut exploiter le NF que par l'intermédiaire du CD.

Outre l'application du "*Branching mechanism*" au pilier réel, notre adaptation réside aussi dans la communication horizontale au niveau des CTP, que nous développons dans le paragraphe suivant.

### 5.3 Flux d'informations dans le modèle adapté

Dans le modèle d'architecture PAC-Amodeus, la réalisation d'une tâche sur l'initiative de l'utilisateur fait intervenir le CTP qui capte les actions physiques, puis le CD et l'ANF et déclenche enfin des actions informatiques au niveau du NF. Les résultats suivent alors le chemin inverse jusqu'à l'utilisateur, pour que ce dernier puisse percevoir un retour d'informations. Il s'agit du flux d'informations classiques au sein de l'arche.

Néanmoins, certaines actions du NF peuvent déclencher de manière totalement automatique le déroulement d'autres actions internes au NF. Ces tâches, dites automatiques, sont alors invisibles pour l'utilisateur et ne servent par exemple qu'à garantir l'intégrité des données stockées dans le NF. Le CD n'est pas informé de ces actions. De plus dans le cas de plusieurs NF, ces actions automatiques peuvent faire intervenir l'ANF si des résultats doivent être répercutés dans d'autres NF.

De manière symétrique, certaines entités du monde réel peuvent avoir une influence sur d'autres entités réelles, sans pour autant requérir les fonctionnalités offertes par le CD. Considérer le modèle PAC-Amodeus dans le cadre des systèmes mixtes, nous amène donc à mettre en évidence l'existence possible d'actions automatiques dans le pilier réel de l'arche.

Par exemple, dans le cas du système *Karma* [Feiner 93], le NF contient les informations du manuel de maintenance de l'imprimante, en particulier les différentes étapes à suivre pour pallier un problème précis. Considérons le cas où l'information "Retirer le bac 1" du NF doit être rendu perceptible par l'utilisateur. L'ANF encode les étapes à réaliser dans un format compréhensible par le CD : par exemple les mots clés "action", "tiroir 1" et "retirer". Le CD dirige l'information vers le CTP correspondant au casque semi-transparent dans lequel ces informations doivent être affichées. D'autres CTP auraient pu être sollicités comme par exemple celui associé à un casque audio si les informations étaient présentées oralement. Le CTP, lié au casque semi-transparent, reçoit l'information. De cette information "normalisée", il en déduit la manière avec laquelle il doit la rendre à l'utilisateur. Il peut s'agir d'un affichage textuel ou d'un affichage 3D graphique, aligné avec le monde réel. Dans ce dernier cas, il est nécessaire d'obtenir des informations de localisation de l'imprimante par rapport au technicien. Pour cela, un adaptateur au sens d'ASUR est utilisé afin de connaître en temps réel la position de l'imprimante ainsi que celle du technicien par rapport à une référence fixe. Le CIBN du localisateur traduit les informations fournies par le dispositif en coordonnées  $x,y,z$  (position) et  $\alpha, \beta, \gamma$  (orientation) et les transmet à son propre CTP. Le CD n'a pas besoin d'être informé de ces valeurs et elles sont donc directement communiquées au CTP du casque semi-transparent. Ce dernier peut alors calculer les objets d'interaction que le CIBN va rendre perceptibles. Cette adaptation est représentée dans la Figure 46 par les flèches horizontales entre les CTP. Tant que l'utilisateur ne spécifie aucune autre action, la mise à jour du rendu graphique en fonction des positions et orientations est la seule action réalisée par le système interactif. Le CTP du localisateur va donc communiquer au CTP du casque semi-transparent tout changement dans l'une de ces données. Il s'agit donc d'une action informatique automatique liée à des actions dans le monde réel et dont la réalisation demeure dans le pilier réel de l'arche du modèle PAC-Amodeus.

## 6 Conclusion

Dans ce chapitre, nous avons proposé une démarche d'analyse ergonomique reposant sur quatre nouvelles propriétés ergonomiques et notre notation ASUR. Après la conception ergonomique, nous avons traité la conception logicielle des systèmes mixtes en proposant un modèle d'architecture logicielle.

Conception ergonomique et conception logicielle de systèmes mixtes se rejoignent. En effet, l'identification d'un problème d'incompatibilité ou de discontinuité perceptuelle, est synonyme d'environnements perceptifs non superposables. Or les environnements perceptifs liés aux entités réelles ne sont pas modifiables et sont des requis à la conception. Il convient donc de modifier les adaptateurs en cause de façon à rendre compatibles les environnements perceptifs requis par le système. Or chaque adaptateur constitue la base d'une branche du pilier réel du modèle d'architecture. Par conséquent, si l'adaptateur est modifié ou retiré au profit d'un autre, les éléments logiciels à modifier ou à retirer sont clairement identifiés puisqu'ils appartiennent au CIBN d'une même branche. **PAC-Amodeus adapté est donc un modèle favorisant une forte modifiabilité du code pour les adaptateurs ASUR.**

Considérons maintenant un problème ergonomique au niveau cognitif. Cela signifie que les référentiels de présentation ou les langages utilisés ont été identifiés comme facteurs d'incompatibilité ou de discontinuité cognitive. La forme des données véhiculées par les relations reliées aux adaptateurs du système doit être modifiée. D'après la définition des composants du modèle d'architecture, le CTP "exprime de manière abstraite le comportement des objets interactionnels". Les relations, pour lesquelles la forme des données doit être modifiée, sont reliées à des adaptateurs constituant les bases de différentes branches du pilier réel. C'est donc au niveau du CTP des branches associées aux adaptateurs en cause, que des éléments logiciels devront être changés. Là encore, la localisation des éléments logiciels à modifier est guidée par le lien existant entre les composants ASUR et les composants logiciels de notre modèle d'architecture logicielle. Comme les propriétés ergonomiques ont été traduites en termes d'ASUR dans le paragraphe 3, la localisation des modifications du code pour traiter un problème ergonomique est donc immédiat. **PAC-Amodeus adapté est donc un modèle favorisant une forte modifiabilité du code pour les langages associées aux relations ASUR.**

En conclusion, un lien clair est donc établi entre la description ASUR d'un système et notre modèle d'architecture : les adaptateurs ASUR d'un système sont présents au niveau du CIBN du modèle d'architecture, alors que les langages utilisés par les relations ASUR interviennent au niveau du CTP du modèle d'architecture. Comme nous l'avons montré, ceci permet d'envisager des modifications du code très ciblées qui favorisent donc la modifiabilité du code. De plus ce lien met en évidence l'indépendance du Contrôleur de Dialogue vis-à-vis des langages et des adaptateurs. Le niveau d'abstraction du CD est la tâche, et il est indépendant des technologies employées.

Les deux Chapitres suivants montrent l'application de nos résultats de conception ergonomique et de conception logicielle pour la mise en œuvre de notre système de chirurgie cardiaque assistée par ordinateur : *CASPER*. Les résultats de l'analyse

ergonomique conduite au Chapitre VI nous conduisent au chapitre VII à proposer des alternatives à la conception de *CASPER*. La réalisation logicielle de ces solutions illustre les éléments de réalisation logicielle, en particulier la mise en œuvre de notre modèle PAC-Amodeus adapté.

## 7 Références

[Abowd 92]

Abowd, D., Coutaz, J., Nigay, L., "*Structuring the Space of Interactive Properties*", Acte de la conférence EHCI'92, (1992), p. 113-128.

[Barnard 85]

Barnard, P., J., "*Cognitive Resources and the Learning of Computer Dialogs*", dans la revue *Interfacing Thought, Cognitive Aspects of Human Computer Interaction*, Carroll, (ed.) MIT Press, (1985), p. 112-158.

[Bernsen 94]

Bernsen, O., "*Foundations of multimodal representations. A taxonomy of representational modalities*", dans la revue *Interacting with Computers*, Vol. 6, 4, (1994), p.347-371.

[Dix 98]

Dix, A., Finlay, A., Abowd, G., Beale, R., "*Human-Computer Interaction*", 2nd Edition, Prentice Hall, (1998).

[Feiner 93]

Feiner, S., MacIntyre, B., Seligmann, D., "*Knowledge-Based Augmented Reality*", Communication de l'ACM n°7, (1993), p. 53-61.

[Gram 96]

Gram, C., Cockton, G., et al., "*Design Principles for Interactive Software*", Chapman et Hall, (1996), 248 pages.

[May 97]

May, J., Scott, S., Barnard, P., "*Modelling Multimodal Interaction: A Theory-based Technique for Design Analysis and Support*", Support de travaux dirigés (Tutorial) de la Conférence INTERACT'97, <http://www.shef.ac.uk/~pc1jm/guide.html>, (1997), 74 pages.

[McGurk 76]

McGurk, H., MacDonald, J., "*Hearing lips and seeing voices*", dans la revue *Nature*, n°264, (1976), p. 746-748.

[Merloz 97]

Merloz, P., Tonnetti, J., Eid, A., Faure, C., Lavalley, S., Troccaz, J., Sautot, P., Hammadeh, A., Cinquin, P., "*Computer Assisted Spine Surgery*", dans la revue *Clinical Orthopaedics and Related Research*, n°337, (1997), p. 86-96.

[Merloz 00]

Merloz, P., Lavalley, S., Tonnetti, J., Pittet, L., "*Image-Guided Spinal Surgery: Technology, Operative Technique and Clinical Practice*", dans la revue *Operative Techniques in Orthopaedics*, Vol 10, n°1, (2000), p. 56-63.

[Nigay 94]

Nigay, L., "*Conception et Modélisation logicielles des systèmes interactifs : applications aux interfaces multimodales*", Thèse de l'Université Joseph Fourier, Grenoble, (1994), 350 pages.

[Nigay 97]

Nigay, L., Coutaz, J., "*Software Architecture Modelling: Bridging Two Worlds Using Ergonomics and Software Properties*", dans Palanque, P., Paterno, F. (eds.) *Formal Methods in Human-Computer Interaction*, Springer Verlag, (1997), p. 49-73.

[Pfaff 85]

Pfaff, G., E., "*User Interface Management Systems*", Eurographics Seminars, Springer Verlag, (1985).

[Radermacher 96]

K. Radermacher, A. Zimolong, M. Anton, B. Thull, "*Recommendations for Ergonomic Analysis*", Deliverable 04 du projet IGOS, Lehrstuhl für Biomedizinische Technik der RWTH Aachen, (1996), 91 pages.

[Rau 96]

Rau, G., Radermacher, K., Thull, B., Pichler, C. v., "*Aspects of an Ergonomic System Design of a Medical Worksystem*", in: Taylor, R., Lavallée, St., Burdea, G., Moesges, R. (eds.): *Computer Integrated Surgery*, MIT-Press, Cambridge, MA, 1996, pp.203-221.

[Salber 95]

Salber, D., "*De l'interaction homme-machine individuelle aux systèmes multi-utilisateurs*", Thèse de doctorat de l'Université Joseph Fourier - Grenoble I, (1995), 305 pages.

[UIMS 92]

The UIMS Workshop Tool Developers : A Metamodel for the Runtime Architecture of an Interactive Systems, SIGCHI Bulletin, 24, 1, (1992), p. 32-37.

[Vernier 01]

Vernier, F., "*La Multimodalité en Sortie et son Application à la visualisation de grandes quantités d'Information*", Thèse de l'Université Joseph Fourier, Grenoble, (2001), 272 pages.



# CHAPITRE VI

## D'UN PROBLEME CLINIQUE A LA CONCEPTION D'UN SYSTEME : CASPER



## Sommaire

<b>Chapitre VI : D'un Problème Clinique à la Conception d'un Système :</b>	
<b>CASPER</b>	<b>157</b>
<b>1 Problème clinique</b>	<b>157</b>
<b>2 Conception et développement de la solution informatisée</b>	<b>160</b>
2.1 CASPER : Principe de fonctionnement et matériels utilisés	160
2.1.1 Principe de fonctionnement	160
2.1.2 Matériels utilisés	161
2.1.3 Etapes d'utilisation	163
2.2 Conception de l'interface homme-machine	166
2.3 Bilan	168
<b>3 Positionnement de CASPER dans notre cadre taxinomique</b>	<b>168</b>
<b>4 Analyse de l'interaction dans CASPER</b>	<b>169</b>
4.1 Description ASUR de CASPER	169
4.2 Caractérisation de l'interaction	170
4.2.1 Caractérisation des composants de l'interaction	171
4.2.2 Caractérisation des relations de l'interaction	171
4.3 Analyse ergonomique de l'interaction	173
<b>5 Conclusion</b>	<b>176</b>
<b>6 Références</b>	<b>178</b>

*CASPER* -Computer ASsisted PERicardial puncture- est un système de chirurgie assistée par ordinateur, développé par le laboratoire TIMC en collaboration avec le service de chirurgie cardiaque du CHU de La Tronche (Grenoble) et le laboratoire CLIPS pour les aspects propres à l'interface homme - machine. Ce système est utilisé pour assister le chirurgien lors de la réalisation d'une ponction péricardique. Nous exposons dans un premier temps le problème clinique et justifions ainsi l'apport d'une assistance informatisée à la réalisation de ce geste chirurgical courant.

## Chapitre VI : D'un Problème Clinique à la Conception d'un Système : CASPER

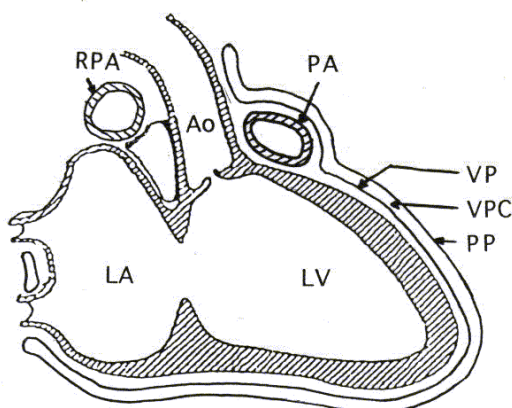
### 1 Problème clinique

Le péricarde est un sac fibro-séreux qui enveloppe le cœur et les gros vaisseaux à leur origine (aorte et artère pulmonaire). Il est composé de 2 portions : le péricarde fibreux et le péricarde séreux.

Le péricarde séreux est composé de 2 feuillets, les feuillets viscéral et pariétal, entre lesquels se trouve une cavité virtuelle, qui ne devient réelle qu'en cas d'épanchement. Le feuillet viscéral est moulé sur le cœur et les vaisseaux et on l'appelle épicaarde. Le feuillet pariétal recouvre l'épicaarde comme l'illustre le schéma de la Figure 47.

Le péricarde fibreux est un sac fibreux épais et hermétique, qui englobe le péricarde séreux. C'est sur lui que se fixent des éléments ligamentaires assurant la fixité du péricarde et du cœur.

Le péricarde évite la brusque dilatation des cavités cardiaques et facilite le remplissage auriculaire lors de la systole ventriculaire (contraction simultanée des deux ventricules pour chasser le sang dans les artères).



#### Légende :

LV : ventricule gauche  
VP : péricarde viscéral  
VPC : Cavité péricardique virtuelle  
PP : Péricarde pariétal  
PA : Artère pulmonaire  
LA : Auricule gauche  
Ao : Aorte  
RPA : Artère pulmonaire droite

Figure 47 Coupe du cœur.

Un épanchement péricardique se manifeste par la présence de liquide, suite à une hémorragie ou à une infection par exemple, à l'intérieur du péricarde et donc à proximité du cœur. La Figure 48 présente une échographie d'un cœur présentant un épanchement péricardique. La conséquence clinique majeure de la présence d'un épanchement péricardique est une augmentation des pressions de remplissage des cavités cardiaques

produisant, dans les cas extrêmes, une tamponnade (compression brutale du cœur, pouvant entraîner une mort subite). La tolérance à un épanchement dépend à la fois de son volume et de sa rapidité de constitution.

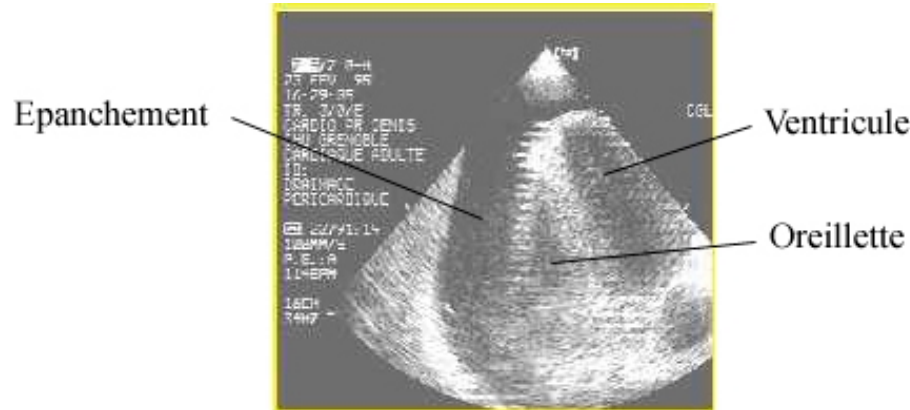


Figure 48 Echographie présentant un épanchement.

Pour évacuer ce liquide, une première solution requiert une chirurgie "lourde", c'est-à-dire nécessite l'ouverture du sternum du patient pour fournir au chirurgien un accès direct au cœur et donc à l'épanchement. Dans le cas d'un épanchement suffisamment important, une alternative à cette chirurgie très invasive et douloureuse pour le patient est la ponction péricardique. La ponction péricardique est le geste chirurgical qui consiste à évacuer ce liquide pathologique de façon minimalement invasive, c'est-à-dire de manière à réduire le plus possible les intrusions au sein du corps du patient. La réalisation de cette opération consiste à introduire précisément une aiguille de ponction au travers de la peau et à l'amener jusqu'à l'intérieur de l'épanchement. Il convient alors de drainer le liquide. La ponction péricardique [Chavanon 98] évite donc de réaliser une chirurgie thoracique lourde. Elle a été pratiquée dès 1840 à l'aveugle et par conséquent avec un fort taux de mortalité lié à la perforation de structures anatomiques voisines (comme le cœur, le foie ou encore les poumons) ou l'endommagement de vaisseaux sanguins principaux. A Grenoble, cette solution n'est envisagée que pour des épanchements dont la taille est supérieure à 15 millimètres. Dans les autres cas, une chirurgie "lourde" est menée.

Grâce à l'échographie, la sécurité de réalisation de la ponction péricardique a pu être améliorée. Dès 1973, un drainage péricardique sous contrôle échocardiographique était envisagé. Par la suite, plusieurs équipes ont eu recours à la ponction péricardique sous contrôle échographique bidimensionnel. Cette technique nécessite la présence de 2 médecins [Chavanon 97] :

- le cardiologue qui réalise l'échographie : dans un premier temps, juste avant la ponction, il repère et caractérise l'épanchement par voie sous-costale, c'est-à-dire qu'il place la sonde échographique de sorte que le plan de l'image échographique passe sous les côtes : la sonde est donc positionnée à la base du sternum. Dans un deuxième temps, le cardiologue transmet ces données au chirurgien.
- le chirurgien qui réalise la ponction : après avoir visualisé l'épanchement péricardique sur l'échographie faite par voie sous-costale, il réalise la ponction par voie sous-xiphœidienne. L'appendice xiphœïde est un petit os légèrement mobile à la base du sternum. La ponction est réalisée sous

contrôle échographique temps réel selon une voie apicale (verticale au cœur, c'est-à-dire perpendiculaire aux côtes).

Néanmoins, malgré l'assistance échographique, cette intervention reste dangereuse et difficile :

- le cœur étant placé dans la cage thoracique, divers organes et structures anatomiques, dont les côtes, peuvent gêner la trajectoire de l'aiguille.
- les risques de ponctions accidentelles d'organes restent élevés, notamment en ce qui concerne le myocarde, le poumon et le foie.
- la visualisation de l'aiguille de ponction sur l'image échographique n'est pas aisée : une sonde échographique ne réalise que des coupes : voir l'aiguille de ponction dans l'image échographique nécessite donc que le plan d'acquisition de la sonde échographique contienne l'aiguille. Dans ces conditions, il n'est pas non plus certain que l'extrémité de l'aiguille soit effectivement présente dans la coupe échographique : si l'orientation de l'aiguille est légèrement différente de l'orientation du plan de l'échographie, alors seule une partie de l'axe de l'aiguille est visible sur l'image.

Cette méthode de ponction avec assistance échographique est actuellement utilisée : le taux d'échec est évalué à 20%, toutes causes confondues. Un système permettant d'atteindre l'épanchement péricardique dans de meilleures conditions est donc nécessaire. Le système à mettre en œuvre doit en particulier traiter les points suivants :

- réduire fortement les risques de dommages causés aux organes voisins (cœur, foie, poumons),
- supprimer les risques de collision de l'aiguille avec une structure osseuse,
- permettre la ponction minimalement invasive de petits épanchements (< 15 mm) pour lesquels l'intervention invasive est la seule solution pratiquée actuellement.

Répondant à ces besoins, le système *CASPER* (Computer Assisted PERicardial puncture) a été conçu et développé. Ce système de ponction péricardique assistée par ordinateur permet de réaliser un geste peu invasif dans de bonnes conditions de précision et donc de sécurité. Il permet, en effet, la planification de la trajectoire d'une aiguille grâce à une série d'images échographiques, puis le guidage de la ponction de l'épanchement péricardique, c'est-à-dire le guidage en temps réel de l'introduction de l'aiguille de ponction. Une première version utilisant un écran a été réalisée grâce à une collaboration entre le TIMC et le CHU de Grenoble. Basée sur un écran, et le système de localisation *Optotrak*, cette première version répondait aux besoins en fournissant une interface de guidage minimale. Depuis septembre 96 cette collaboration pluridisciplinaire s'est étendue au laboratoire CLIPS, associant ainsi au projet des spécialistes en conception et développement d'interfaces homme-machine du laboratoire CLIPS. Dans la suite de ce chapitre, nous présentons les deux facettes de cette collaboration entre le TIMC, le CLIPS et le CHU : aussi, au paragraphe 2.1, nous expliquons les principes de fonctionnement et les dispositifs techniques utilisés, puis au paragraphe 2.2, nous exposons la conception de l'interface homme-machine.

## 2 Conception et développement de la solution informatisée

### 2.1 CASPER : Principe de fonctionnement et matériels utilisés

L'objectif de *CASPER* est d'optimiser la ponction péricardique en fournissant un guidage précis vers une cible déterminée. Ce guidage s'appuie sur l'utilisation d'un localisateur permettant de repérer la position d'un outil chirurgical en 3D. En regard de la taxonomie des systèmes de GMCAO présentée au chapitre I, *CASPER* est donc un système passif.

#### 2.1.1 Principe de fonctionnement

Les étapes de fonctionnement de *CASPER* correspondent à la boucle "Perception - Décision - Action", décrite dans le chapitre I. Nous détaillons ces trois étapes de *CASPER* :

- **Acquisition** et enregistrement préopératoire de données échographiques : cette phase s'effectue à l'aide d'une sonde d'échographie couplée à un système de localisation 3D (Polaris - Northern Digital Incorporated). Chaque image est alors localisée par rapport à une référence placée dans le bloc opératoire. Dans *CASPER*, préopératoire signifie "juste avant le geste", donc dans le bloc opératoire où aura lieu la ponction.
- **Segmentation** des structures anatomiques sur les images échographiques acquises précédemment : l'objectif est, lors de cette phase, de réaliser un modèle du cœur et de l'épanchement péricardique. Sur chaque image, le contour de l'épanchement est défini manuellement ou automatiquement. A chaque contour correspond donc un plan lié à une image échographique. Le calcul d'un plan médian permet ensuite de projeter tous les contours sur le même plan. L'intersection des zones segmentées de l'épanchement détermine alors une "zone stable" quelle que soit la position du cœur. Le chirurgien peut alors définir précisément sa trajectoire selon ce modèle.
- **Guidage** de l'outil en temps réel selon cette trajectoire prédéfinie : la position de l'aiguille est connue en permanence grâce à son couplage avec le localisateur. Cette stratégie implique toutefois l'immobilisation du patient tout le long de la procédure.

La difficulté particulière à *CASPER* par rapport à d'autres systèmes de GMCAO (en particulier en orthopédie) est due à la nature mobile de l'environnement et à la "déformabilité des tissus" mous. Or *CASPER* requiert, comme déjà mentionné, l'immobilité du patient entre le début de la phase d'acquisition des images échographiques et la fin de la ponction. Cependant, au cours de l'intervention, le cœur et donc l'épanchement sont soumis à des mouvements liés aux cycles cardiaque et pulmonaire, ainsi qu'à des mouvements moins prévisibles liés au patient :

- Les mouvements cardiaques étant cycliques, leur modélisation est envisageable, par le biais de l'acquisition d'un nombre suffisamment grand d'images échographiques du cœur. L'ensemble de ces images (environ 30) permet de supposer connue la position du cœur et de l'épanchement à chaque instant du cycle cardiaque.

- Les mouvements respiratoires peuvent être "maîtrisés" par un travail en apnée ; plusieurs apnées étant nécessaires, le repositionnement de la cage thoracique dans la position initiale est possible grâce à des capteurs.
- Les mouvements imprévisibles du patient (douleur, toux, etc.) sont réduits par une prémédication du patient et une anesthésie générale. Néanmoins, en cas de mouvement anormal au-dessus d'un seuil prédéfini par rapport au référentiel de base, une alarme se déclenche, grâce à des capteurs placés sur le thorax.

### 2.1.2 Matériels utilisés

*CASPER* repose sur la connaissance de la position en 3D des images échographiques acquises, de l'outil au cours de l'intervention et du patient, tous les calculs étant effectués par un ordinateur PC. Les aspects techniques et médicaux de *CASPER* sont décrits en détail dans [Chavanon 97] et [Barbe 97]. Le principe de localisation retenu est un système optique. Il se compose de deux caméras qui peuvent localiser en temps réel un nombre important de marqueurs : avec le système Polaris, ces marqueurs sont des diodes émettrices d'infrarouges ou des pastilles réfléchissant les infrarouges. Chaque élément à localiser (aiguille, sonde échographique, etc.) est accolé fermement à un groupe de marqueurs (entre 3 et 6 en général), qui forme un corps-rigide. La Figure 49 montre un foret équipé d'un corps-rigide muni de 3 pastilles réfléchissantes, ainsi que les deux caméras du système Polaris permettant de les localiser en temps réel dans un volume délimité.



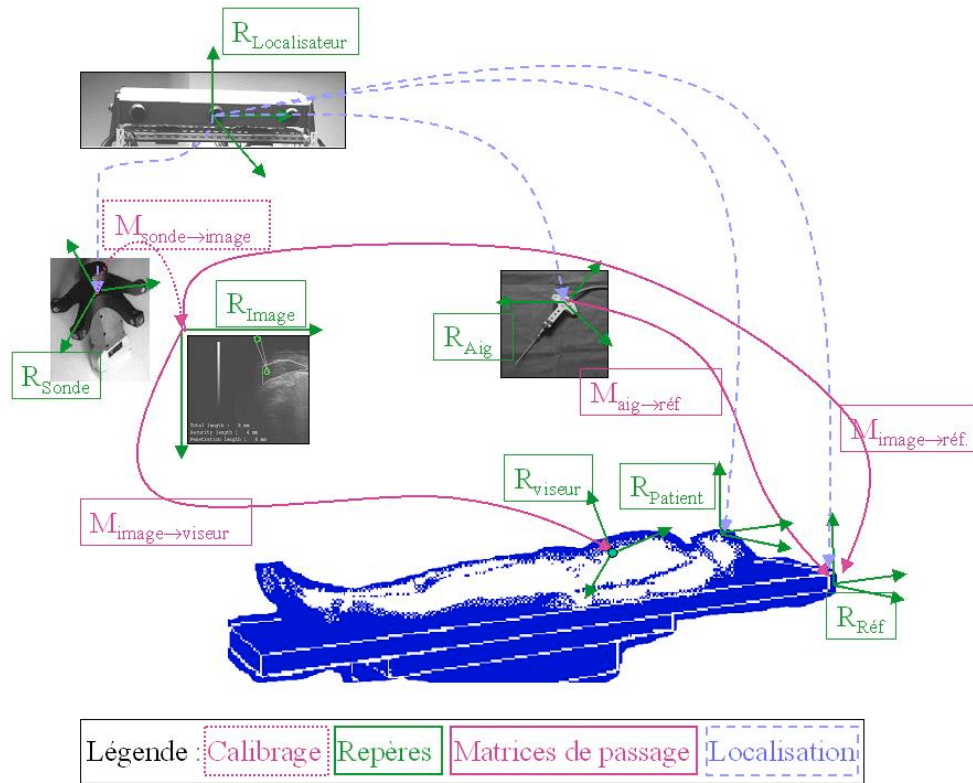
**Figure 49** Matériel de localisation utilisé dans *CASPER* : les caméras du Polaris (à gauche), un foret muni d'un corps-rigide doté de 3 pastilles réfléchissantes (à droite).

Un modèle du corps-rigide, connu du système informatique, permet lors de la phase de calibrage requise pour toute entité localisée, de déduire la position d'un point de l'objet accolé à ce corps-rigide (l'extrémité de l'aiguille par exemple) ou/et son orientation (plan d'acquisition de la sonde échographique par exemple), à partir de la position des diodes captée par le localisateur.

Dans *CASPER*, un premier corps-rigide fixé sur le patient permet de détecter d'éventuels mouvements du patient. Le repère associé au patient est noté  $R_{\text{Patient}}$  dans la Figure 50.

Lors de l'acquisition d'images échographiques, un corps-rigide est également fixé sur la sonde échographique. Le calibrage de la sonde échographique par rapport à son corps-rigide permet d'établir une matrice de changement de repère : les coordonnées des points de l'image exprimées dans le repère de l'image échographique, noté  $R_{\text{Image}}$ , peuvent donc être transformées en coordonnées exprimées dans le repère de référence (lié à la table

d'opération). Dans la Figure 50, nous notons cette matrice de changement de repère  $M_{\text{image} \rightarrow \text{réf}}$ .



**Figure 50** Repères et matrices de passage utilisés pour la localisation de l'image échographique, de l'aiguille, du patient et de la référence dans CASPER.

Lorsque le clinicien définit une trajectoire sur l'image échographique, un repère lui est associé : c'est le repère viseur noté  $R_{\text{viseur}}$ . Dans ce repère, l'axe  $z$  est confondu avec la trajectoire de telle sorte que le point de coordonnées  $(0,0,0)$  est le point cible qui doit être atteint par l'extrémité de l'aiguille. Les coordonnées de la trajectoire étant initialement exprimées dans l'image échographique, une matrice de passage du repère de l'image échographique  $R_{\text{Image}}$  vers le repère de la trajectoire  $R_{\text{viseur}}$  est établie : nous notons cette matrice de passage  $M_{\text{image} \rightarrow \text{viseur}}$ . Il est donc possible de transformer les coordonnées d'un point exprimées dans le repère de référence en coordonnées exprimées dans le repère du viseur : la relation permettant ce changement de repère est donnée par :

$$\begin{aligned} \mathbf{P}_{\text{viseur}} &= \mathbf{M}_{\text{image} \rightarrow \text{viseur}} \times \mathbf{M}_{\text{réf} \rightarrow \text{image}} \times \mathbf{P}_{\text{réf}}, \text{ ou encore} \\ \mathbf{P}_{\text{viseur}} &= \mathbf{M}_{\text{image} \rightarrow \text{viseur}} \times \mathbf{M}_{\text{image} \rightarrow \text{réf}}^{-1} \times \mathbf{P}_{\text{réf}} \quad (1) \end{aligned}$$

Cette relation est en particulier appliquée pour transformer les coordonnées des points, appartenant à l'aiguille chirurgicale, fournies par le localisateur. Chaque coordonnée est en effet exprimée dans le repère de référence et doit donc être exprimée dans le repère du viseur. Pour cela, un corps-rigide est fixé à l'aiguille de ponction ; une matrice de passage du repère associé au corps-rigide vers le repère de référence est construite par le système ( $M_{\text{aig} \rightarrow \text{réf}}$ ). Une procédure de calibrage établit, dans un premier temps, les coordonnées d'un point de l'axe de l'aiguille par rapport au corps-rigide ( $\text{Axe}_{\text{Aiguille}}$ ). Le même processus est repris afin d'établir les coordonnées de l'extrémité de l'aiguille dans le même repère, c'est-à-dire celui du corps-rigide de l'aiguille ( $\text{Extrémité}_{\text{Aiguille}}$ ). Lorsque le localisateur est activé, il détecte la position du corps-rigide de l'aiguille, de laquelle sont



déduites les coordonnées de l'extrémité et d'un point de l'axe de l'aiguille dans le repère de référence, en appliquant les relations suivantes :

$$\mathbf{Axe}_{\text{Réf}} = \mathbf{M}_{\text{aig} \rightarrow \text{réf}} \times \mathbf{Axe}_{\text{Aiguille}} \quad (2)$$

$$\mathbf{Extrémité}_{\text{Réf}} = \mathbf{M}_{\text{aig} \rightarrow \text{réf}} \times \mathbf{Extrémité}_{\text{Aiguille}} \quad (2')$$

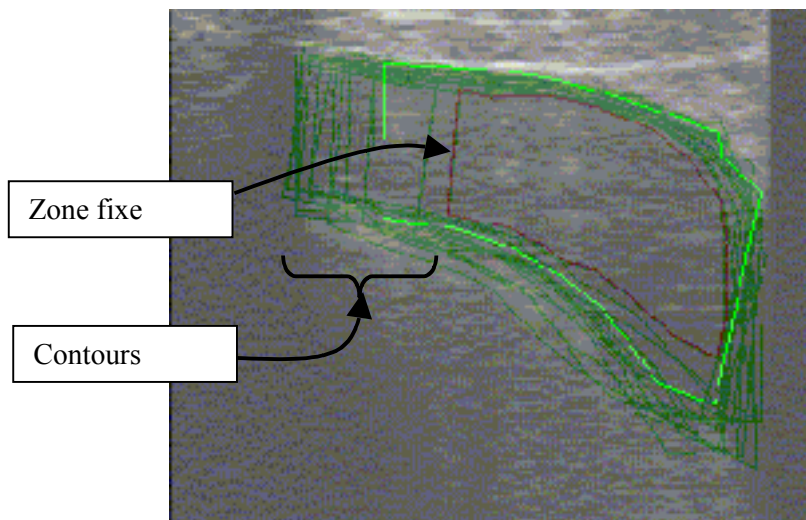
Une fois exprimées dans le repère de référence, les coordonnées des points sont converties dans le repère du viseur par la relation (1) et peuvent être utilisées pour afficher des données de guidage. La Figure 50 présente les repères et matrices de passage utilisés dans *CASPER*.

### 2.1.3 Etapes d'utilisation

Nous détaillons dans ce paragraphe les étapes d'utilisation de *CASPER*. Nous les présentons selon les trois étapes de fonctionnement présentées au paragraphe 2.1.1.

La première étape consiste en l'**acquisition** d'images échographiques du cœur, suffisamment nombreuses pour pouvoir prendre en compte l'ensemble des positions du cœur au cours du cycle cardiaque. Pour chaque acquisition, l'image ainsi que la position du plan d'acquisition dans le référentiel global sont stockées. La position des plans est connue grâce au suivi de la sonde échographique par un localisateur. A l'issue de la phase d'acquisition, un plan médian est calculé, et toute image dans un plan d'acquisition trop loin du plan médian est invalidée. Les données de guidage sont exprimées dans ce plan. Le chirurgien et le système doivent donc s'assurer que le nombre d'images valides restantes est suffisant.

Lors de la phase de **segmentation**, la zone d'intérêt est définie manuellement ou automatiquement sur chaque image échographique. En l'occurrence, il s'agit d'identifier le contour de l'épanchement péricardique. Une fois toutes les images segmentées, le système projette sur le plan médian les contours définis sur chaque image échographique et calcule l'intersection de ces zones. Cette intersection constitue la zone fixe. Cette zone est telle que, à chaque instant du cycle cardiaque, une partie de l'épanchement se trouve à cet endroit, repéré en 3D par rapport au référentiel global. Le chirurgien peut donc en toute sécurité amener l'extrémité d'une aiguille dans cette zone pour ponctionner l'épanchement. La Figure 51 illustre le tracé sur le plan médian des contours (en vert) définis sur chaque image, ainsi que de la zone fixe (en rouge).



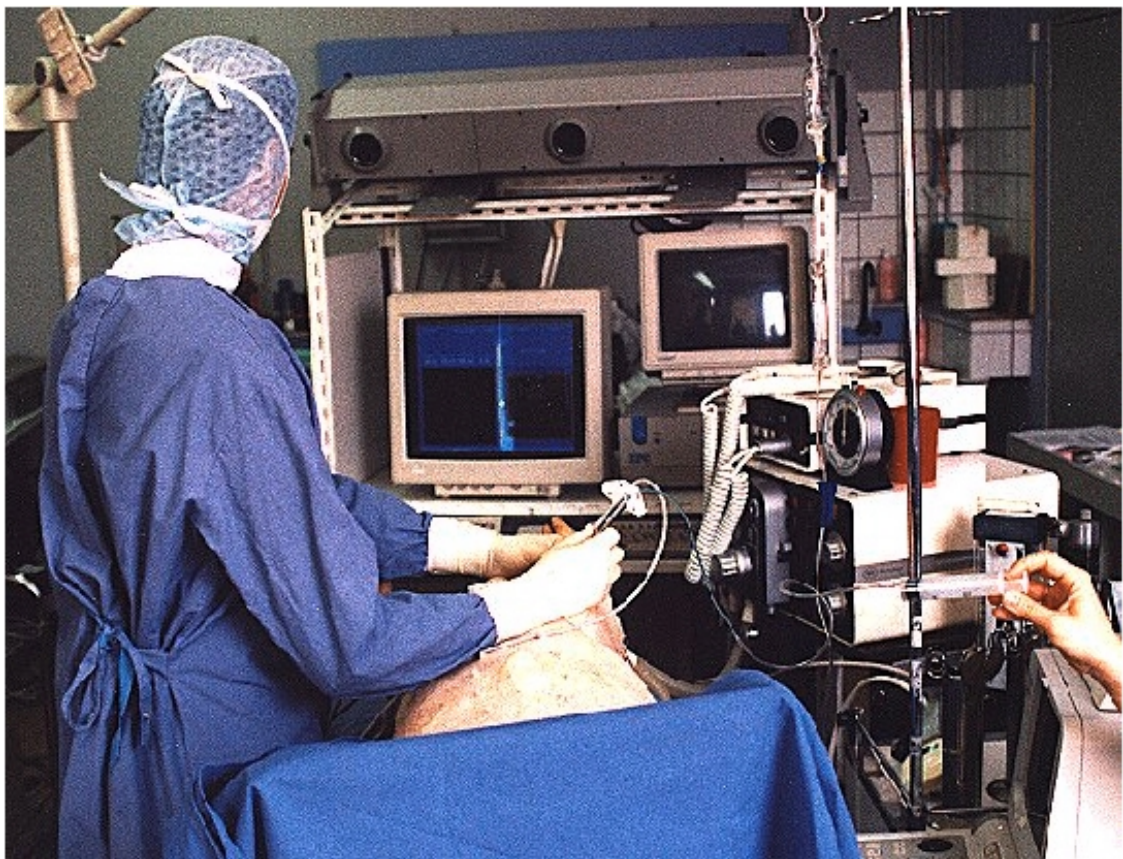
**Figure 51** Affichage des contours (verts) et de la zone fixe (rouge), issus de la phase de segmentation.



L'étape suivante est l'étape de **définition de la trajectoire**. Celle-ci est complètement définie en choisissant deux points sur l'image échographique du plan médian :

- l'un à l'intérieur de la zone fixe représentant la position souhaitable de l'extrémité de l'aiguille au moment de la ponction. Ce point constitue la cible à atteindre.
- le second situé à l'extérieur de la zone fixe, représentant la position souhaitée par le chirurgien pour un point de l'axe de l'aiguille au moment de la ponction. Ce deuxième point définit l'orientation de la trajectoire à reproduire.

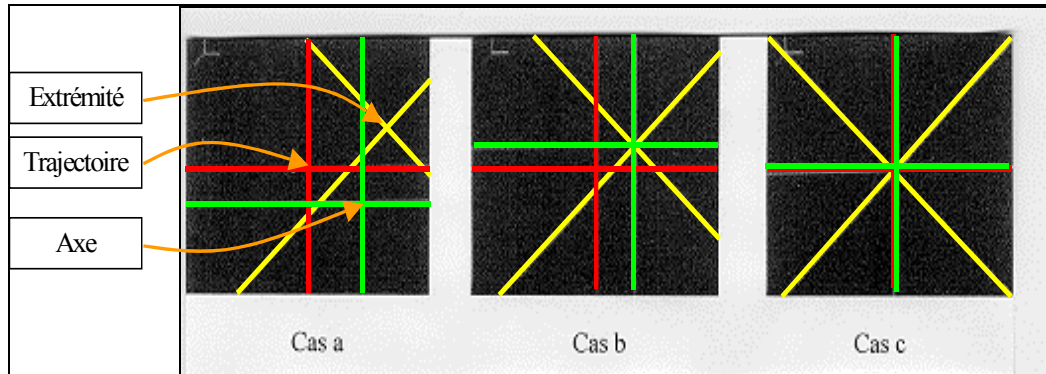
La dernière étape, celle de **guidage**, présente alors simultanément deux informations au chirurgien : la position courante de son aiguille ainsi que la trajectoire planifiée. La Figure 52 montre la configuration matérielle du bloc opératoire lors de l'utilisation de CASPER au moment de la réalisation de la phase de guidage.



**Figure 52** Organisation du bloc opératoire lors de l'utilisation de CASPER en phase de guidage.

Comme lors de la phase d'acquisition, un localisateur fournit en temps réel la position de l'aiguille chirurgicale en 3D. Cette position est donnée par rapport à la même référence que celle utilisée lors de l'acquisition. Dans cette version de CASPER, un système de viseur, combiné à une jauge, est utilisé pour fournir les données de guidage à l'écran. Dans le viseur, une croix rouge représente la trajectoire à reproduire : elle reste statique durant toute l'intervention. A cette croix s'ajoutent deux autres croix, l'une, jaune, représentant la position de l'extrémité de l'aiguille par rapport à la trajectoire, l'autre, verte, représentant la position d'un point de l'axe de l'aiguille. La position de ces trois points est présentée à l'utilisateur dans une fenêtre graphique 2D, représentant un plan

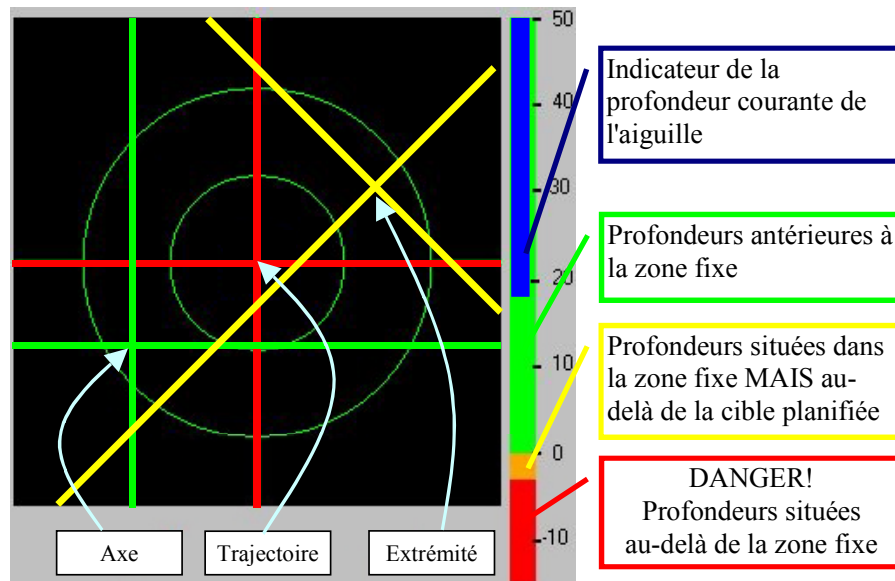
perpendiculaire à la trajectoire planifiée, avec en son centre la croix représentative de la trajectoire. Les deux autres croix sont affichées par rapport au centre. Si la croix jaune est au-dessus de la croix rouge, le chirurgien doit bouger l'aiguille de telle sorte qu'il abaisse l'extrémité de l'aiguille. La Figure 53 illustre différents cas de fonctionnement de ce viseur.



**Figure 53** Représentation courante de la position de l'aiguille par rapport à la trajectoire planifiée dans CASPER. **Cas a)** : trajectoires différentes, **cas b)** : trajectoires parallèles, **cas c)** : trajectoires identiques.

Dans le premier cas (cas a), l'extrémité de l'aiguille est au-dessus et trop à droite de la trajectoire planifiée. Dans le même temps, la croix représentant un point de l'axe de l'aiguille est au-dessous et trop à droite de la trajectoire planifiée. La trajectoire courante est donc différente de la trajectoire planifiée : le chirurgien doit déplacer son aiguille vers la gauche, abaisser l'extrémité de l'aiguille et au contraire légèrement relever l'axe. Dans le cas b, les croix représentant l'extrémité de l'aiguille et un point de l'axe sont superposées. Par contre elles sont au-dessus et à droite de la trajectoire planifiée. Cela signifie que l'aiguille est en train de suivre une trajectoire parallèle à celle planifiée par le chirurgien : pour corriger la position de l'aiguille, il suffit donc au chirurgien de descendre l'aiguille et de la déplacer sur la gauche. Enfin le cas c symbolise un alignement parfait de l'aiguille avec la trajectoire planifiée. C'est la configuration recherchée par le chirurgien avant de faire pénétrer l'aiguille à l'intérieur du corps du patient.

Lors de l'avancée de l'aiguille dans le corps du patient, le chirurgien cherche à maintenir au mieux l'alignement de ces croix, et avance l'aiguille jusqu'à ce que la jauge, à droite dans la Figure 54, indique que l'extrémité de l'aiguille (curseur bleu) se trouve dans la zone fixe (partie orange de la jauge).



**Figure 54** Données de guidage présentées par CASPER :  
le viseur (à gauche), la jauge de profondeur (à droite).

Ces données graphiques de guidage sont complétées par des données numériques indiquant la distance en millimètres entre l'extrémité de l'aiguille et la cible visée, la distance de l'aiguille à la trajectoire planifiée et une indication angulaire indiquant la déviation entre la trajectoire suivie et la trajectoire à réaliser. L'image échographique du plan médian est également affichée à côté du viseur et de la jauge de profondeur. La position de l'extrémité de l'aiguille est superposée à cette image échographique de contrôle.

## 2.2 Conception de l'interface homme-machine

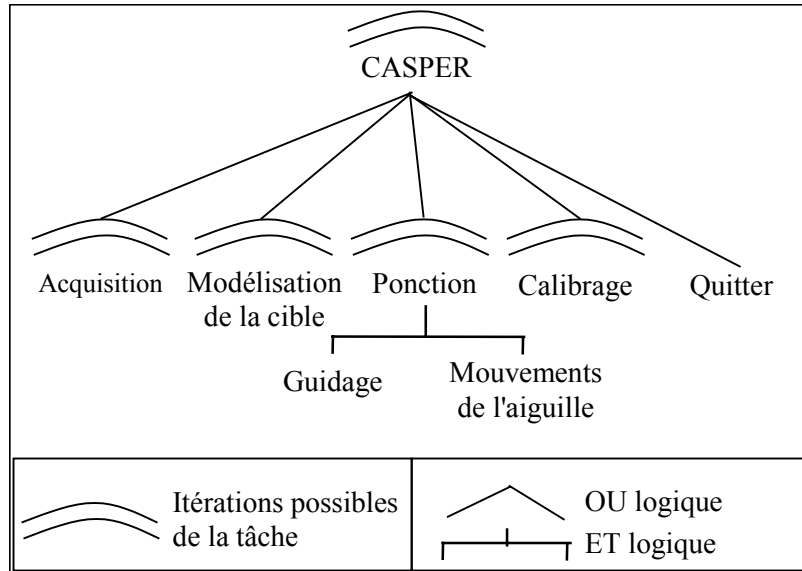
Pour la conception de l'interface de *CASPER*, nous avons appliqué la méthode présentée au Chapitre V qui se compose de six étapes : esquisse du problème à résoudre, modélisation de l'utilisateur, modélisation de la tâche et des objets spécifiques au domaine, c'est-à-dire identification des sous-tâches et de leur ordonnancement, définition des concepts informatiques manipulés par le système, définition de la présentation de l'interface, c'est-à-dire des fenêtres de l'application et de l'organisation des données et fonctionnalités regroupées dans chacune de ces fenêtres, et enfin évaluation expérimentale.

L'analyse du problème et la modélisation des utilisateurs ont été réalisées lors de la définition du cahier des charges du système *CASPER*. Nous développons ici l'analyse de la tâche conduite. Celle-ci repose sur des observations directes en bloc opératoire, que nous avons complétées par des entretiens avec le chirurgien. Ces observations et entretiens ont mis en évidence deux solutions possibles :

- soit l'interface est peu directive et très souple d'utilisation : l'arbre de tâches est alors minimal.
- soit au contraire l'interface impose des barrières lors de l'utilisation : l'arbre de tâches est alors effectif et correspond à une logique d'utilisation.

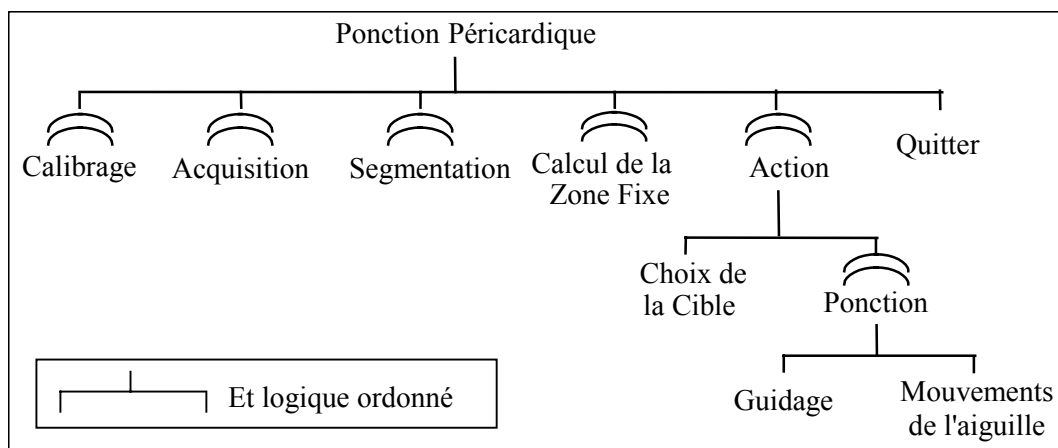
En utilisant la notation HTA ("*Hierarchical Task Analysis*"), l'arbre de tâche minimal correspondant à une ponction péricardique assistée par ordinateur est représenté à la

Figure 55 : chaque étape composant l'application peut être réalisée à tout moment au cours de la réalisation de l'intervention. Cet arbre n'implique aucun ordre dans les actions à entreprendre. Le chirurgien n'étant pas expert dans la manipulation d'un système informatique d'une part, et compte tenu de la criticité de la tâche d'autre part, nous avons choisi de concevoir une interface plus directive qui met en œuvre un arbre de tâches effectif.



**Figure 55** Arbre de tâches minimal d'une ponction assistée par ordinateur, décrit avec la notation HTA.

L'arbre de tâches effectif, de la Figure 56, impose un ordre dans l'utilisation des différentes étapes requises pour effectuer une ponction péricardique assistée par ordinateur. Ainsi en phase de guidage, la seule autre tâche que le chirurgien puisse réaliser est la modification de la cible. Il n'est en effet pas utile de permettre une nouvelle segmentation des images par exemple, à ce niveau de la réalisation de la tâche. Les différentes barrières mises en œuvre par cet arbre de tâches ont pour objectif de mieux guider le chirurgien dans la réalisation de sa tâche, mais aussi d'éviter que le chirurgien ne commette des erreurs de manipulation du système. Par exemple le lancement du calibrage des outils au moment du guidage invalide les étapes d'acquisition, de segmentation et de définition de la cible.



**Figure 56** Arbre de tâches effectif d'une ponction assistée par ordinateur, décrit avec la notation HTA.

L'arbre de tâches effectif ayant été validé par le chirurgien, nous avons ensuite défini la présentation des objets informatiques manipulés par les tâches. L'arbre de tâches nous a permis d'identifier la nécessité d'affichage de certaines données lors de phases particulières. Par exemple avant de segmenter les images, donc à l'issue de la phase d'acquisition, il est nécessaire de connaître le nombre d'images valides avant de quitter cette phase d'acquisition. Par ailleurs, nous avons privilégié la mise en évidence des barrières et de messages d'aides au chirurgien. Les barrières permettent de mettre en œuvre l'arbre de tâches effectif et donc constituent une aide à la "navigation" dans l'interface, tandis que les messages d'aide assistent le chirurgien dans la réalisation d'une tâche particulière. La définition de la présentation est plus largement décrite dans [Dubois 97].

### 2.3 Bilan

Cette version de *CASPER* est en phase de validation clinique. Après les étapes de validation sur modèles physiques puis sur cadavres, *CASPER* a été validé cliniquement sur animaux. Le protocole expérimental de validation clinique, détaillé dans [Chavanon 00], a permis de réaliser 10 ponctions péricardiques pour lesquelles la taille du plus petit épanchement était de 7,5 mm et celle du plus large était de 12.6 mm. Dans les 10 cas, l'intégralité de la procédure de ponction péricardique assistée par ordinateur a nécessité entre 10 et 15 minutes. Pour chacun des cas, le liquide de l'épanchement obtenu était sans présence de sang (ce qui aurait été synonyme de perforation de la paroi du cœur). Sans assistance informatique, la ponction de petits épanchements requiert la mise en œuvre d'une chirurgie cardiaque ouverte, basée sur l'ouverture du thorax, une thoracotomie, donc très invasive pour le patient. L'intérêt clinique a donc été démontré, ainsi que sa sécurité puisque au cours des 10 ponctions aucune n'a donné lieu à la perforation du cœur ou d'un organe voisin. L'étape suivante de la validation clinique consiste à utiliser *CASPER* sur des patients. Cette étape est en cours de réalisation.

Toutefois, si les résultats cliniques sont très satisfaisants, le fonctionnement de *CASPER* et sa mise en place requièrent la présence d'un ingénieur informaticien en plus de l'équipe chirurgicale. Les deux critères d'utilisabilité définis par la FDA efficacité ("*effectiveness*") et sécurité ("*safety*") ont donc été évalués, mais les critères de "facilité d'utilisation" ("*ease of use*") et d'"acceptation par l'utilisateur" ("*acceptance by the user*") n'ont pas été évalués. Néanmoins, l'acceptation du système par l'utilisateur est fortement favorisée par l'implication de l'utilisateur dans la conception : validation de l'arbre de tâches et de la présentation (principe de la conception participative [Reich 96]). Pour étudier le critère de "facilité d'utilisation", nous présentons aux paragraphes suivants la caractérisation et l'analyse ergonomique de *CASPER*, lors de sa phase la plus critique : le guidage. Pour cela, nous utilisons notre cadre de classification présenté au Chapitre III et notre approche de conception développée aux chapitres IV et V.

## 3 Positionnement de *CASPER* dans notre cadre taxinomique

*CASPER* ayant pour objectif d'aider à réaliser la ponction d'un épanchement péricardique sur un patient réel, le patient apparaît clairement comme étant l'entité sans laquelle la tâche n'a aucun sens. Par conséquent l'objet de la tâche est réel et *CASPER* fait donc partie des systèmes de Réalité Augmentée, comme définis dans le Chapitre III.

Le second axe de notre classification vise à caractériser l'apport du système informatique dans la réalisation d'une tâche avec le système mixte. La tâche que nous avons choisie d'étudier est celle de ponction, au cours de laquelle le chirurgien doit introduire une aiguille dans le corps du patient jusqu'au centre de l'épanchement. Par rapport à une intervention conventionnelle, le chirurgien perçoit en plus la position de son aiguille par rapport à la trajectoire planifiée, affichée sur l'écran. Par conséquent *CASPER*, lors de la réalisation de la ponction, augmente la phase d'évaluation de la réalisation de la tâche.

En regard de notre cadre taxinomique, *CASPER* est donc un système de Réalité Augmentée pour lequel la phase de réalisation de la tâche qui est augmentée est l'évaluation. En nous référant à la synthèse des systèmes mixtes existants, présentée dans la Figure 31 et la Figure 32 du Chapitre III, *CASPER* appartient à la même classe de systèmes que la *visée pédiculaire* et les systèmes basés sur le principe d'*Image Overlay*. *CASPER* présente également les mêmes caractéristiques interactionnelles que *KARMA* ou le *Musée Augmenté* qui eux mettent en œuvre un casque semi-transparent. Enfin, *Audio Notebook* et le système chirurgical de Wegner, basés sur l'utilisation de données sonores répondent aussi à la même caractérisation de la tâche.

Basé sur notre espace taxinomique, nous déduisons donc par comparaison avec des systèmes mixtes existants, que *CASPER* pourrait utiliser des données autres que graphiques pour augmenter l'évaluation lors de la réalisation de la ponction, comme des données haptiques ou sonores. Enfin l'usage d'un casque semi-transparent est une autre solution technique à envisager dans *CASPER*, pour remplacer l'écran de l'ordinateur.

## 4 Analyse de l'interaction dans *CASPER*

Notre analyse ergonomique de l'interaction dans *CASPER* repose sur la notation ASUR présentée au Chapitre IV et notre démarche d'analyse exposée au Chapitre V. Nous décrivons et caractérisons l'interaction dans *CASPER* avec notre notation ASUR aux paragraphes suivants 4.1 et 4.2, puis nous l'analysons selon notre démarche au paragraphe 4.3.

### 4.1 Description ASUR de *CASPER*

Nous identifions d'abord les composants du système *CASPER*.

Appartenant au monde réel, l'objet de la tâche (composant  $R_{\text{task}}$ ) est le patient pour lequel est réalisée l'intervention. Le chirurgien (composant U) est l'utilisateur du système, responsable de mener l'aiguille au point de ponction défini dans une phase préalable et selon la trajectoire pré-déterminée. Au cours de l'utilisation du système, l'aiguille de ponction est un objet du monde réel utilisé comme outil : il s'agit donc d'un composant de type  $R_{\text{tool}}$ .

Du côté du monde informatique ou virtuel mis en jeu dans *CASPER*, un ordinateur maintient les images médicales et les coordonnées de la trajectoire à reproduire et est doté de capacités de calculs (changements de repères par exemple) : c'est donc le système au sens d'ASUR (composant S). De plus, un écran est utilisé pour afficher les données : cet écran permet de transférer des données du monde virtuel vers le monde réel et constitue à ce titre un adaptateur de sortie (composant  $A_{\text{out}}$ ). A l'inverse, le localisateur, composé des



caméras infrarouges et des corps-rigides fixés sur plusieurs objets réels, permet de transférer au système informatique la position des objets réels dans le champ opératoire. Il s'agit donc d'un adaptateur d'entrée (composant  $A_{in}$ ).

CASPER comprend donc six composants ASUR comme le montre la Figure 57:

- un composant  $R_{task}$ , le patient,
- un composant  $R_{tool}$ , l'aiguille de ponction,
- un composant  $U$ , le chirurgien, réalisant la ponction péricardique,
- un composant  $S$ , l'ordinateur,
- un composant  $A_{in}$ , le localisateur optique,
- un composant  $A_{out}$ , l'écran.

Examinons maintenant les relations mises en jeu, c'est-à-dire, les couples de composants qui échangent des données au cours de la réalisation de la ponction.

Pendant l'opération, le chirurgien, composant  $U$ , déplace l'aiguille, composant  $R_{tool}$ , et peut l'observer, ce qui est décrit par la double relation  $U \leftrightarrow R_{tool}$ . La position de l'aiguille est repérée dans l'espace par le localisateur, composant  $A_{in}$ , ( $R_{tool} \rightarrow A_{in}$ ), qui transmet la position de l'outil au système informatique ( $A_{in} \rightarrow S$ ). Le système informatique traite cette donnée et en définit la forme graphique en vue de l'affichage à l'écran, composant  $A_{out}$ , ( $S \rightarrow A_{out}$ ). L'affichage peut alors être perçu par l'utilisateur ( $A_{out} \rightarrow U$ ). Enfin, en permanence le chirurgien peut percevoir le patient ( $R_{task} \rightarrow U$ ) et l'aiguille de ponction est au contact du patient, ce est représenté dans la description ASUR de CASPER à la Figure 57 par la relation mécanique  $R_{tool} \text{---} R_{task}$ .

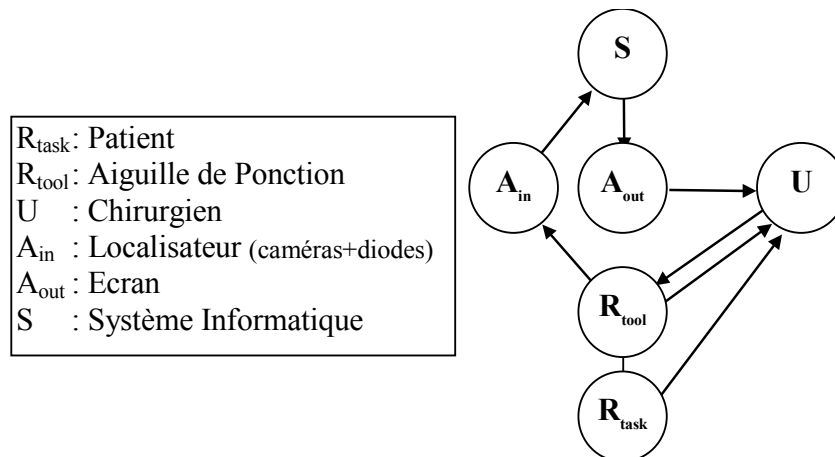


Figure 57 Description avec ASUR de la tâche de ponction réalisée avec CASPER.

Pour compléter la description ASUR de la tâche de ponction avec CASPER afin d'en étudier l'interaction, il convient de caractériser les composants et relations en contact avec l'utilisateur.

## 4.2 Caractérisation de l'interaction

Caractériser l'interaction induite par un système mixte nécessite l'étude des caractéristiques des différents éléments constituant l'interaction, que nous avons présentées au chapitre IV :

- les composants de l'interaction, c'est-à-dire les adaptateurs et objets du monde réel connectés à l'utilisateur par des relations ASUR,
- les différentes relations connectées à l'utilisateur.

### 4.2.1 Caractérisation des composants de l'interaction

Pour un composant ASUR, nous avons identifié trois caractéristiques : le sens perceptif et actionnel, la localisation et le niveau de partage. *CASPER* met trois composants en jeu :

- le patient, objet de la tâche (composant  $R_{task}$ )
- l'aiguille chirurgicale, outil utilisé pour la ponction (composant  $R_{tool}$ )
- l'écran, adaptateur de sortie (composant  $A_{out}$ )

Ces trois composants utilisent la vue comme sens perceptif requis pour transmettre des données au chirurgien. L'aiguille est manipulée par le chirurgien : le sens actionnel requis est donc composé d'actions physiques. En conséquence des mouvements appliqués à l'aiguille, le chirurgien perçoit aussi des données tactiles.

Le patient et l'aiguille sont perceptibles au niveau du champ opératoire. Par contre, compte tenu des contraintes propres au bloc opératoire (stérilité, espace réduit, etc.) l'écran est nécessairement distant du champ opératoire. Par conséquent, le lieu de perception des données affichées à l'écran est distant du champ opératoire. Quant au lieu d'action, il est localisé au niveau du champ opératoire.

Enfin, le niveau de partage est grand pour les données fournies par tous les composants de l'interaction. Par contre, il est limité à un seul utilisateur en ce qui concerne les mouvements appliqués à l'aiguille. En effet il n'est pas envisageable que deux personnes ou plus puissent agir simultanément sur l'aiguille pour la déplacer.

Le tableau de la Figure 58 résume ces caractéristiques.

	Sens Perceptif et Actionnel	Lieu de perception et d'actions	Niveau de Partage
<b>Patient</b>	Vue	Champ opératoire	N
<b>Aiguille de ponction</b>	Vue	Champ opératoire	N
	Action physique/haptique	Champ opératoire	1
<b>Ecran</b>	Vue	Ecran	N

**Figure 58** Caractérisation des composants participant à l'interaction dans *CASPER*.

### 4.2.2 Caractérisation des relations de l'interaction

Pour une relation ASUR connectée à l'utilisateur et donc participant à l'interaction, nous avons retenu trois caractéristiques également : le langage utilisé, le poids de la relation en regard de la tâche et le référentiel de présentation des données véhiculées. Lors de la réalisation de la tâche de ponction avec *CASPER*, les quatre relations concernées sont celles permettant :

- la perception des données de guidage affichées à l'écran,  $A_{out} \rightarrow U$
- l'observation de l'aiguille de ponction,  $R_{tool} \rightarrow U$
- les déplacements appliqués à l'aiguille de ponction,  $U \rightarrow R_{tool}$
- la visualisation du patient,  $R_{task} \rightarrow U$

L'échange de données de l'écran vers l'utilisateur est dans cette version de *CASPER* basé sur l'affichage :

- du viseur, défini par un langage 2D, arbitraire, non analogique, non linguistique et dynamique,



- de la jauge de profondeur, indiquant la distance entre l'extrémité de l'aiguille et la cible, définie par un langage 1D, non arbitraire, analogique, non linguistique, et dynamique,
- de valeurs numériques représentatives de l'écart séparant l'aiguille de la trajectoire, expressions d'un langage textuel, non arbitraire, non analogique, linguistique et dynamique.
- de l'image échographique, qui est une coupe 2D, non arbitraire, analogique, non linguistique et statique.

La perception de l'aiguille réelle ainsi que du patient relève de la perception du monde réel. Nous caractérisons donc ces deux relations comme basées sur un langage 3D (perception 3D du monde réel). Le chirurgien manipule l'aiguille : les actions physiques requises s'expriment selon un langage 3D. Enfin, la manipulation de l'aiguille entraîne nécessairement la perception par le clinicien de sensations tactiles, au travers de l'aiguille. Ces sensations expriment en particulier la résistance des tissus traversés et s'expriment également selon les trois dimensions, même si la dimension située dans l'axe de l'aiguille est privilégiée. Ces premières caractéristiques sont résumées dans la première colonne du tableau de la Figure 59.

	Langage	Poids	Référentiel de présentation
$A_{out} \rightarrow U$	2D, arb., non anal., non ling., dyn	Elevé	Trajectoire
Viseur	1D, non arb., anal., non ling., dyn.	Elevé	Trajectoire
Jauge	Textuel, non arb., non anal., ling., dyn.	Elevé	Trajectoire
Valeurs Num.	2D, non arb., anal., non ling., statique	Faible	Sonde
Echographie			
$R_{tool} \rightarrow U$	3D, réel	Elevé	Utilisateur
$U \rightarrow R_{tool}$	3D, réel	Elevé	Utilisateur
$R_{task} \rightarrow U$	3D, réel	Elevé	Utilisateur

**Figure 59** Caractérisation des relations participant à l'interaction dans CASPER.

Parmi les relations constituant l'interaction, toutes sont relatives à la tâche et importantes pour la réalisation de celle-ci. Cependant l'image échographique reçoit une attention moins grande. La composante de la relation portant les données de l'image échographique a donc un poids faible alors que les autres ont toutes un poids élevé.

Enfin, les données de guidage fournies par le viseur, la jauge et par les valeurs numériques sont présentées selon le référentiel de présentation lié à la trajectoire planifiée. A l'opposé, les autres relations reliées à l'utilisateur sont porteuses de données exprimées selon un référentiel utilisateur.

Le tableau de la Figure 59 résume l'analyse des caractéristiques de ces relations.

La description ASUR de CASPER met en évidence les composants participant à l'interaction mais aussi les différents échanges de données entre ces composants. Les

caractéristiques de ces composants et de ces relations affinent cette description. Ainsi nous obtenons, pour la tâche de ponction avec *CASPER*, une description des différentes facettes de l'interaction du chirurgien avec le système mixte *CASPER*. Sur la base de cette description, il est alors possible de conduire une analyse ergonomique de l'interaction du chirurgien avec le système. Cette analyse est l'objet du paragraphe suivant.

### 4.3 Analyse ergonomique de l'interaction

En nous basant sur notre démarche d'analyse présentée au chapitre V, nous étudions ici les propriétés de compatibilité et de continuité de l'interaction en sortie dans *CASPER*. La première étape consiste à décrire le système avec notre notation ASUR. C'était l'objet du paragraphe précédent, aussi nous commençons l'analyse à partir du second point de notre démarche.

#### 2. Identifier les concepts principaux manipulés par le système.

Le patient et l'aiguille de ponction constituent les deux principaux concepts. La trajectoire à suivre, c'est-à-dire sa direction ainsi que la cible visée constitue également un concept manipulé et représenté par le système.

#### 3. Identifier les relations ASUR composant l'interaction de sortie du système.

La description ASUR de *CASPER*, présentée au paragraphe précédent, met en évidence trois relations ASUR constituant l'interaction en sortie. Il s'agit de la perception des données de guidage ( $A_{out} \rightarrow U$ ), du patient ( $R_{task} \rightarrow U$ ) et de l'outil chirurgical ( $R_{tool} \rightarrow U$ ).

#### 4. Caractériser les environnements perceptifs des composants ASUR à l'origine des relations identifiées à l'étape 3.

La caractérisation des composants participant à l'interaction, résumée à la Figure 58, souligne trois environnements perceptifs liés à l'adaptateur de sortie, à l'outil manipulé et au patient. Les données de guidage requièrent le couple (vue, écran), la perception du patient sollicite le couple (vue, champ opératoire) et l'aiguille de ponction implique l'environnement perceptif représenté par le couple (vue/toucher, champ opératoire) pour la perception visuelle et tactile de l'aiguille.

#### 5. Caractériser les langages mis en œuvre par les relations identifiées à l'étape 3.

La Figure 59 résume les différents langages mis en œuvre dans l'interaction en sortie du système. Quatre langages aux caractéristiques différentes sont utilisés pour afficher les données sur l'écran à destination du clinicien. Par contre, le même langage est mis en jeu lors de la perception de l'outil chirurgical et du patient, tous deux appartenant au monde réel.

#### 6. Analyser la compatibilité de l'interaction en sortie aux niveaux perceptuel et cognitif.

Pour étudier la compatibilité au niveau perceptuel de l'interaction en sortie, il convient d'analyser les environnements perceptifs requis par les relations identifiées à l'étape 4. Dans *CASPER*, l'analyse de tous les environnements perceptifs mis en jeu par les composants à l'origine de l'interaction en sortie nous conduit à identifier deux lieux géographiques différents de sources de données relatives à la tâche : l'écran et le champ opératoire où se trouvent le

patient et l'aiguille de ponction. La configuration matérielle actuelle ne permet pas d'envisager une superposition de ces deux environnements perceptifs. En effet dans un bloc opératoire, la place est relativement restreinte, et les contraintes de position des différents composants sont très fortes. Par conséquent il n'est pas envisageable de positionner l'écran actuel suffisamment près du champ opératoire afin de rendre compatible ces deux environnements perceptifs. CASPER présente donc une incompatibilité perceptuelle, liée à la localisation des données à percevoir.

L'étude de la compatibilité au niveau cognitif repose sur l'analyse des langages utilisés par les relations constituant l'interaction en sortie du système. La combinaison de langages de différentes dimensions (textuel, 1D, 2D et 3D), de types différents (arb. et non arb., anal. et non anal., ling. et non ling.) ainsi que la combinaison de différents référentiels de présentation, sont autant de facteurs propices à l'incompatibilité cognitive.

**7. Déduire parmi les relations identifiées à l'étape 3, celles relatives à un même concept : l'aiguille.**

Deux relations ASUR sont véhicules de données relatives à l'aiguille de ponction : les données de guidage fournies par le viseur et affichées à l'écran ( $A_{out} \rightarrow U$ ), ainsi que la perception directe qu'a le chirurgien de l'aiguille ( $R_{tool} \rightarrow U$ ).

**8. Analyser la continuité de l'interaction en sortie aux niveaux perceptuel et cognitif pour l'aiguille.**

Au niveau perceptuel, la Figure 58 met en évidence deux lieux de perception distincts pour les deux composants  $R_{tool}$  et  $A_{out}$ , sources des deux relations identifiées à l'étape précédente : respectivement le champ opératoire et l'écran. Nous identifions donc une **discontinuité perceptuelle** concernant la perception des données liées au concept d'aiguille. Cette discontinuité est liée à de multiples lieux de perception.

Au niveau cognitif, la Figure 59 met en évidence trois langages différents : celui utilisé pour présenter la trajectoire (viseur), celui utilisé pour présenter la cible (jauge) et celui que nous assimilons à la présentation du monde réel, en l'occurrence l'aiguille réelle. Ces trois langages ont des caractéristiques différentes tant au niveau de la dimension (respectivement 2D, 1D et 3D), que du type de la représentation associée (respectivement arb./non anal., non arb./anal. et non arb./anal.) ainsi que du référentiel de présentation (référentiel lié à la trajectoire pour le viseur et la jauge, référentiel lié à l'utilisateur pour la perception de l'aiguille réelle). Ces différences sont des facteurs propices à la **discontinuité cognitive** lors de l'interprétation des données perçues relatives au concept de l'aiguille.

**9. Recommencer à l'étape 7 pour un autre concept.**

Dans cette version de CASPER, les données concernant le patient sont également portées par deux relations. Par contre les données concernant la trajectoire et la cible sont entièrement exprimées par une seule relation ASUR : ( $A_{out} \rightarrow U$ ).

Pour la complétude de l'analyse, il convient donc d'appliquer les étapes 7 et 8 pour le patient.

### 10. Déduire parmi les relations identifiées à l'étape 3, celles relatives à un même concept : le patient

Les données relatives au patient proviennent de la visualisation directe du patient par le clinicien ( $R_{task} \rightarrow U$ ) mais aussi de l'écran, par le biais de l'image échographique affichée ( $A_{out} \rightarrow U$ ), et de la perception tactile obtenue au travers de l'aiguille ( $R_{tool} \rightarrow U$ ).

### 11. Analyser la continuité de l'interaction en sortie aux niveaux perceptuel et cognitif pour le patient.

Comme pour le cas de l'aiguille, les environnements perceptifs requis par ces trois relations nécessitent deux lieux de perception qui ne sont pas superposables. Par conséquent, la perception des données relatives au patient souffre d'une **discontinuité perceptuelle**.

Au niveau cognitif, les référentiels de présentation adoptés sont différents, puisqu'il s'agit d'un référentiel de présentation utilisateur pour la perception directe (visuelle et tactile) et d'un référentiel de présentation lié à la sonde pour l'image échographique. De plus, la dimension des langages est différente dans la mesure où l'image échographique ne représente par définition qu'une coupe (2D). Ces différences au niveau des langages et référentiels de présentation sont susceptibles d'induire une **discontinuité cognitive** pour l'interprétation des données relatives au patient.

En synthèse de l'analyse ergonomique de *CASPER*, notons toutefois que pour le patient, les caractéristiques des langages sont moins différentes que dans le cas de l'aiguille. Notons aussi que l'image échographique apporte au clinicien une représentation du patient qu'il peut associer aux sensations tactiles perçues au travers de l'aiguille. En effet, les transitions entre tissus sont perceptibles par le retour tactile perçu via l'aiguille, mais aussi dans l'image échographique. Or, c'est en faisant progresser l'extrémité de l'aiguille le long de la trajectoire, que les différents tissus sont traversés. La perception tactile des transitions entre tissus est donc liée à la profondeur de pénétration de l'aiguille dans le corps du patient. La profondeur de pénétration étant représentée par la jauge affichée dans l'interface de guidage, l'image échographique établit un lien, possiblement discontinu au niveau cognitif, entre la perception tactile des transitions de tissus liées à la position de l'extrémité de l'aiguille réelle et la représentation de la cible par la jauge. Notons en outre, que dans une procédure conventionnelle, c'est-à-dire sans *CASPER*, le clinicien visualise toujours les images échographiques sur un échographe situé à proximité du patient. La discontinuité perceptuelle mise en évidence ci-dessus correspond donc à une situation de travail habituelle en cardiologie. Enfin, nous avons mentionné que l'image échographique n'est que faiblement utilisée par le clinicien. Par conséquent, l'impact des discontinuités perceptuelle et cognitive au niveau du patient n'est sans doute pas le problème majeur dans l'utilisation de *CASPER*.

Parallèlement à cette analyse ergonomique de *CASPER*, nous avons assisté à plusieurs utilisations de *CASPER* lors d'expérimentations menées par un chirurgien du service de chirurgie cardiaque sur des animaux, dans le cadre du processus de validation clinique de l'application. Au cours de ces observations, ainsi que lors d'entretiens informels menés avec le chirurgien, nous avons pu constater que le chirurgien devait réellement observer

alternativement l'écran puis le champ opératoire. Le fait de devoir alternativement regarder l'écran puis le patient et l'aiguille de ponction a effectivement été mentionné comme problématique par le chirurgien, notamment au moment de la ponction à proprement dit : à ce moment précis, l'extrémité de l'aiguille est située à l'intérieur de l'épanchement et le liquide pathologique constituant l'épanchement doit être aspiré. Etant très proche du cœur, l'aiguille doit rester dans la zone de sécurité définie avant l'intervention ; pour cela le chirurgien doit se baser sur les informations de guidage affichées à l'écran. Mais en parallèle il doit s'assurer que le liquide aspiré n'est pas coloré, ce qui signifierait que le liquide aspiré est du sang et que l'extrémité de l'aiguille ne se trouve pas dans l'épanchement mais possiblement dans le cœur. Le problème d'incompatibilité perceptuelle identifié avec notre méthode d'analyse ergonomique est donc corroboré par ces observations informelles.

De plus, avant de débiter la phase de guidage de chacune de ces interventions, le chirurgien consacre quelques instants à se remémorer l'interprétation des indications de guidage fournies par le système. Au cours de ce délais, certes limité à quelques dizaines de secondes, le chirurgien se positionne à distance de la peau du patient et s'aligne avec la trajectoire. Il procède alors à plusieurs mouvements de l'aiguille pour percevoir les modifications des informations de guidage qui en résultent. C'est donc une courte phase de ré-apprentissage qui est nécessaire lors de chaque intervention. Le chirurgien a d'ailleurs exprimé le souhait de disposer d'un gabarit dans lequel positionner avant chaque intervention son aiguille de ponction dans des positions prédéfinies, permettant ainsi de réapprendre rapidement à interpréter les informations de guidage. Du fait de problèmes d'asepsie de ce gabarit, un tel dispositif n'a pas encore pu être réalisé. Toutefois, ceci confirme les problèmes de discontinuité cognitive mis en évidence par notre analyse ergonomique. En effet, l'apprentissage requis est consécutif à une difficulté d'interprétation des données de guidage, que nous pouvons mettre en relation avec les facteurs de discontinuité cognitive identifiés, en particulier la différence entre les dimensions des langages requis et l'aspect arbitraire versus non arbitraire de ces différents langages.

## 5 Conclusion

Nous avons consacré ce chapitre à la première version de *CASPER*, un système pour assister le chirurgien lors d'une ponction péricardique. Nous avons motivé ce système par son apport clinique puis nous avons expliqué sa conception et réalisation. Après avoir expliqué son principe de fonctionnement basé sur des dispositifs de suivi (localisation), nous avons détaillé la conception de l'interface homme-machine en décrivant en détail l'analyse de la tâche. Cette conception s'appuie sur les étapes présentées au Chapitre V. Nous avons ensuite jeté un regard critique à ce système en le comparant, au sein de notre cadre taxinomique, à d'autres systèmes mixtes existants et en conduisant une analyse ergonomique de l'interaction du chirurgien avec le système mixte. Pour cela nous avons décrit *CASPER* avec notre notation ASUR présentée au Chapitre IV et appliqué la démarche d'analyse prédictive expliquée au Chapitre V. Ce chapitre dédié à *CASPER* a permis d'illustrer avec un même système nos résultats des Chapitres III, IV et V.

De l'analyse ergonomique de cette version de *CASPER*, nous retenons une incompatibilité et une discontinuité perceptuelles ainsi que des éléments facteurs d'incompatibilité et de discontinuité cognitives. La discontinuité cognitive et perceptuelle

est intimement liée à la perception et l'interprétation des données liées au concept de l'aiguille. Or cette entité est essentielle à la réalisation de la tâche de ponction. L'objet du chapitre suivant est d'étudier les problèmes d'incompatibilité et de discontinuité perceptuelles et cognitives mis en évidence dans ce chapitre, en proposant une nouvelle version de *CASPER*.

## 6 Références

[Barbe 97]

Barbe-Zoppis, C., "*Gestes Médico-Chirurgicaux sur tissus mous : application à la ponction péricardique assistée par échographie 2.5D*", Thèse de l'Institut National Polytechnique de Grenoble, France, (1997).

[Chavanon 97]

Chavanon, O., Barbe, C., Troccaz, J., Carrat, L., Ribuoat, C., Blin, D., "*Computer Assisted PERicardial punctures : animal feasibility study*", Acte de CVRMed/MRCAS'97, Grenoble, (1997), p. 285-291.

[Chavanon 98]

Chavanon, O., "*Ponction péricardique assistée par ordinateur. Guidage informatisé de la ponction péricardique*", Rapport soumis au Comité de Consultation pour la Protection des Personnes pour la Recherche Biomédicale (CCPPRB) en vue d'avoir l'autorisation de faire des tests sur des patients, (1998).

[Chavanon 00]

Chavanon, O., Carrat, L., Pasqualini, C., Dubois, E., Blin, D., Troccaz, J., "*Computer Guided Pericardiocentesis: Experimental results and Clinical Perspectives*", publiée par Urban & Vogel dans la revue *Herz* 25, 8, (2000), p. 761-768.

[Dubois 97]

Dubois, E., "*Interaction Personne-Système : le cas particulier des Gestes Médico-Chirurgicaux Assistés par Ordinateur*", rapport de DEA de l'Université Joseph Fourier, Grenoble - France, (1997), 113 pages.

[Reich 96]

Reich, Y., Konda, S., L., Levy, S., N., Monarch, I., A., Subrahmanian, E., "*Varieties and Issues of Participation and Design*", dans la revue *Design Studies*, 17, 2, (1996), p. 165-180.

# **CHAPITRE VII**

## **NOUVELLE VERSION DE CASPER : CONCEPTION, REALISATION ET EVALUATION EXPERIMENTALE**



## Sommaire

<b>Chapitre VII : Nouvelle Version de CASPER :</b>	
<b>Conception, Réalisation et Evaluation Expérimentale _____ 181</b>	
<b>1</b>	<b>Re-conception _____ 181</b>
1.1	Incompatibilité et discontinuité perceptuelles _____ 181
1.2	Facteurs d'incompatibilité et de discontinuité cognitives _____ 184
<b>2</b>	<b>Analyse ergonomique prédictive de CASPER-V2 _____ 187</b>
<b>3</b>	<b>Réalisation logicielle _____ 191</b>
3.1	Architecture logicielle _____ 191
3.1.1	Pilier informatique _____ 191
3.1.2	Contrôleur de Dialogue _____ 192
3.1.3	Pilier réel _____ 194
3.2	Représentation 3D _____ 196
3.2.1	Représentation dans un référentiel de présentation trajectoire _____ 197
3.2.2	Représentation dans un référentiel de présentation aiguille _____ 198
3.2.3	Représentation dans un référentiel de présentation utilisateur _____ 200
3.3	Calibrage du dispositif d'affichage _____ 204
<b>4</b>	<b>Evaluation Expérimentale _____ 207</b>
4.1	Cadre expérimental _____ 208
4.1.1	Utilisateurs _____ 208
4.1.2	Tâche _____ 208
4.1.3	Dispositif technique _____ 209
4.1.4	Mesures _____ 209
4.2	Méthodologie expérimentale _____ 210
4.2.1	Facteurs ou variables indépendantes _____ 210
4.2.2	Mesures ou variables dépendantes _____ 211
4.2.3	Hypothèses concernant le signal _____ 212
4.3	Analyse statistique du signal global _____ 214
4.3.1	Effets principaux et effets d'interaction _____ 214
4.3.2	Identification des effets _____ 216
4.3.3	Analyse des effets principaux _____ 218
4.3.3.1	Le support _____ 218
4.3.3.2	Le type de présentation _____ 219
4.3.3.3	Le référentiel de présentation _____ 220
4.3.3.4	Les mouvements selon les trois axes _____ 220
4.3.4	Analyse des interactions _____ 221
4.3.4.1	L'interaction entre représentation et axes _____ 221
4.3.4.2	L'interaction entre représentation, axes et référentiel de présentation _____ 222
4.3.4.3	L'interaction entre représentation et référentiel de présentation _____ 224
4.3.5	Bilan de l'analyse du signal global _____ 224
4.4	Analyse statistique de chaque composante du signal _____ 225
4.5	Bilan de l'évaluation expérimentale _____ 226
<b>5</b>	<b>Conclusion _____ 228</b>
<b>6</b>	<b>Références _____ 231</b>

Ce chapitre est motivé par le constat d'incompatibilité et de discontinuité dans la réalisation de la tâche de ponction avec le système *CASPER*. Nous proposons des solutions pour pallier à ce problème de discontinuité et parmi celles-ci, nous optons pour une nouvelle version que nous justifions. Une évaluation prédictive de la nouvelle version de *CASPER* ainsi conçue, que nous nommons *CASPER-V2*, est ensuite conduite avant d'aborder la réalisation logicielle, étape suivant la conception dans un cycle de vie du logiciel. Nous finissons ce chapitre par les résultats d'une évaluation expérimentale de *CASPER-V2*, dernière étape du cycle de vie du logiciel.

## Chapitre VII :

### Nouvelle Version de CASPER :

### Conception, Réalisation et Evaluation

### Expérimentale

## 1 Re-conception

### 1.1 Incompatibilité et discontinuité perceptuelles

L'incompatibilité et la discontinuité perceptuelles représentent le fait que des données relatives à la réalisation de la tâche requièrent deux environnements perceptifs non superposables, chaque environnement étant constitué du couple (sens perceptif, lieu de perception). Dans *CASPER*, les deux environnements perceptifs requis sont tous deux basés sur la vue, mais l'un est localisé au niveau du champ opératoire, tandis que l'autre se trouve au niveau de l'écran d'affichage des données de guidage. Ces dernières incluent la trajectoire, la cible, l'échographie de contrôle et les données numériques.

Pour supprimer l'incompatibilité et la discontinuité perceptuelles, plusieurs solutions sont envisageables : nous organisons notre analyse selon les trois sens perceptifs humains déjà utilisés en GMCAO : haptique, auditif et visuel.

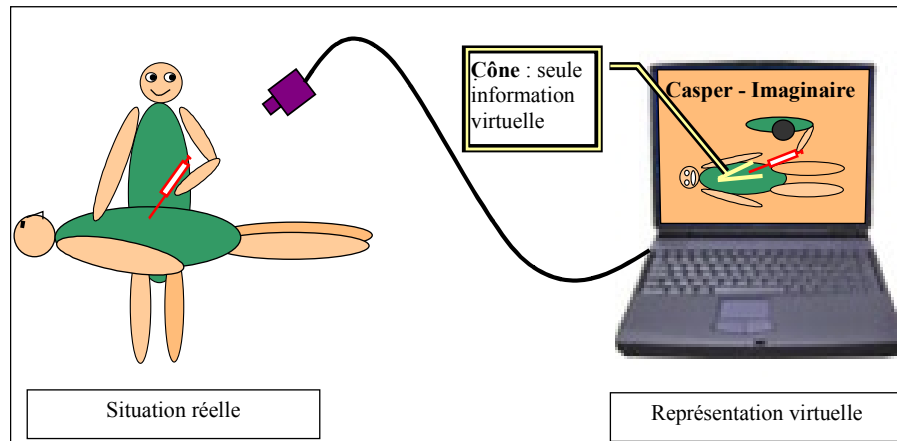
Au niveau **haptique**, l'usage d'un gabarit semblable à ceux de [Radermacher 97] n'est pas possible puisque la structure anatomique visée, le cœur, est une structure molle donc déformable.

La solution basée sur une action conjointe entre le clinicien et un robot est par contre tout à fait concevable. C'est une solution en cours de développement [Schneider 99]. Le bras robotisé maintient l'outil du clinicien dans l'alignement et l'empêche de sortir de la zone de sécurité, tandis que le clinicien imprime des mouvements à l'aiguille de ponction en vue d'amener son extrémité dans l'épanchement à ponctionner.

Une solution basée sur des données **sonores** pour présenter la trajectoire est aussi concevable. Un retour sonore traduirait à tout instant la position de l'aiguille par rapport à la trajectoire planifiée. Nous n'avons pas examiné plus en avant cette solution ; il est néanmoins important de noter que l'usage de sons reste très intrusif dans un bloc opératoire, où la communication entre les membres de l'équipe chirurgicale est fréquente et importante.

Considérons enfin le sens **visuel**. L'environnement perceptif lié à la perception de l'aiguille et du patient repose aussi sur la vue. Il convient donc dans les solutions envisagées d'obtenir un seul lieu de perception des données visuelles liées à l'aiguille, au patient et aux données de guidage. Pour cela, trois solutions sont envisageables : en regard du principe actuel de *CASPER*, il conviendrait soit de placer l'écran dans le champ opératoire, soit de rendre perceptible le champ opératoire dans l'écran. Enfin, la troisième solution consiste à placer dans le champ opératoire les données affichées jusqu'alors à l'écran. Nous détaillons ci-dessous ces trois solutions :

- 1. Placer l'écran dans le champ opératoire** : cette première solution n'est matériellement pas simple à mettre en œuvre. En effet, les contraintes de placement des différents dispositifs au sein d'un bloc opératoire sont très fortes. Ainsi, les deux bords de la table d'opération, de part et d'autre du patient, doivent être accessibles afin que le chirurgien et les membres de l'équipe chirurgicale puissent observer et éventuellement accéder au patient. De plus, la taille de l'écran est bien trop importante pour qu'il puisse être positionné suffisamment près du champ opératoire. Il conviendrait donc d'utiliser des minis écrans éventuellement fixés sur les outils chirurgicaux. Outre le fait que la technologie ne propose pas encore de tels dispositifs, cette solution impliquerait pour l'instant la présence de fils d'alimentation et de connexion au système informatique. Or placer des fils dans le champ opératoire encombrerait trop l'accès au patient et cela n'est pas, d'un point de vue de la sécurité, la solution la plus satisfaisante. Cette solution ne sera donc possible que lorsque des minis écrans pourront être fixés sur les outils manipulés et seront capables de fournir des données sans aucune connexion filaire, ni pour l'alimentation, ni pour la communication avec le système. Celle-ci pourraient être assurée par exemple par une communication infrarouge ou au moyen d'un réseau ethernet sans fil. Ces derniers pré-requis sont parfaitement remplis par certains modèles de *Pocket-PC*, tel que le *iPAQ* de Compaq par exemple. Mais la taille de ces dispositifs reste encore un obstacle à leur mise en œuvre en bloc opératoire.
- 2. Placer le champ opératoire dans l'écran** : cette seconde solution consiste en fait à "numériser" le champ opératoire sur lequel travaille le clinicien, afin de pouvoir le combiner aux données virtuelles au sein d'une représentation informatique. Cela correspond à l'approche de la RA du domaine de la synthèse d'image présentée au chapitre II. Une caméra filme le champ opératoire et envoie en temps réel le flux vidéo sur le dispositif affichant les données de guidage (un écran). Un traitement automatique du flux vidéo, basé par exemple sur la détection de marqueurs, permet d'identifier position et orientation des entités réelles par rapport auxquelles doivent être produites les données de guidage. Les données réelles et virtuelles sont ainsi superposées sur l'écran. La Figure 60 schématise ce principe.



**Figure 60** Schématisation de CASPER où réel et virtuel seraient fusionnés au sein d'une représentation virtuelle.

Ce principe correspond à la solution retenue par [Grimson 96] pour le système, présenté au chapitre I, assistant le clinicien dans la définition du contour d'une craniotomie à accomplir. Le problème majeur de cette solution de conception réside dans le fait que le clinicien ne perçoit plus visuellement et directement le patient : l'utilisation d'une caméra entraîne une indirection peu souhaitable en chirurgie.

**3. Placer les données de guidage dans le champ opératoire** : cette dernière solution est une mise en œuvre du paradigme de RA comme appréhendé dans la communauté IHM. Il s'agit de rendre perceptibles les données de guidage, générées par le système informatique, au niveau du champ opératoire. Ainsi les lieux de perception des différentes données fournies à l'utilisateur sont unifiés et l'utilisateur est toujours en contact direct avec le monde réel. Les problèmes d'incompatibilité et de discontinuité perceptuelles sont donc résolus tout en respectant les requis du domaine, ce qui n'était pas le cas de la solution de conception précédente.

Pour présenter les données de guidage dans le champ opératoire, plusieurs solutions techniques sont envisageables. Une première solution consisterait à projeter les données de guidage sur le patient. Cependant cette solution n'est pas facilement envisageable en milieu clinique pour deux raisons : tout d'abord le champ opératoire n'est pas net et par conséquent la perception des données projetées peut s'avérer difficile. D'autre part un scialitique est toujours présent au-dessus du champ opératoire afin de l'illuminer correctement. Intégrer un projecteur à ce scialitique constitue un enjeu technique.

Une seconde solution consiste alors à ajouter un support semi-transparent entre le clinicien et le champ opératoire. Le système de [Peuchot 95] présenté au chapitre I, est un exemple de mise en œuvre de cette solution. Cependant, pour des raisons d'organisation et d'encombrement du bloc opératoire déjà opposées à la solution précédente, cette solution n'est pas tout à fait adaptée au bloc opératoire.

Une dernière solution consiste à utiliser un casque semi-transparent porté par le clinicien. Cela correspond exactement à la solution précédente sans le problème de l'encombrement, puisque le casque est posé sur la tête du

clinicien. Le port d'un casque entraîne cependant un problème d'asepsie ainsi qu'une gêne éventuelle pour le clinicien, liée par exemple à la diminution de la largeur du champ de vision imposé par le casque. Notons cependant qu'en chirurgie cardiaque, le chirurgien est habitué au port de lunettes grossissantes pour effectuer des gestes de microchirurgie. De plus, la technologie des casques évoluant très rapidement, la largeur du champ de vision et le poids du casque ne constitueront bientôt plus un réel obstacle à leur utilisation.

C'est la dernière solution que nous avons choisie de mettre en œuvre dans *CASPER-V2*, afin de résoudre les problèmes d'incompatibilité et de discontinuité perceptuelles identifiés dans *CASPER* au chapitre précédent. Le sens perceptif que nous avons choisi de privilégier est la vue. Or, parmi les trois solutions techniques basées sur la vue que nous avons présentées précédemment, l'utilisation du casque semi-transparent nous semble la plus prometteuse et la moins intrusive au niveau du champ opératoire. L'incompatibilité et la discontinuité perceptuelles seront donc supprimées en réduisant à un seul le nombre d'environnements perceptifs requis : il s'agit du couple (vision, champ opératoire) dans lequel seront perçues les données de guidage grâce au casque semi-transparent, l'aiguille de ponction et le patient réel.

## 1.2 Facteurs d'incompatibilité et de discontinuité cognitives

Outre une incompatibilité et une discontinuité perceptuelles, nous avons souligné, au chapitre précédent, des caractéristiques de *CASPER* qui contribuent à une incompatibilité et une discontinuité cognitives. Comme nous l'avons expliqué, le problème est lié aux différents langages et référentiels de présentation utilisés pour rendre les données que le chirurgien doit interpréter lors de la tâche de ponction. En effet :

- l'aiguille de ponction et le patient sont perçus directement par le clinicien. Le langage considéré est donc 3D, non arbitraire et le référentiel de présentation utilisé est celui lié à l'utilisateur. Il s'agit de l'interprétation du monde réel perçu par l'utilisateur. La perception du patient est en outre complétée par l'image échographique, c'est-à-dire un langage 2D non arbitraire, dont le référentiel de présentation est lié à la sonde échographique.
- les données de guidage sont représentées par le viseur, langage 2D, arbitraire, par la jauge, langage 1D non arbitraire et les valeurs numériques, langage textuel. Toutes les données de guidage adoptent le référentiel de présentation lié à l'aiguille de ponction.

Ainsi, la mise en correspondance des données nécessite un effort cognitif d'interprétation qui peut être à l'origine d'incompatibilité et de discontinuité cognitives.

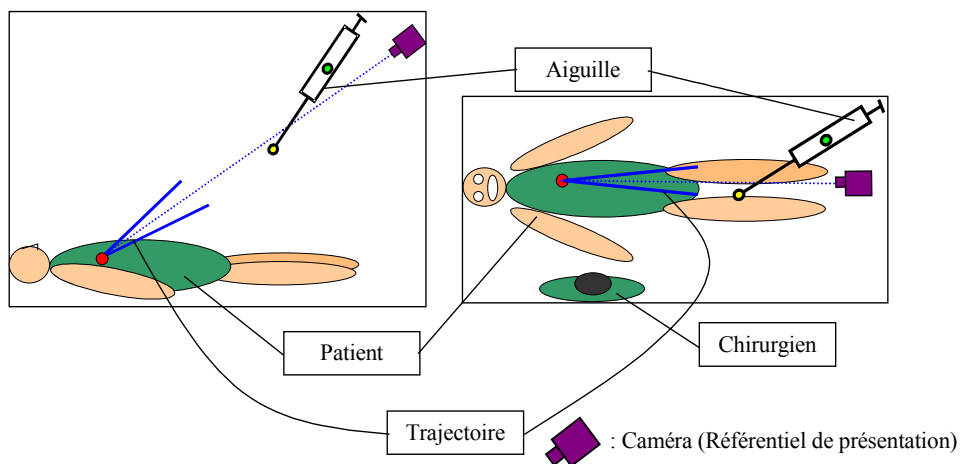
Pour résoudre ce problème, il est donc nécessaire d'agir sur les langages et les référentiels de présentation des données relatives à chacun des concepts mis en cause. Parmi ces données, la perception de l'aiguille réelle, celle du patient et celle de l'échographie ne peuvent pas être modifiées : il s'agit respectivement de la perception d'objets réels et d'une image médicale, incontournables pour la réalisation de la tâche. Nous nous concentrons donc dans la suite de cette analyse sur les données de guidage. Les langages et les référentiels de présentation utilisés pour exprimer ces données doivent

être similaires. En outre, les langages utilisés pour exprimer ces données doivent être similaires au langage associé à la perception 3D réelle du patient et de l'aiguille réelle. La représentation des données de guidage doit donc privilégier un langage 3D, non arbitraire.

La représentation 3D doit rendre compte de la trajectoire à suivre, de la cible à atteindre et de données plus précises quant aux distances séparant l'extrémité de l'aiguille de la trajectoire idéale. Pour représenter ces données, nous avons choisi d'utiliser un cône 3D tronqué. L'orientation du cône traduit la trajectoire à suivre, tandis que la plus petite base représente une zone de quelques millimètres autour du point cible. La largeur de la grande base du cône est définie en fonction de la position de structures anatomiques à éviter lors de l'avancée de l'aiguille à l'intérieur du corps du patient. De plus, le cône est constitué de  $n * m$  facettes où  $n$  représente le nombre de tranches horizontales et  $m$  le nombre de tranches verticales réalisées dans le cône. Chaque tranche du cône ayant une épaisseur de 1 millimètre, cette découpe fournit des indications précises de distance à la cible, exprimées dans la version précédente par les valeurs numériques à l'écran. Cette représentation est non arbitraire, dans la mesure où elle s'apparente à un objet du monde réel dans lequel l'utilisateur souhaite glisser l'aiguille, analogique, non linguistique et statique. Elle regroupe donc les données nécessaires à l'expression de la trajectoire. La forme adoptée pour cette représentation possède en outre des caractéristiques similaires à celles du langage utilisé par le clinicien lors de la perception réelle et directe de son aiguille et du patient.

Pour représenter l'aiguille il convient également d'adopter une représentation 3D et, si possible, non arbitraire. Nous avons donc défini un objet 3D modélisant l'aiguille chirurgicale utilisée. Cependant, la nécessité de représenter l'aiguille dépend du référentiel de présentation adopté pour produire les données de guidage. Pour expliquer ces différentes solutions, nous adoptons l'analogie à une caméra : la position et l'orientation de la caméra définissent l'origine et l'orientation du référentiel de présentation. Nous avons donc envisagé trois référentiels de présentation différents :

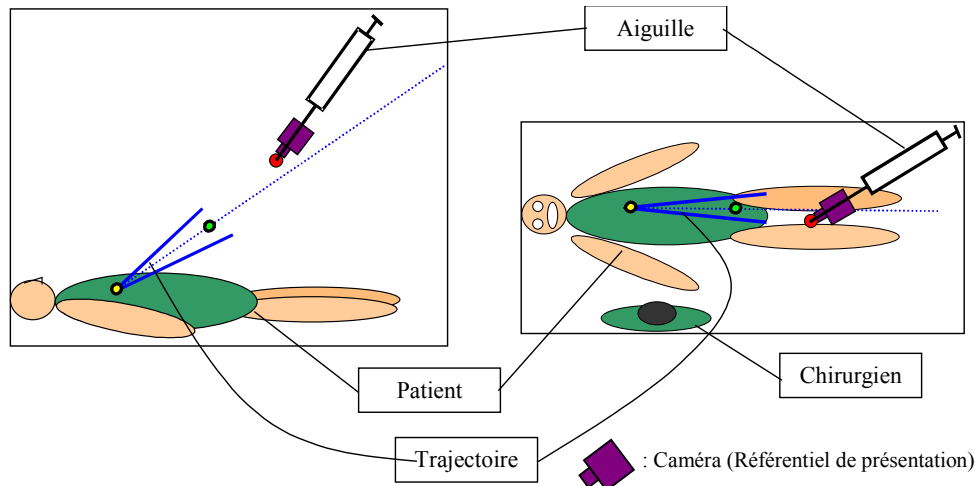
- **le référentiel de présentation trajectoire** : la caméra est ici positionnée dans l'alignement de la trajectoire, comme l'illustre la Figure 61.



**Figure 61** Schématisation du référentiel de présentation trajectoire pour la représentation des données dans CASPER-V2.

La représentation de celle-ci est donc statique et située au centre de la zone de rendu. Seule la représentation de l'aiguille doit apparaître et être mise à jour en fonction des mouvements de l'utilisateur. C'est le référentiel de présentation qui avait été adopté dans la première version de *CASPER*.

- **le référentiel de présentation aiguille** : la caméra est positionnée sur l'aiguille de ponction. Concrètement, l'aiguille n'a plus besoin d'être représentée puisque le centre de l'objectif est en fait situé à l'extrémité de l'aiguille. Adopter ce référentiel place le clinicien dans une situation analogique à celle d'une chirurgie laparoscopique : le clinicien a l'impression "d'entrer lui-même à l'intérieur du patient". Ce référentiel de présentation est particulièrement intéressant lorsque la visualisation de structures anatomiques modélisées est ajoutée à la scène. Par conséquent, seule la représentation de la trajectoire doit être rendue. Cette représentation est alors dynamique car elle dépend de la position et de l'orientation de l'aiguille, et donc de la caméra fixée sur elle. Cette configuration est schématisée dans la Figure 62.



**Figure 62** Schématisation du référentiel de présentation aiguille pour la représentation des données dans *CASPER-V2*

- **le référentiel de présentation utilisateur** : dans ce dernier cas, l'objectif de la caméra est "fixé" sur les yeux de l'utilisateur. La vue représentée correspond donc à la vue que l'utilisateur a naturellement sur le monde réel. Les données virtuelles sont alors alignées avec le monde réel. En conséquence, si on choisit de faire apparaître une représentation de l'aiguille, celle-ci doit être exactement alignée avec l'aiguille réelle. Dans cette configuration, la représentation virtuelle de l'aiguille n'est donc pas absolument nécessaire, tant que l'aiguille réelle est visible naturellement, sauf pour permettre au clinicien de vérifier la correspondance entre le modèle virtuel et les données réelles.

Afin de réduire les facteurs d'incompatibilité et de discontinuité cognitives identifiés dans *CASPER*, nous adopterons dans *CASPER-V2* la représentation des données de guidage basée sur un cône 3D. Ainsi toutes les données de guidage (trajectoire, cible et valeurs numériques) sont regroupées au sein d'une même représentation, qui utilise un seul langage 3D non arbitraire.

Nous avons également envisagé trois référentiels de présentation pour ces données. Adopter un référentiel utilisateur pour présenter les données de guidage diminue encore les facteurs d'incompatibilité et de discontinuité cognitives, puisque ce référentiel est requis pour visualiser le patient et l'aiguille réelle. Cependant, nous avons choisi de mettre en œuvre les trois référentiels de présentation dans *CASPER-V2* et de laisser le choix final à l'utilisateur.

## 2 Analyse ergonomique prédictive de CASPER-V2

L'analyse ergonomique menée au chapitre précédent a mis en évidence des faiblesses de *CASPER*. Pour pallier à ces faiblesses, nous avons présenté dans le paragraphe 1 des alternatives à la conception. Avant de passer à la réalisation logicielle de cette nouvelle version du système, *CASPER-V2*, il convient de mener une analyse ergonomique prédictive prenant en considération les solutions de conception envisagées. Pour cela, nous appliquons les 9 étapes de notre démarche d'analyse, présentée au chapitre V.

### 1. Description du système avec ASUR

La description ASUR de *CASPER-V2* est proche de celle de *CASPER*. En effet seul l'adaptateur de sortie (composant  $A_{out}$ ) a été modifié : il s'agit d'un casque semi-transparent au lieu d'un écran. Le casque comme l'écran fournissent des données au clinicien. Par conséquent la représentation des composants et des relations reste inchangée. Néanmoins, le casque doit être localisé en position et en orientation par l'adaptateur d'entrée (composant  $A_{in}$ ) formé de caméras et de diodes. Ainsi, par rapport à la description ASUR de la version précédente, seule une relation est ajoutée, du casque vers le localisateur ( $A_{out} \rightarrow A_{in}$ ). La description ASUR de *CASPER-V2* est présentée à la Figure 63.

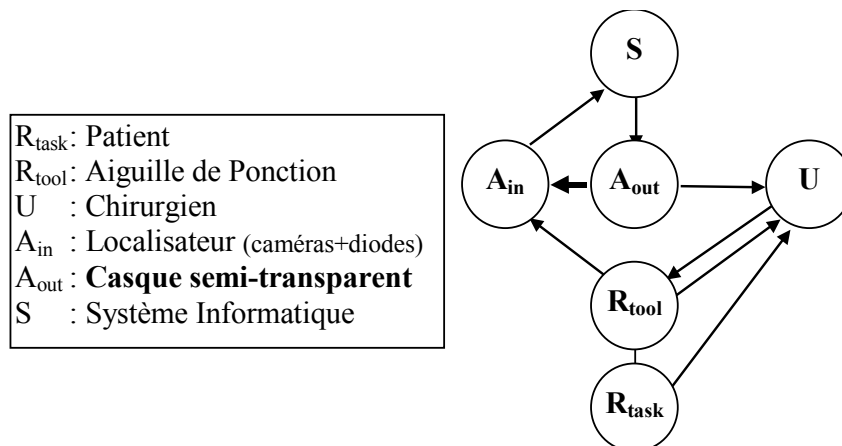


Figure 63 Description avec ASUR de *CASPER-V2*.

Au niveau des propriétés des relations et composants participant à l'interaction Utilisateur-Système Mixte, seuls l'adaptateur de sortie et la relation entre celui-ci et l'utilisateur ( $A_{out} \rightarrow U$ ) sont modifiés par rapport à *CASPER*.

Le **sens** requis par le nouvel adaptateur de sortie est toujours la vision. Par contre, le **lieu** de perception est limité au champ opératoire. Quant à son niveau de **partage**, il est réduit à un, puisque seul le clinicien porteur du casque peut percevoir les données affichées dans ce casque.



La relation entre le casque et l'utilisateur met désormais en jeu un **langage** 3D, non arbitraire, composé d'une représentation de la trajectoire à suivre et, selon le référentiel de présentation choisi, d'une représentation de l'aiguille manipulée. Son **poids** est toujours élevé et le **référentiel de présentation** qu'elle adopte peut être lié à la trajectoire, à l'aiguille ou à l'utilisateur (cf. paragraphe 1.2).

Les tableaux de la Figure 64 résument l'analyse des propriétés des composants et relations formant l'interaction de l'utilisateur avec *CASPER-V2*.

a) Composants	Sens Perceptif et Actionnel	Lieu	Niveau de Partage
Patient	Visuel	Champ opératoire	N
Aiguille de ponction	Visuel/Haptique	Champ opératoire	N
	Action Physique	Champ opératoire	1
Casque semi-transparent	Visuel	Champ opératoire	1

b) Relations	Langage	Poids	Point de vue
$A_{out} \rightarrow U$ Cône	3D, non arb., anal., non ling., dyn.	Elevé	Utilisateur, Trajectoire ou Aiguille
Echographie	2D, non arb., anal., non ling., stat.	Faible	Sonde
$R_{tool} \rightarrow U$	3D, réel	Elevé	Utilisateur
$U \rightarrow R_{tool}$	3D, réel	Elevé	Utilisateur
$R_{task} \rightarrow U$	3D, réel	Elevé	Utilisateur

**Figure 64** Caractérisation des composants (a) et relations (b) participant à l'interaction dans *CASPER-V2*.

## 2. Identification des concepts principaux

Les concepts manipulés par *CASPER-V2* sont les mêmes que dans la version précédente du système mixte : il s'agit du patient, de la trajectoire à reproduire, de la cible et de l'aiguille chirurgicale utilisée pour ponctionner l'épanchement.

## 3. Identification des relations ASUR, de poids élevé, formant l'interaction de sortie du système

Ces relations sont là encore les mêmes que dans la précédente version. Il s'agit des relations provenant du casque semi-transparent ( $A_{out} \rightarrow U$ ), du patient ( $R_{task} \rightarrow U$ ) et de l'aiguille de ponction ( $R_{tool} \rightarrow U$ ).

## 4. Caractérisation des environnements perceptifs des composants ASUR à l'origine des relations identifiées à l'étape 3

Les trois environnements perceptifs requis par les composants intervenant dans l'interaction en sortie du système sont identiques : il s'agit du couple composé de la vue (sens perceptif utilisé) et du champ opératoire (lieu de perception).

## 5. Caractérisation des langages et des référentiels de présentation mis en œuvre par les relations identifiées à l'étape 3

Les langages sur lesquels se basent les différentes facettes de l'interaction en sortie ont tous des caractéristiques communes. Dans les trois cas (en ne

considérant pas l'échographie), il s'agit d'un langage non arbitraire, analogique, non linguistique et 3D.

Les référentiels de présentation sont les mêmes pour les relations ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{U}$ ) et ( $\mathbf{R}_{\text{tool}} \rightarrow \mathbf{U}$ ). Par contre, les données fournies par l'adaptateur de sortie peuvent être présentées selon le même référentiel de présentation ou un référentiel lié à la trajectoire ou à l'aiguille, comme expliqué au paragraphe 1.2.

#### 6. Analyse de la compatibilité de l'interaction en sortie

Compte tenu de l'analyse des environnements perceptifs requis (étape 4), dans *CASPER-V2* le problème de compatibilité perceptuelle est résolu. L'analyse des propriétés des langages mis en jeu (étape 5) montre quant à elle que *CASPER-V2* réduit le nombre de facteurs contribuant à l'incompatibilité cognitive. En effet, le seul facteur restant est lié au référentiel de présentation mis en œuvre par la relation entre le casque et l'utilisateur ( $\mathbf{A}_{\text{out}} \rightarrow \mathbf{U}$ ). Il peut être différent ou identique au référentiel de présentation requis par les deux relations ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{U}$ ,  $\mathbf{R}_{\text{tool}} \rightarrow \mathbf{U}$ ) représentant respectivement la perception du patient réel et de l'aiguille réelle. Néanmoins, dans le cas où il est différent, toutes les données (trajectoire, cible, aiguille) sont exprimées selon ce même référentiel.

#### 7. Identification des relations relatives à un même concept : l'aiguille

Tout comme précédemment, seule une relation ASUR est le véhicule de données relatives au patient : elle représente la perception directe que le clinicien a du patient ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{U}$ ). La trajectoire est également totalement exprimée et perçue par le biais d'une seule relation ASUR. La nouvelle conception n'a donc pas modifié le nombre de relations relatives à chacun de ces deux concepts.

Concernant le concept de l'aiguille, trois cas sont possibles, selon le référentiel de présentation adopté. En effet, comme nous l'avons expliqué dans le paragraphe 1.2, si le référentiel de présentation choisi est celui lié à l'aiguille, aucune donnée relative au concept de l'aiguille n'est fournie par le système. Dans ce premier cas, aucun il n'existe plus aucun concept pour lequel des données seraient véhiculées par des relations multiples.

Le second cas correspond à une situation dans laquelle les données sont présentées dans un référentiel utilisateur. La représentation de l'aiguille dans les données fournies par le système est presque inutile, néanmoins, dans notre implémentation de ce référentiel, nous avons choisi de la représenter afin de permettre un contrôle de l'alignement des données réelles avec les données virtuelles. Deux relations dans la description ASUR du système sont alors porteuses de données relatives à l'aiguille : il s'agit d'une part de la perception directe de cette aiguille par le clinicien ( $\mathbf{R}_{\text{task}} \rightarrow \mathbf{U}$ ) et d'autre part de la perception de sa position et de son orientation par rapport à la trajectoire à reproduire ( $\mathbf{A}_{\text{out}} \rightarrow \mathbf{U}$ ).

Enfin, le troisième cas correspond aux situations basées sur un référentiel de présentation lié à la trajectoire. Une représentation de l'aiguille est obligatoirement contenue dans les données fournies par le système. Comme

dans le second cas, deux relations, ( $R_{task} \rightarrow U$ ) et ( $A_{out} \rightarrow U$ ) sont porteuses de données relatives à l'aiguille.

#### 8. Analyse de la continuité de l'interaction en sortie

Comme expliqué précédemment, cette analyse n'est valide que dans le cas où le référentiel de présentation associé à la relation entre l'adaptateur de sortie (le casque) et l'utilisateur est le référentiel de présentation lié à la trajectoire ou le référentiel utilisateur. En effet, l'étape 7 a montré que seuls ces deux cas mettent en œuvre de multiples relations porteuses de données relatives à un même concept, l'aiguille. Les relations dispensant des données relatives à l'aiguille sont dans ce cas celle provenant du casque semi-transparent ( $A_{out} \rightarrow U$ ) et celle représentant la vision directe que le chirurgien a de l'aiguille réelle ( $R_{tool} \rightarrow U$ ).

Dans le cas du référentiel trajectoire, l'analyse des environnements perceptifs menée à l'étape 4 nous permet de conclure que la continuité perceptuelle est assurée. L'étape 5 quant à elle nous permet d'établir qu'un facteur de discontinuité cognitive demeure : il s'agit d'une différence entre les référentiels de présentation requis. En effet, les données relatives à l'aiguille, portées par la relation entre le casque et l'utilisateur, sont exprimées dans un référentiel lié à la trajectoire, alors que la vue directe que l'utilisateur a de l'aiguille réelle correspond à un référentiel de présentation lié à l'utilisateur.

Lors de l'utilisation d'un référentiel de présentation lié à l'utilisateur, la continuité perceptuelle est également assurée. De plus, le référentiel de présentation de toutes les données est cette fois-ci le même. Dans ce cas, il n'y a plus aucun facteur de discontinuité cognitive.

#### 9. Recommencer à l'étape 7 pour un autre concept

*CASPER-V2* présente à l'utilisateur une image échographique de contrôle au clinicien. Il conviendrait donc de poursuivre l'analyse ergonomique en appliquant les étapes 7 et 8 au concept du patient : il est en effet perçu directement par le chirurgien, mais aussi par le biais de l'image échographique. Cette analyse aboutirait aux mêmes conclusions que celles obtenues pour *CASPER* dans le chapitre VI : une continuité perceptuelle et des risques de discontinuité cognitive. En effet, la juxtaposition de données 2D (l'image échographique) et 3D (le patient) constitue un facteur de discontinuité cognitive. Néanmoins, nous avons minoré ce facteur dans la mesure où la manipulation d'une échographie en chirurgie cardiaque est très courante. Notons également, que le référentiel utilisateur réduit encore d'avantage ce facteur de discontinuité cognitive : cette dernière est principalement due à l'utilisation d'un langage 2D (l'échographie) et d'un langage 3D (interprétation du patient perçu en 3D). Or, le référentiel permet d'aligner exactement l'image échographique 2D sur la partie 3D correspondante du patient. Par conséquent, l'image 2D apparaît comme une coupe de l'objet 3D, ce qui favorise le processus d'interprétation des données et réduit donc le risque de discontinuité cognitive.

En synthèse de l'analyse ergonomique de *CASPER-V2*, il apparaît que le problème d'incompatibilité perceptuelle relevé pour *CASPER* est supprimé. De plus, dans la

première version, des langages aux caractéristiques très différentes étaient mis en œuvre, alors que dans la nouvelle version ceux-ci ont été limités. Il en résulte une réduction du nombre de facteurs d'incompatibilité cognitive. En effet, au niveau cognitif, les facteurs d'incompatibilité sont réduits à un seul : le référentiel de présentation.

La re-conception de *CASPER* a également un effet bénéfique en ce qui concerne les données relatives à l'aiguille. En effet, la discontinuité cognitive mise en évidence dans *CASPER* est éludée dans *CASPER-V2*. L'interprétation des données relatives à ce concept, essentiel pour la réalisation de la tâche de ponction, semble également facilitée par *CASPER-V2*, puisque le seul facteur identifié comme pouvant accroître la discontinuité cognitive est là aussi le référentiel de présentation. En outre, ce facteur de discontinuité cognitive disparaît si le référentiel de présentation choisi pour les données de la relation du casque vers l'utilisateur est un référentiel utilisateur. Dans ce cas, plus aucun facteur de discontinuité cognitive n'est identifié par notre démarche concernant l'aiguille.

### 3 Réalisation logicielle

Au paragraphe précédent, nous avons montré par une évaluation prédictive que *CASPER-V2* réduisait, voire supprimait les problèmes de compatibilité et de discontinuité perceptuelle et cognitive. Il s'agit là de résultats d'une évaluation prédictive, avant le développement de *CASPER-V2*, qu'il convient d'attester ou d'infirmer par une évaluation expérimentale. La conception étant aboutie et validée par une évaluation prédictive, l'étape à considérer à ce stade du cycle de vie du logiciel *CASPER-V2* est donc la réalisation logicielle. Cette étape fait l'objet de ce paragraphe, le paragraphe suivant exposant l'évaluation expérimentale et ses conclusions.

Dans ce paragraphe, nous présentons d'abord l'architecture logicielle de *CASPER-V2* selon le modèle PAC-Amodeus exposé au Chapitre V. Nous détaillons ensuite les aspects logiciels liés à la représentation 3D des données de guidage ainsi qu'au calibrage du casque semi-transparent.

#### 3.1 Architecture logicielle

La réalisation logicielle de *CASPER-V2* repose sur le modèle d'architecture PAC-Amodeus. Ce modèle et nos adaptations à la réalisation de systèmes mixtes ont été exposés au chapitre V : en particulier deux piliers logiciels, notés piliers informatique et réel, ont été mis en exergue avec comme point d'ancrage le composant logiciel noté Contrôleur de Dialogue. Nous présentons dans les paragraphes suivants les composants logiciels de *CASPER-V2*, en détaillant le pilier informatique, puis le composant Contrôleur de Dialogue et enfin le pilier réel.

##### 3.1.1 Pilier informatique

Dans le Noyau Fonctionnel (NF), se trouvent les fonctions nécessaires au fonctionnement de *CASPER-V2* :

- le processus d'acquisition d'images, nécessaire pour obtenir et corriger les données fournies par l'échographe et transmises au système informatique (composant S dans la description ASUR de *CASPER-V2*),
- les fonctions d'identification des images valides,

- les fonctions permettant d'analyser la trajectoire proposée par le clinicien et d'en déduire les paramètres tels que la distance de sécurité ou la largeur de la zone de sécurité,
- les fonctions de gestion des informations à propos d'un patient : identité, images échographiques utilisées, trajectoires choisies, observations du clinicien, etc.,
- les fonctions de calibrages des outils.

Dans *CASPER-V2*, l'ANF joue donc essentiellement le rôle d'interface entre le CD et le NF. Il absorbe ainsi les modifications apportées au NF de sorte qu'elles n'aient aucune répercussion sur le code du CD.

### 3.1.2 Contrôleur de Dialogue

Le modèle PAC-Amodeus préconise de peupler le composant Contrôleur de Dialogue (CD) d'agents PAC. Le CD a la responsabilité de contrôler l'enchaînement des actions proposées par l'utilisateur. Il devrait donc comprendre un agent PAC par phase de la ponction péricardique assistée par ordinateur : calibrage des outils, acquisition des images, segmentation, choix d'une cible, puis guidage du geste de ponction. Ces agents devraient en outre être regroupés sous un agent PAC ciment, chargé d'établir la communication des données entre chacune de ces phases : nombres d'images valides, contour de la zone d'intérêt identifiée sur les images échographiques, position de la trajectoire planifiée. Enfin sous chacun des agents correspondant à une tâche, une hiérarchie d'agents PAC gèrerait le déroulement de la tâche considérée. La structure générale du CD selon le modèle PAC-Amodeus est illustrée à la Figure 65.

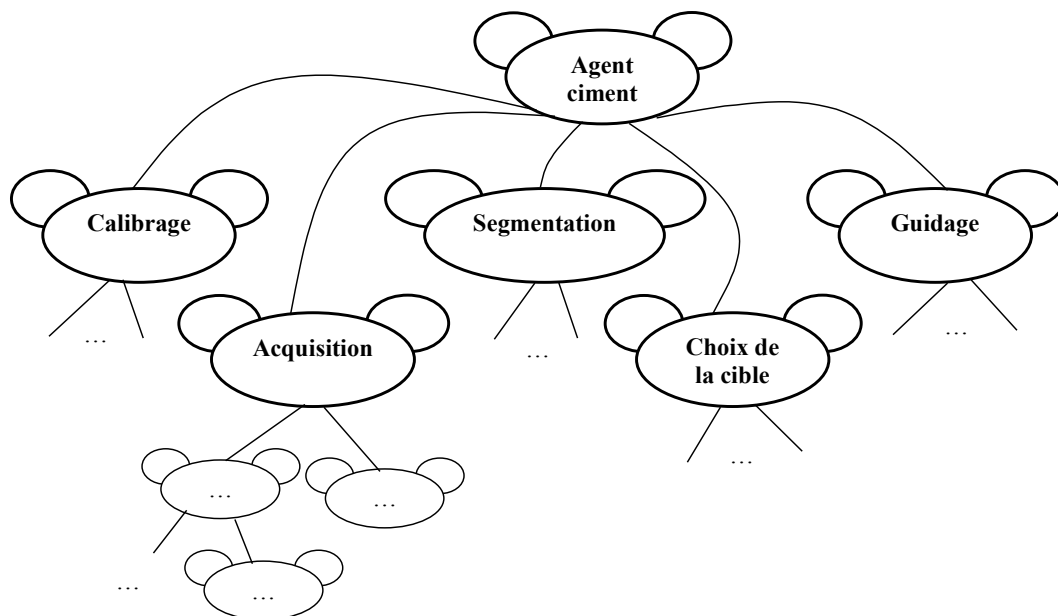
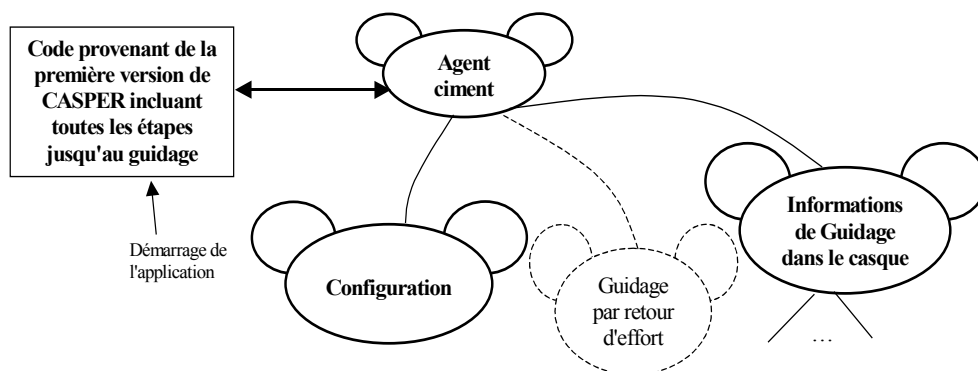


Figure 65 Structure du CD de *CASPER-V2* selon le modèle PAC-Amodeus.

Toutefois le code produit pour la première version de *CASPER* ne mettait pas en œuvre un modèle d'architecture particulier. Il aurait donc été nécessaire de reprogrammer la totalité du système pour respecter le modèle PAC-Amodeus. Or *CASPER* est en phase de validation clinique et tout changement dans le code doit impliquer de refaire tout le processus de validation technique et clinique. Par contre si un module seulement est ajouté dans le programme, par exemple la segmentation automatique ou l'affichage des

données de l'agent PAC "Information de guidage dans le casque", les modules précédemment validés ne sont pas réintégrés dans le nouveau processus de validation clinique. Nous nous sommes donc limités à mettre en œuvre le modèle d'architecture pour la partie nouvelle du code, celle dédiée à l'affichage des données dans le casque. Ainsi l'architecture effective du CD de *CASPER-V2* est présentée à la Figure 66. La structure du CD de la Figure 65 n'est donc que partiellement réalisée. Les intérêts de la structuration effective du code du CD de la Figure 66 sont les suivants :

- réutiliser les parties du système déjà validées cliniquement sans les modifier.
- intégrer selon le modèle PAC-Amodeus les éléments logiciels permettant la visualisation dans le casque, le choix du référentiel de présentation et le choix de la représentation.
- envisager l'intégration d'autres modules, par exemple un retour d'effort, sans devoir de nouveau repenser l'architecture totale du système.



**Figure 66** Structure effective du CD de *CASPER-V2*, intégrant le code existant de *CASPER*.

En appliquant ainsi le modèle d'architecture PAC-Amodeus, le lancement du système *CASPER-V2* s'effectue comme dans la version précédente. Une ligne a été insérée pour créer, dynamiquement au lancement de l'application, l'agent ciment, racine de la hiérarchie des agents PAC du CD.

L'agent dédié au guidage s'assure que les étapes précédentes (calibrage, acquisition, segmentation, définition de la cible) ont bien été menées lorsqu'il est sollicité, puis communique au pilier réel la position et l'orientation de la trajectoire établie par le clinicien à l'issue des phases précédentes. C'est également par ce composant que transitent les interactions de l'utilisateur avec l'interface, pour initialiser et arrêter le localisateur et, modifier la configuration des données affichées (référentiel de présentation, représentation, espace interpupillaire).

Parmi ces données, seul l'espace interpupillaire est maintenu dans la facette Abstraction de l'agent PAC "Informations de Guidage dans le casque". Les autres données citées, sont relatives à la configuration générale du système, puisqu'elles déterminent la forme sous laquelle les données doivent être fournies à l'utilisateur : le référentiel de présentation, la représentation utilisée (cône ou croix). Ces données relèvent de l'agent PAC "Configuration". Ainsi, lorsque l'utilisateur souhaite changer de représentation pour percevoir les données, le message est conduit au travers de la hiérarchie d'agents PAC jusqu'à l'agent "Configuration", qui s'assure de la cohérence de la représentation demandée. Ces choix sont alors sauvegardés dans la facette Abstraction de l'agent "Configuration".

Conformément à l'adaptation du modèle PAC-Amodeus aux systèmes mixtes, présentée dans le chapitre V, les données de position et d'orientation de l'aiguille, du patient et éventuellement de la tête du clinicien ne transitent pas par ce composant : elles restent dans le pilier réel de l'arche, que nous détaillons dans le paragraphe suivant.

### 3.1.3 Pilier réel

Le pilier réel de l'arche du modèle d'architecture PAC-Amodeus, est composé de deux *branches* comme le montre la Figure 67 :

- une branche est dédiée à l'interaction de l'utilisateur avec l'interface homme-machine de *CASPER-V2*, entrées de l'utilisateur vers le système et sorties du système vers l'utilisateur.
- l'autre branche gère la gestion des données du localisateur, données relatives aux positions de l'aiguille de ponction, du clinicien et du patient.

La mise en œuvre d'une branche spécifique au localisateur dans l'arche réelle du modèle PAC-Amodeus est justifiée par la nature des données manipulées. En effet les données de localisation des différentes entités ne font pas évoluer la tâche en cours, qui est le guidage. Ces données n'ont donc pas lieu d'être communiquées à la facette contrôle de l'agent PAC "informations de guidage dans le casque" du CD. De plus, ces données n'influencent pas la stratégie établie par les étapes précédentes de la ponction assistée par ordinateur. Donc ces données n'ont pas non plus lieu d'être communiquées à l'abstraction de l'agent PAC du CD.

La branche dédiée à l'interaction Utilisateur-*CASPER-V2* est composée de deux facettes : celle relative à l'interface graphique composée de boutons, fenêtres, zones de texte, et celle permettant l'affichage des données de guidage. La première facette, l'interface graphique, s'appuie sur la bibliothèque "*Microsoft Foundation Class*" (MFC). Celle-ci constitue les éléments d'un CIBN et, dans notre implémentation, le CTP relatif à cette facette n'a pas été développé, comme l'illustre le rectangle noir de la Figure 67 : le CD est donc dépendant des MFC. Mais, toutes nos applications étant développées sur la base des MFC, introduire un intermédiaire entre le CD et MFC ne se révèle pas utile dans notre contexte de travail.

La seconde facette regroupe donc tous les éléments nécessaires au rendu des données de guidage. Elle se compose donc des fonctionnalités de rendu graphique dans notre cas. Pour cela, nous avons utilisé la bibliothèque graphique OpenGL, qui constitue le CIBN de cette facette. Nous avons donc conçu un CTP qui rend le CD indépendant de la bibliothèque OpenGL. Il se compose de trois volets illustrés dans la Figure 67. Ces trois volets permettent d'instancier les données abstraites fournies par le CD concernant la représentation choisie, le référentiel de présentation et la dimension de la représentation utilisée. Chaque volet du CTP est donc chargé de traduire les données fournies par le CD en faisant appel aux fonctions du CIBN-OpenGL requises. Ainsi, si la représentation adoptée est un cône, le volet "Représentation" du CTP-Utilisateur est chargé de construire la scène 3D correspondant aux coordonnées de la trajectoire fournies par le CD. Il doit donc calculer les coordonnées de chaque facette constituant le cône, puis utiliser les commandes OpenGL permettant la construction de la scène 3D. Traduire le référentiel de présentation revient dans notre cas à positionner la caméra OpenGL dans la scène 3D

construite. Enfin une ou deux images sont générées si la dimension exprimée par le CD correspond respectivement à une visualisation monoscopique ou stéréoscopique.

La délégation au CTP de l'interprétation des préférences utilisateurs par le CD se justifie par le fait que tant que l'utilisateur n'agit pas sur la définition de la trajectoire à reproduire, ou plus généralement la stratégie opératoire, l'abstraction des agents PAC du CD n'est pas modifiée. Il n'y a donc aucun intérêt pour le CD de savoir si la représentation 3D de la trajectoire se base sur un cône ou un cylindre par exemple.

De plus, cette indépendance permet de faire évoluer le code relatif à la branche dédiée à l'interaction Utilisateur-CASPER-V2. Par exemple, si une évolution de CASPER-V2 met en œuvre une représentation plus complexe de la trajectoire, supposons un cône représentant la phase d'approche en dehors du corps du patient, puis un cylindre représentant la trajectoire à suivre à l'intérieur du corps du patient, il suffit de modifier les éléments logiciels chargés de la construction de la représentation de la trajectoire. Dans la Figure 67, ces éléments correspondent à ceux du volet "Représentation" de la branche utilisateur. Le CD n'est pas modifié et les autres composants du CTP-Utilisateur (référentiel de présentation, dimension) demeurent également inchangés.

Dans notre implémentation de la branche dédiée aux données fournies par le localisateur, le lien représenté à la Figure 67 entre le CTP-Localisateur et le CD n'est utilisé que pour deux types de messages seulement :

- recevoir les données de mise en route et d'arrêt du suivi des entités réelles : celles-ci sont en effet émises par l'agent PAC "Informations de guidage dans le Casque".
- notifier un dysfonctionnement dans le processus de localisation des entités. En effet, un tel dysfonctionnement perturbe voire même suspend la tâche en cours de réalisation. Il est donc opportun que le CD en soit informé. Le CD transmet ensuite cette donnée à l'utilisateur par le biais des éléments de l'interface.

Les données de localisation ne sont donc pas transmises au CD mais directement communiquées au CTP-Utilisateur, comme le représente le lien entre le CTP-Localisateur et le CTP-Utilisateur dans la Figure 67. Le CTP-Utilisateur réalise les actions requises puis teste l'arrivée de nouvelles valeurs fournies par le localisateur. Dans notre implémentation, le CTP-Localisateur constitue donc un intermédiaire entre le CD et les bibliothèques de fonctionnement du localisateur (les CIBN de la Figure 67 : CIBN-Polaris, CIBN-FreeD et CIBN-FlockOfBird), ainsi qu'entre le CTP-Utilisateur et les bibliothèques de fonctionnement du localisateur. Ainsi l'interrogation du localisateur est la même quel que soit le localisateur utilisé. Seul le CTP utilise l'appel aux fonctions spécifiques au localisateur utilisé. Cette architecture offre donc au CD une indépendance vis-à-vis des localisateurs utilisés : en effet le changement d'un localisateur *Flock of Bird* au profit d'un localisateur *Polaris* par exemple, nécessite seulement des modifications au niveau des composants logiciels internes au CTP-Localisateur. Les fonctions abstraites de manipulation du localisateur restent les mêmes et le changement de localisateur n'induit donc aucun changement au niveau du CD. Dans CASPER-V2, nous avons implémenté les trois volets du CTP-Localisateur permettant ainsi l'utilisation d'un localisateur *FreeD*, *Flock of Bird*, ou *Polaris*. Ce dernier est celui que nous utilisons désormais.



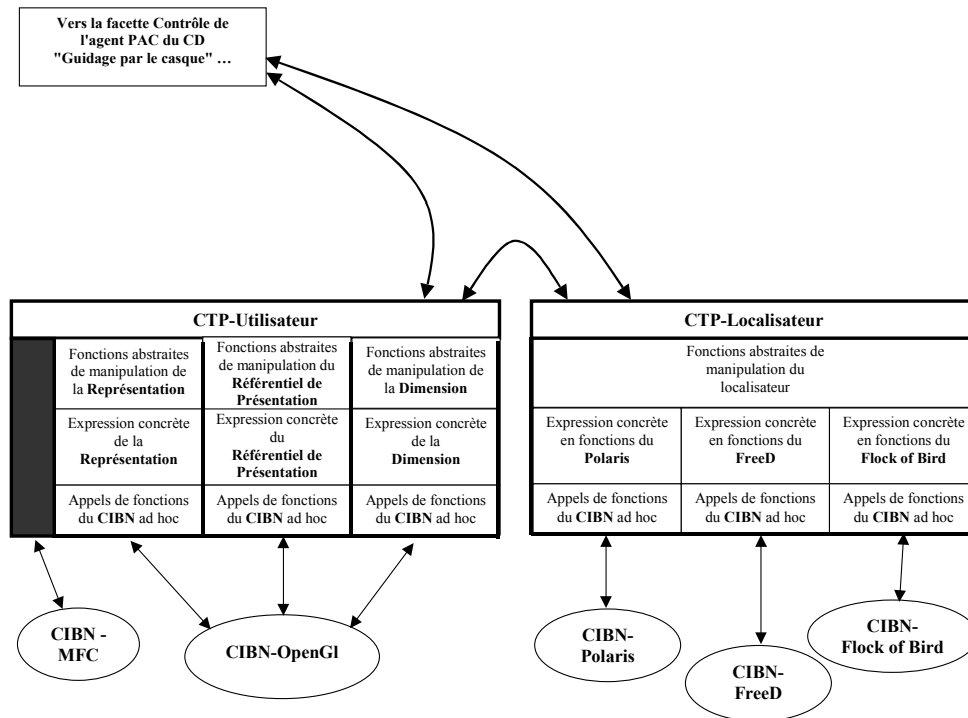


Figure 67 Architecture logicielle du pilier réel de CASPER-V2.

L'architecture logicielle globale de CASPER-V2 étant exposée et justifiée, nous détaillons deux aspects de réalisation logicielle qui sont nouveaux par rapport à CASPER : le paragraphe suivant 3.2 explique la réalisation logicielle de la représentation 3D des données de guidage puis le paragraphe 3.3 expose le calibrage du casque semi-transparent.

### 3.2 Représentation 3D

Nous avons identifié, au paragraphe 1.2, trois référentiels de présentation pour les données de guidage : le référentiel de présentation lié à la trajectoire, conforme à la première version de CASPER, le référentiel de présentation lié à l'aiguille et le référentiel de présentation lié au clinicien. Chacune des solutions de conception requiert une présentation des données différentes, mais les composants propres à la scène graphique affichée restent les mêmes.

Conformément aux pré-requis exprimés en phase de conception, cette représentation doit inclure un modèle de la trajectoire à suivre et de la cible à atteindre, ainsi qu'un modèle informant l'utilisateur de la position et de l'orientation de l'aiguille chirurgicale.

La trajectoire correspond à une demi-droite dans l'espace dont l'extrémité représente le point cible de l'épanchement à atteindre. Tout autour de la trajectoire se trouve une zone de sécurité, dans laquelle l'aiguille peut se situer sans risque pour les structures anatomiques voisines. Le modèle de la trajectoire est donc composé d'un cône, centré sur la trajectoire, représentant la zone de sécurité et dont la pointe tronquée représente une zone de quelques millimètres autour du point cible.

Le modèle de l'aiguille est constitué d'une sphère représentant l'extrémité de l'aiguille, d'un segment représentant l'axe de l'aiguille et d'un tore représentant un point de l'axe de l'aiguille. Ce point est le second point de l'aiguille dont les coordonnées sont fournies par le localisateur.

Ces deux modèles ont été réalisés avec des primitives de la bibliothèque graphique OpenGL et s'appuient sur le principe de localisation présenté au paragraphe 2.1.2 du Chapitre VI. L'algorithme repose sur le repère du viseur ( $R_{\text{viseur}}$ ) qui constitue le repère de base pour le calcul de la scène de guidage. Ainsi, le centre de la sphère, représentant l'extrémité de l'aiguille, a pour coordonnées dans la scène OpenGL celles de l'extrémité de l'aiguille réelle, ces dernières étant exprimées dans le repère du viseur ( $\text{Extrémité}_{\text{viseur}}$ ). Similairement, le centre du tore, représentant le point de l'axe de l'aiguille, a pour coordonnées celles du point de l'axe de l'aiguille exprimées dans le repère du viseur ( $\text{Axe}_{\text{viseur}}$ ).

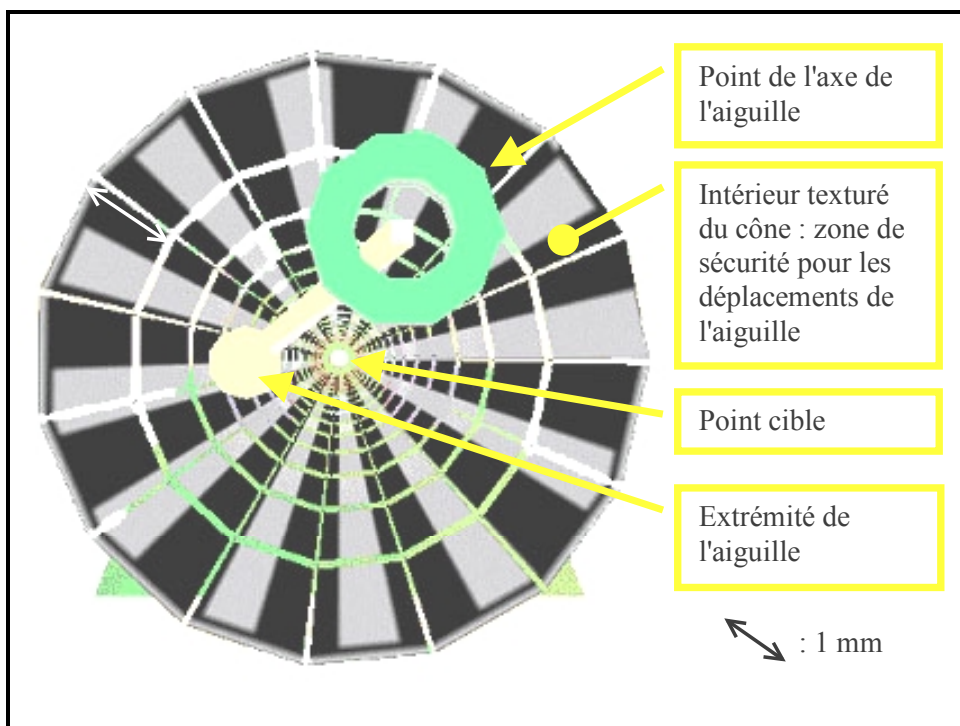
La seule différence entre les différents référentiels de présentation envisagés réside dans la position et l'orientation de la caméra OpenGL positionnée dans la scène : la position de la caméra OpenGL dans une scène 3D détermine ce qui est affiché à l'écran. De plus, une bonne perception de l'affichage, est directement liée à l'utilisation adéquate de lumières et de textures sur les éléments graphiques constituant la scène.

Dans les paragraphes suivants, nous décrivons la réalisation logicielle des trois référentiels identifiés ainsi que les résultats affichés dans le casque.

### 3.2.1 Représentation dans un référentiel de présentation trajectoire

Adopter un référentiel de présentation lié à la trajectoire consiste à positionner la caméra OpenGL sur la trajectoire à reproduire et à l'orienter en direction de la cible à atteindre, comme l'illustre la Figure 61. Le repère choisi pour représenter la scène étant celui du viseur, il convient de donner à la caméra OpenGL les coordonnées  $(O, O, z)$  où  $z$  varie en fonction de la position de l'aiguille réelle.

De plus, dans la scène de guidage affichée, la caméra doit être placée de sorte que l'aiguille se trouve entre l'objectif de la caméra et la cible à atteindre. En calculant les coordonnées de la projection orthogonale du point représentant l'axe de l'aiguille sur la trajectoire, on obtient la valeur minimale à donner à  $z$ . Nous avons ensuite choisi arbitrairement de reculer la caméra d'une distance fixe afin que l'utilisateur ait l'impression d'être placé juste derrière l'aiguille dans la représentation affichée. La Figure 68 montre la scène affichée lorsque ce référentiel de présentation est choisi.



**Figure 68** Présentation des données dans le référentiel de présentation trajectoire.

Lors de la ponction, il convient d'une part de déplacer l'aiguille de sorte que sa représentation ne sorte jamais du cône et jusqu'à ce que la sphère représentant l'extrémité de l'aiguille pénètre dans un volume OpenGL représentant l'épanchement et situé à l'extrémité du cône. D'autre part, pour reproduire au mieux la trajectoire, il convient de positionner la sphère appartenant au modèle de l'aiguille exactement sur le point cible représenté par le fond du cône. De plus, le tore doit être centré sur la trajectoire à reproduire. Un alignement idéal amène naturellement la sphère au milieu du tore, permettant ainsi à l'utilisateur d'être sûr que l'orientation actuelle de son aiguille est la même que celle de la trajectoire à reproduire.

Dans le cas où cet alignement n'est pas établi, le principe de correction de la position d'un des deux points de l'aiguille (extrémité ou axe) consiste :

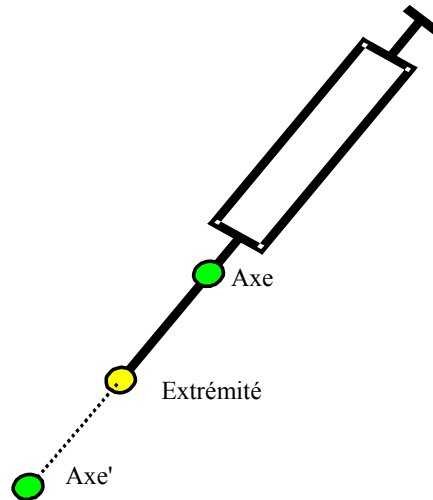
- Si la représentation d'un des deux points de l'aiguille est trop haute (resp. basse) par rapport au centre du cône : baisser (resp. monter) le point considéré de l'aiguille réelle.
- Si la représentation d'un des deux points de l'aiguille est trop à gauche (resp. à droite) par rapport au centre du cône : ramener le point considéré de l'aiguille réelle vers la droite (resp. la gauche).

L'écran de visualisation des données de guidage étant situé sur la tête du patient, les mouvements que le chirurgien applique à l'aiguille sont reflétés dans le même sens à l'écran : s'il bouge vers la gauche de la trajectoire, la représentation de l'aiguille se déplace vers la gauche de l'écran, et ainsi vice-versa.

### 3.2.2 Représentation dans un référentiel de présentation aiguille

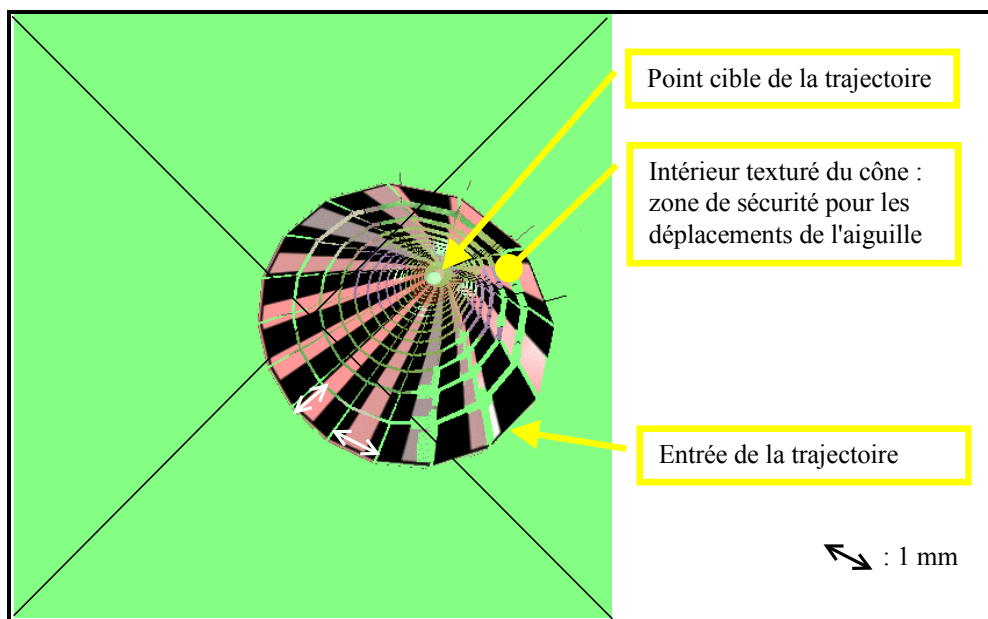
Adopter un référentiel de présentation aiguille revient à positionner la caméra OpenGL sur l'aiguille et à l'orienter selon la même direction que l'aiguille réelle, comme l'illustre la Figure 62. Le repère choisi pour représenter la scène étant celui du viseur, la position de

la caméra OpenGL est celle de l'extrémité de l'aiguille exprimée dans le repère du viseur ( $Aiguille_{\text{viseur}}$ ). Pour que la caméra soit orientée selon la même direction que celle de l'aiguille, la caméra doit être orientée en direction du point  $Axe'$ , où  $Axe'$  est le symétrique du point de l'axe de l'aiguille par rapport à l'extrémité de l'aiguille, comme l'illustre la Figure 69.



**Figure 69** Schématisation des points  $Axe$ ,  $Extrémité$  et  $Axe'$  de l'aiguille de ponction.

Dans ce référentiel, la représentation de l'aiguille est implicite. La caméra étant fixée sur l'aiguille à son extrémité, le centre de l'objectif de la caméra, c'est-à-dire le centre de la zone graphique affichée à l'écran, représente l'extrémité de l'aiguille. Ce point est mis en évidence par l'intersection des deux diagonales, comme le montre la Figure 70. L'orientation de l'aiguille est également implicite : elle est représentée par la direction de la caméra; l'orientation de l'aiguille est donc perpendiculaire aux données de guidage affichées.



**Figure 70** Présentation des données dans le référentiel de présentation aiguille.

Lors de la ponction, il convient d'une part de déplacer l'aiguille de sorte que le centre de la zone d'affichage ne soit jamais hors du cône et ce, jusqu'à ce que l'utilisateur

visualise l'intérieur du volume OpenGL situé à l'extrémité du cône et représentant l'épanchement à ponctionner. D'autre part, pour reproduire au mieux la trajectoire, et avant tout autre déplacement, l'extrémité la plus petite du cône doit être située le plus près possible du centre de la zone d'affichage (intersection des deux diagonales). Si l'extrémité du cône est décalée vers la droite par rapport au centre, cela représente le cas où l'aiguille est trop à gauche par rapport à la cible. De plus, les surfaces du cône doivent être équitablement réparties de part et d'autre du centre de la zone d'affichage. La répartition de la représentation des parois du cône autour du centre de la zone d'affichage traduit la correspondance entre l'orientation de l'aiguille et celle de la trajectoire. En effet, le référentiel de présentation lié à l'aiguille entraîne la même visualisation de la scène 3D que celle que l'utilisateur aurait s'il était lui-même à l'extrémité de l'aiguille. Par conséquent, voir la partie inférieure de la paroi plus grande que la partie supérieure, signifie que le regard est dirigé vers le bas du cône, c'est-à-dire que l'aiguille est trop inclinée vers le devant. La difficulté apparente de compréhension des données présentées dans ce référentiel vient du fait que ce qui est immobile dans cette vue, le centre, représente ce que l'utilisateur peut déplacer, l'aiguille. Cela nécessite donc un processus d'apprentissage ; ce processus de manipulation peut être relié au processus utilisé lorsque l'utilisateur veut viser quelque chose du regard. La combinaison des translations et des rotations à appliquer à l'aiguille est donc similaire à celles à appliquer aux mouvements de la tête pour porter le regard sur un point précis :

- Si l'extrémité la plus petite du cône est trop haute (resp. à gauche) par rapport au centre de la zone d'affichage, cela signifie que l'origine du référentiel, c'est-à-dire l'extrémité de l'aiguille, est trop basse (resp. à droite) : il convient alors de monter (resp. déplacer vers la gauche) l'aiguille réelle, et inversement.
- Si l'extrémité du cône est correctement positionnée par rapport au centre de la zone d'affichage, et si les parois du cône apparaissent plus larges en bas (resp. à gauche) par rapport au centre du cône, cela signifie que l'orientation de l'aiguille est dirigée trop bas (resp. à gauche) : il convient alors de déplacer le talon de l'aiguille, c'est-à-dire l'extrémité opposée à la pointe de l'aiguille, vers le bas (resp. la gauche).

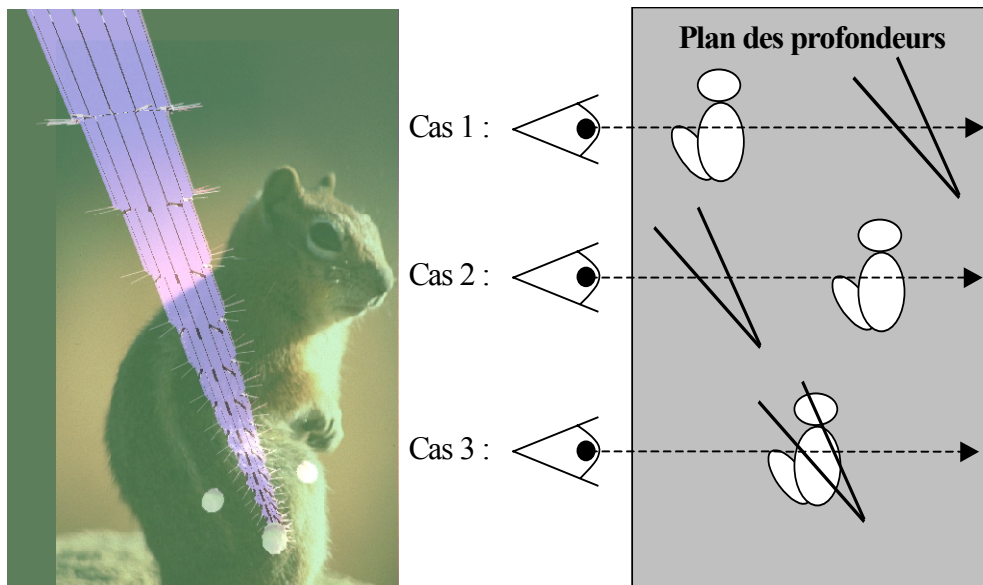
En conclusion, dans ce référentiel, les corrections doivent donc être effectuées dans le même sens que l'erreur observée : ainsi si la cible est plus basse que le centre, il convient de baisser l'aiguille; si les parois sont plus larges en bas, il faut baisser "l'arrière" de l'aiguille.

### **3.2.3 Représentation dans un référentiel de présentation utilisateur**

Le référentiel de présentation utilisateur fournit le résultat sans doute le plus réaliste : les données de guidage sont directement alignées avec le patient. La position de la caméra doit donc correspondre à la position des yeux du clinicien par rapport à son patient. La caméra est ensuite orientée dans la direction de l'extrémité de l'aiguille. Positionner la caméra OpenGL nécessite donc de connaître avec précision la position des yeux du clinicien dans le repère de référence. La technique que nous avons mise au point est plus largement détaillée dans le paragraphe 3.3.

Lors de l'utilisation, le cône représentant la trajectoire à reproduire, étant directement aligné avec le patient et le casque étant semi-transparent, le clinicien doit juste incliner l'aiguille réelle de sorte qu'elle ait la même orientation que celle du cône affiché dans le casque. Une fois l'aiguille réelle alignée avec le cône virtuel, le clinicien translate l'aiguille en conservant l'alignement avec le cône jusqu'à ce que l'extrémité de l'aiguille réelle atteigne la cible, ou si la cible est à l'intérieur du corps du patient, jusqu'à ce que l'extrémité de la représentation de l'aiguille affichée dans le casque atteigne le volume 3D OpenGL représentant l'épanchement.

Cependant, l'ajout de données virtuelles sur le monde réel soulève le problème de la mise en correspondance des profondeurs des champs de vision des scènes réelle et virtuelle. La Figure 71 illustre ce problème : lors de la superposition de l'écureuil (données réelles) et du cône (données virtuelles), comment savoir si l'écureuil est devant (cas 1), derrière (cas 2) ou dans le cône (cas 3) ? Dans notre cas, la question est de savoir si l'aiguille réelle est devant ou derrière le cône, c'est-à-dire entre les yeux du clinicien et la trajectoire, ou au-delà de la trajectoire. Ce problème est spécifique à ce référentiel de présentation car dans les autres cas, l'aiguille ou la trajectoire constituent le référentiel. La position de l'une de ces entités par rapport à l'autre est donc matérialisée dans les données produites.



**Figure 71** *Problème de superposition des profondeurs de champs de vision des scènes réelle et virtuelle.*

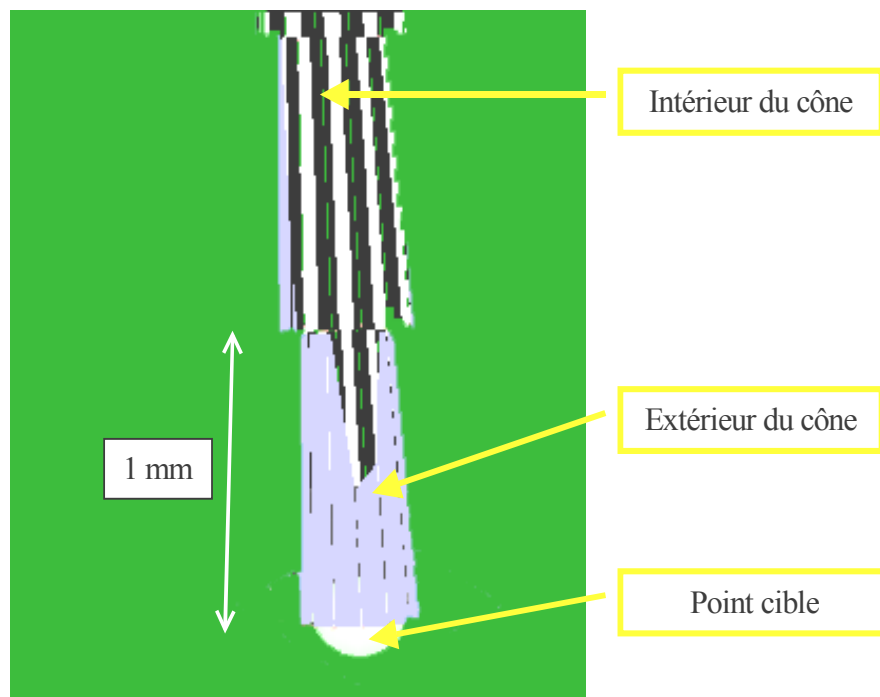
L'appréciation de la profondeur dans une scène virtuelle n'étant pas possible, la mise en correspondance des profondeurs des champs réel et virtuel ne l'est pas non plus. Aussi, la solution adoptée s'appuie sur la technique de "*Clipping Plane*". Cette technique permet de cacher tous les éléments de la scène 3D situés entre la caméra OpenGL, point de vue sur la scène, et un plan défini par son équation  $Ax + By + Cz + D = 0$ . L'équation du plan doit être donnée dans le même repère que celui dans lequel sont exprimées les coordonnées de la position de la caméra OpenGL.

A chaque rendu de la scène, l'équation d'un plan est calculée : l'extrémité de l'aiguille appartient à ce plan et la normale à ce plan est égale au produit vectoriel de  $V1$  et  $V2$ , où le vecteur  $V1$  a pour origine l'extrémité de l'aiguille et pour fin le point de l'axe de

l'aiguille, et le vecteur V2 a pour origine l'œil gauche et pour fin l'œil droit de l'utilisateur. Concrètement, le plan ainsi défini contient l'aiguille et est parallèle à la droite passant par les deux yeux de l'utilisateur.

Grâce à cette technique, la scène rendue affiche :

- l'extérieur violet du cône si la distance entre les yeux du clinicien et l'aiguille est inférieure à la distance entre les yeux du clinicien et la paroi du cône la plus proche,
- rien du tout, si la distance entre les yeux du clinicien et l'aiguille est supérieure à la distance entre les yeux du clinicien et la paroi du cône la plus éloignée,
- un écorché du cône dans les autres cas. Ce cas est illustré dans la Figure 72. La partie hachurée représente l'intérieur du cône tandis que la partie unie représente la paroi extérieure du cône.



**Figure 72** Visualisation des données de guidage dans le référentiel de présentation utilisateur.

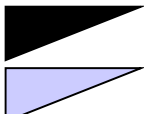
Ce plan, passant par les deux points définissant l'aiguille, fournit à l'utilisateur une indication permettant d'évaluer la position de l'aiguille réelle par rapport à la trajectoire planifiée dans le plan des profondeurs. L'indication lui permet aussi de s'assurer que l'orientation de l'aiguille dans le plan des profondeurs est en accord avec l'orientation de la trajectoire dans ce même plan. En effet s'il n'y a pas correspondance d'orientation, le cône sera "écorché" en biais.

En conclusion, les trois référentiels ont été développés dans *CASPER-V2*, selon les méthodes que nous venons d'exposer. En outre, nous avons implémenté la visualisation dans le casque de la représentation des données de guidage utilisée dans *CASPER*, c'est-à-dire le viseur basé sur les croix et la jauge, ainsi que la visualisation sur l'écran de la représentation 3D basée sur un cône. Ainsi, *CASPER-V2* propose deux représentations possibles : le cône et le viseur basé sur les croix. Pour chacune de ces représentations, le

référentiel de présentation lié à l'aiguille ou à la trajectoire est utilisable. Les données peuvent alors être affichées sur l'écran ou dans le casque. Enfin le référentiel de présentation lié à l'utilisateur est utilisable si la représentation choisie est le cône et dans ce cas les données sont nécessairement affichées dans le casque. Dans tous les cas, la configuration matérielle requise est la même que celle de *CASPER*. Seule l'utilisation du casque semi-transparent modifie légèrement la configuration matérielle : dans ce cas, le chirurgien porte le casque sur la tête et ce casque est localisé dans l'espace au moyen d'un corps-rigide fixé sur l'armature du casque. Nous détaillons plus avant le principe de localisation dans le paragraphe 3.3. La Figure 73 résume les différentes combinaisons du référentiel de présentation, de la représentation et du support d'affichage permises par *CASPER-V2* pour visualiser les données de guidage.

**Référentiels de Présentation**  
(Points de vue)

		Aiguille	Trajectoire	Utilisateur
Représentations	Croix et Jauge	Ecran Casque	Ecran Casque	
	Cône 3D	Ecran Casque	Ecran Casque	Casque



Configuration impossible

Configuration déjà en place dans *CASPER*

**Figure 73** Combinaisons du référentiel de présentation, de la représentation et du support d'affichage, utilisables pour présenter les données de guidage dans *CASPER-V2*.

Enfin, pour permettre à l'utilisateur de choisir un mode de présentation des données de guidage, la boîte de dialogue présentée à la Figure 74 permet de passer à tout moment d'une représentation à l'autre et d'un référentiel de présentation à l'autre. La phase de guidage n'est alors interrompue que pendant la durée d'affichage de cette boîte de dialogue.



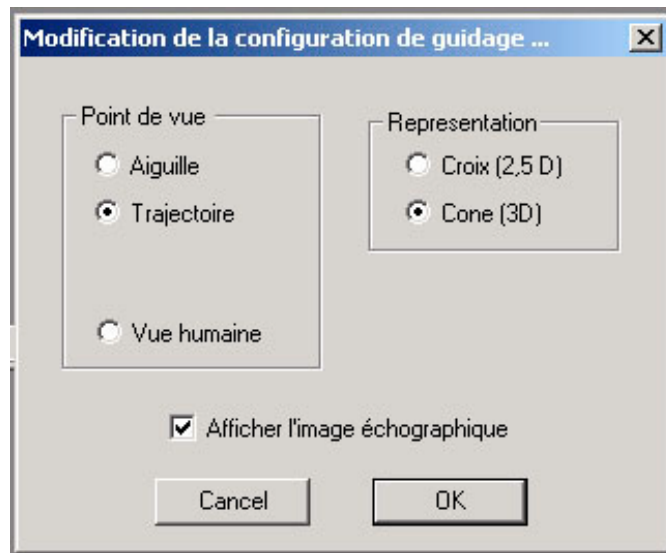


Figure 74 Boîte de dialogue permettant le choix de la représentation et du référentiel de présentation (point de vue) dans CASPER-V2.

### 3.3 Calibrage du dispositif d'affichage

Comme expliqué dans le paragraphe 1.1, la solution de conception retenue se base sur l'utilisation d'un casque semi-transparent. Le casque utilisé dans CASPER-V2 est le SONY LDI-D100 BE, semi-transparent et permettant la visualisation stéréoscopique. L'angle de vue horizontal est de 28 degrés et la résolution des écrans à cristaux liquides positionnés devant chaque œil est de 832 \* 624 \* 3. Le signal vidéo d'entrée est le même que celui fourni par un PC standard (VGA, ou VGA entrelacé pour l'affichage stéréoscopique). La fiche technique du casque est jointe dans l'annexe A.



Figure 75 Casque SONY équipé d'un corps-rigide blanc.

Pour le référentiel de présentation utilisateur, la localisation des yeux du clinicien est nécessaire, comme nous l'avons expliqué au paragraphe 3.2.3. Aussi, nous avons approximé la localisation des yeux du clinicien à la localisation du casque comme suit : la

position des yeux du chirurgien est considérée comme étant celle du centre des écrans semi-transparents composant le casque. Pour cela nous avons fixé un corps-rigide sur l'armature du casque comme le montre la Figure 75.

Comme pour les autres corps-rigides, un repère noté  $R_{\text{casque}}$  est associé au corps-rigide du casque. Il est nécessaire à chaque utilisation de calibrer le casque, c'est-à-dire de définir la position du centre des deux écrans du casque dans le repère  $R_{\text{casque}}$  : les coordonnées de ces deux points dans ce repère sont notées  $OeilGauche_{\text{Casque}}$  et  $OeilDroit_{\text{Casque}}$ . Après cette phase de calibrage, les écrans du casque demeurent à la même position par rapport au corps-rigide, sauf si l'utilisateur modifie la taille de l'armature du casque. Dans ce cas, le calibrage doit être refait.

Pour réaliser ce calibrage, nous adoptons une méthode en trois étapes :

**1. Détermination des coordonnées du centre de l'écran gauche par palpation des coins de l'écran**

Pour cela, nous utilisons un pointeur sur lequel est fixé un corps-rigide. L'extrémité de ce pointeur est positionnée successivement sur les quatre coins de l'écran gauche et pour chacun, les coordonnées de l'extrémité du pointeur sont enregistrées. Le calcul du barycentre de ces quatre points nous donne alors les coordonnées  $OeilGauche_{\text{Casque}}$  exprimées dans le repère du casque.

**2. Détermination des coordonnées du centre de l'écran droit par palpation des coins de l'écran**

La méthode est exactement la même et cette étape fournit les coordonnées  $OeilDroit_{\text{Casque}}$  exprimées dans le repère du casque.

**3. Détermination de la direction du regard**

Lors de cette troisième étape, l'utilisateur doit obligatoirement positionner le casque sur sa tête. La scène affichée se compose d'une sphère de petit rayon. L'utilisateur doit aligner la sphère avec l'extrémité du palpeur, tout en regardant droit devant lui. La direction du regard de l'utilisateur est alors déterminée comme étant la droite orthogonale à la droite passant par le centre des deux écrans et l'extrémité du palpeur. Une seconde solution envisagée mais non réalisée consiste à calculer la somme des normales aux plans des écrans droit et gauche. Cette somme des normales fournit alors une "normale moyenne" aux plans d'affichage et représente la direction du regard.

La Figure 76 montre les points palpés lors des trois étapes de calibrage du casque.

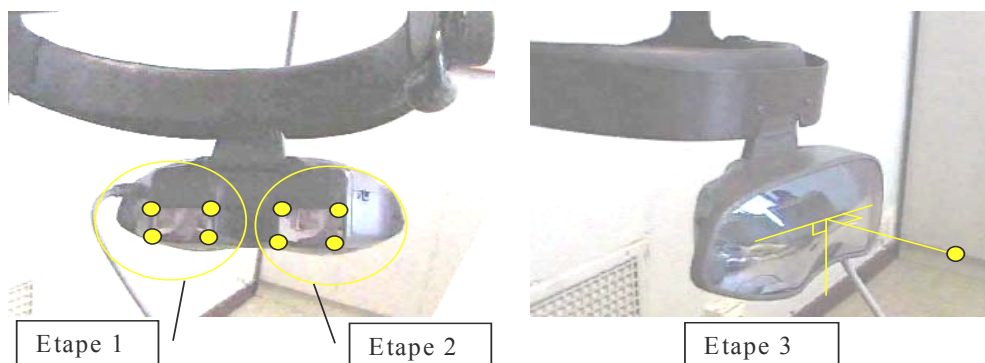


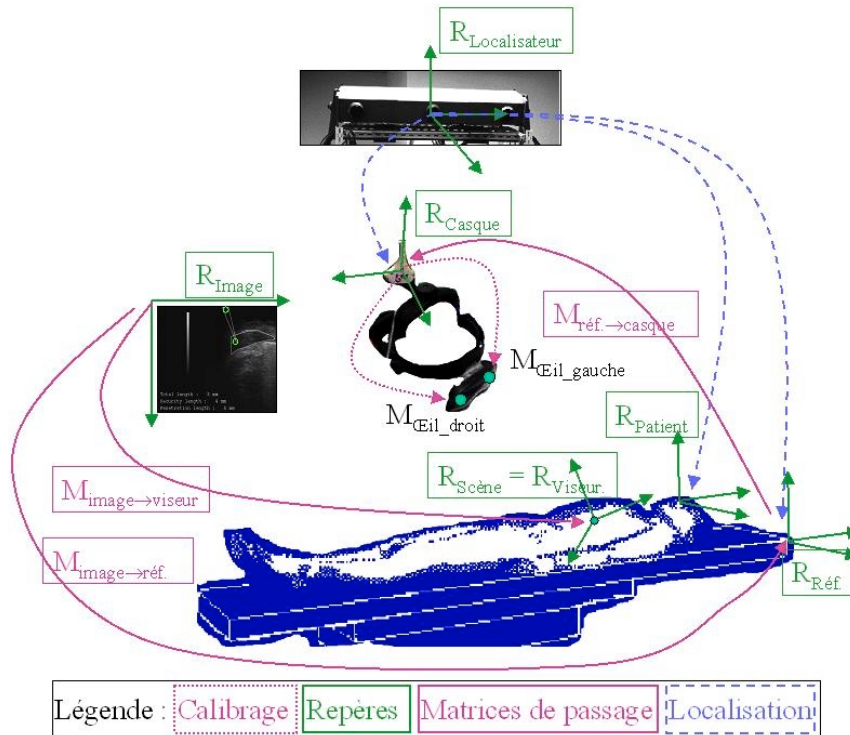
Figure 76 Schématisation des étapes de calibrage du casque.

Lors de l'utilisation, le localisateur fournit, à chaque itération du processus de localisation, une matrice de passage entre le repère de référence et le repère du casque :  $M_{Référence \rightarrow Casque}$ . Grâce à cette matrice, les mouvements de la tête du clinicien, et donc de ses yeux, sont pris en compte dans l'affichage des données de guidage. Les caméras OpenGL ayant les mêmes coordonnées que celles des écrans, il est nécessaire d'exprimer ces dernières dans le repère du viseur, puisque la construction de la scène OpenGL et la position des caméras OpenGL sont basées sur ce repère. On obtient donc :

$$\begin{aligned} \text{OeilGauche}_{\text{Viseur}} &= M_{\text{image} \rightarrow \text{viseur}} \times M_{\text{image} \rightarrow \text{Réf.}}^{-1} \times M_{\text{Référence} \rightarrow \text{Casque}}^{-1} \times \text{OeilGauche}_{\text{Casque}}, \text{ et} \\ \text{OeilDroit}_{\text{Viseur}} &= M_{\text{image} \rightarrow \text{viseur}} \times M_{\text{image} \rightarrow \text{Réf.}}^{-1} \times M_{\text{Référence} \rightarrow \text{Casque}}^{-1} \times \text{OeilDroit}_{\text{Casque}} \end{aligned}$$

Par la suite, le positionnement des caméras OpenGL dans la scène réutilise directement les coordonnées de l'œil droit et de l'œil gauche, exprimées dans le repère associé au viseur.

La Figure 77 représente les repères et les matrices de passage nécessaires pour l'affichage dans le casque des données de guidage alignées sur le monde réel.



**Figure 77** Repères et matrices de passage requis pour l'affichage de données alignées avec le monde réel dans le casque semi-transparent.

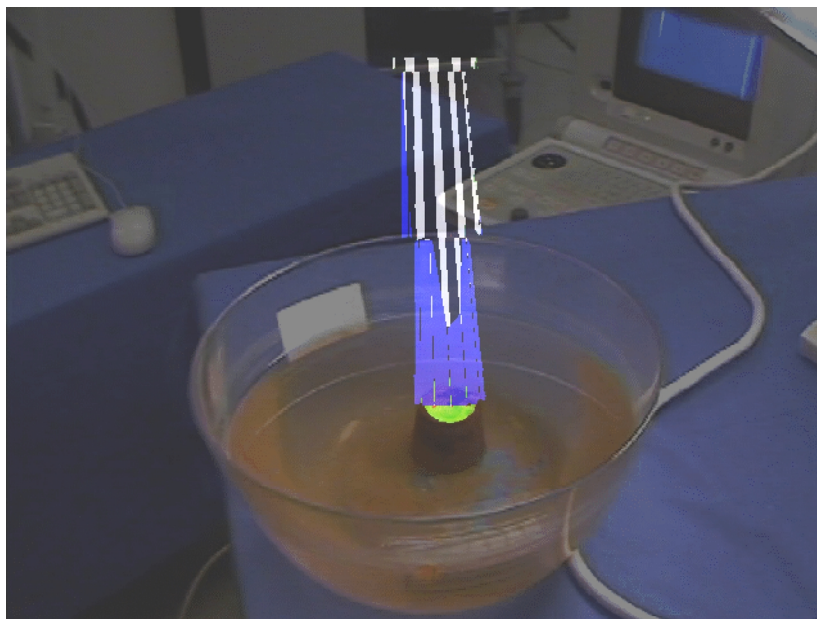
Deux types de représentation 3D sont possibles : stéréoscopique ou monoscopique.

Le principe de la visualisation stéréoscopique s'appuie sur la production d'une image pour l'œil gauche et d'une image pour l'œil droit. Les deux images sont entrelacées :

- ligne par ligne : les lignes paires correspondent à l'image gauche, les lignes impaires à l'image droite.
- image par image : à chaque instant, une seule image (gauche ou droite) est affichée devant l'œil correspondant. Rien n'est affiché devant l'autre œil. Les images gauches et droites sont ainsi affichées alternativement à une fréquence de 120 Hz.

Avec l'une ou l'autre forme d'entrelacement des images gauche et droite, le mécanisme du casque permet seulement à l'œil gauche (resp. droit) de percevoir l'image gauche (resp. droite). Lors de la perception, il y a alors reconstruction mentale de la scène, aboutissant à la perception de la profondeur du champ visuel. La réalisation de l'image gauche consiste à construire la scène de guidage dans le repère associé au viseur, puis de positionner une caméra OpenGL aux coordonnées de l'œil gauche, exprimées dans le repère du viseur, en l'orientant selon la direction donnée par le casque (le dernier point du calibrage) : ceci fournit alors l'image gauche. On procède de même pour l'image droite. L'écart interpupillaire de chaque utilisateur n'étant pas toujours le même, nous avons en outre ajouté un facteur permettant de réduire ou d'accroître la distance entre les deux images. Ce facteur est ajusté par l'utilisateur de façon à obtenir une vision stéréoscopique confortable, par simple manipulation des flèches "gauche" et "droite" du clavier, ou des boutons gauche et droit de la souris.

Dans le cas de la visualisation monoscopique de la scène 3D, une seule image est requise. Une seule caméra OpenGL nécessite donc d'être positionnée dans la scène 3D. Nous avons choisi de prendre le milieu des deux écrans comme position pour cette caméra. Le résultat monoscopique est illustré par la Figure 78, qui montre une photographie prise derrière l'écran gauche du casque. Les données de guidage sont clairement visibles dans le casque, au travers duquel on distingue également la réalité, ici un saladier symbolisant l'objet réel de la tâche (le patient) et un plot en métal symbolisant la cible à atteindre (l'épanchement).



**Figure 78** *Vue au travers du casque : superposition des données de guidage virtuelles à la vue du monde réel (saladier et bloc métallique).*

## 4 Evaluation Expérimentale

Dans le paragraphe 2 nous avons effectué une analyse ergonomique prédictive de *CASPER-V2*. Suite à la réalisation logicielle de cette version et si l'on se reporte au cycle de vie du logiciel, il convient de mener des tests. Les tests unitaires et d'intégration ont été réalisés pour vérifier la précision des données de guidage fournies. De plus, *CASPER-V2*

s'appuie entièrement sur le noyau fonctionnel existant dans *CASPER*, déjà testé et validé sur cette première version. Les changements notables entre les deux versions résident dans la présentation des données, autrement dit dans l'interface. Les tests les plus importants doivent donc concerner l'utilisabilité de cette nouvelle interface.

Nous avons tout d'abord présenté au chirurgien avec qui nous collaborons les solutions conçues et développées, compte tenu des différents référentiels de présentation. Cette première étape avait pour but de nous assurer de l'acceptation de *CASPER-V2* par l'utilisateur final, un des critères de l'utilisabilité définis par la FDA ("*Food and Drug Administration*") et présenté au chapitre III.

Devant la très bonne réaction de notre chirurgien, prêt à utiliser *CASPER-V2* en bloc opératoire, nous avons souhaité élargir l'évaluation. Nous avons alors fait tester notre système à un échantillon d'utilisateurs plus large, afin de recueillir un nombre important de données, nécessaire pour mener des études statistiques qui aient un sens. Le but de cette expérimentation est d'évaluer le critère de facilité d'utilisation ("*ease of use*"), qui, comme le critère d'acceptation par l'utilisateur, fait partie des quatre composantes de l'utilisabilité identifiées par la FDA.

L'évaluation expérimentale a été menée en collaboration avec le Laboratoire de Psychologie Expérimentale (LPE) de Grenoble (en particulier avec le Professeur Ohlmann, Dominique Ribeiro et Joëlle Neimer). L'objectif de cette expérimentation est d'évaluer l'impact de notre système sur le comportement de l'utilisateur. L'expérimentation doit révéler si certaines combinaisons des supports d'affichage (écran ou casque), des référentiels de présentation (lié à l'aiguille, à la trajectoire ou à l'utilisateur) et des représentations (croix et jauge ou cône) permettent à l'utilisateur de réaliser la tâche de façon plus favorable. Nous prenons également en compte le profil comportemental des utilisateurs, notion développée dans le paragraphe suivant. La tâche considérée consiste à reproduire une trajectoire en procurant à l'utilisateur des données de guidage. Au travers de l'étude du comportement de l'utilisateur, c'est également l'analyse de la qualité de la réalisation de la tâche qui est visée. Nous présentons dans les paragraphes suivants le cadre expérimental, la méthodologie, puis les résultats obtenus suite à une analyse statistique des données obtenues.

## 4.1 Cadre expérimental

### 4.1.1 Utilisateurs

Les utilisateurs finaux de *CASPER-V2* sont des chirurgiens cardiaques. Leur faible nombre, au CHU de Grenoble trois seulement sont impliqués dans ce projet, et leur disponibilité très réduite ne nous permettaient pas de mener une évaluation expérimentale uniquement avec des chirurgiens. Nous avons donc testé *CASPER-V2* et ses différents modes de représentation avec 12 personnes étrangères au domaine chirurgical et même plus généralement étrangères au domaine clinique.

### 4.1.2 Tâche

Outre le fait que les utilisateurs sont étrangers au domaine chirurgical, les conditions de réalisation de la tâche ne peuvent demeurer similaires à celles en bloc opératoire. Tout d'abord, la longueur moyenne de la trajectoire de l'aiguille de ponction ne dépasse guère

les quelques centimètres en situation réelle. Cette manipulation repose donc sur une nécessaire habileté dans le geste, dont seuls les chirurgiens disposent du fait de leur formation et de leur expérience. Par ailleurs, en situation réelle l'aiguille est introduite dans le corps du patient, procurant ainsi un support aux mouvements. Une situation analogue n'était pas envisageable dans notre expérimentation. Enfin, les chirurgiens sont confrontés à des facteurs de stress que l'expérimentation ne pouvait pas non plus recréer.

Aussi, au cours de notre expérimentation, la tâche confiée à nos utilisateurs consiste à reproduire une trajectoire environ 10 fois plus longue que la trajectoire suivie en situation réelle. Les tremblements et imprécisions de nos utilisateurs par rapport à ceux des chirurgiens, sont donc absorbés par ce changement d'échelle.

De plus, dans le cas de l'écran (*CASPER*, version 1) le chirurgien perçoit les données de guidage sur un écran positionné à la tête du patient. La réalisation de la ponction péricardique nécessite donc de la part du chirurgien de regarder cet écran. Pour aligner au mieux l'aiguille réelle avec la trajectoire prédéfinie, le chirurgien applique régulièrement des contraintes sur l'aiguille afin de la réaligner avec la trajectoire planifiée. Tant que l'aiguille n'a pas pénétré le corps du patient, ces mouvements sont totalement libres. Par contre, une fois l'extrémité de l'aiguille introduite dans le patient, les mouvements imprimés par le chirurgien peuvent l'amener à tordre cette aiguille : le modèle de l'aiguille par rapport aux diodes permettant sa localisation devient alors erroné et les indications de guidage sont faussées. En conséquence, le chirurgien doit régulièrement regarder l'aiguille afin de s'assurer que celle-ci n'a subi aucune déformation. Son regard fait donc l'aller-retour entre l'écran et le champ opératoire. Afin de reproduire cette contrainte lors de l'expérimentation, nous avons généré un signal sonore à intervalles de temps aléatoires, compris entre 7 et 15 secondes : à chaque signal sonore l'utilisateur doit regarder la pseudo-aiguille et confirmer oralement l'avoir vue. La personne chargée de mener l'expérimentation s'assurait que l'utilisateur avait respecté cette consigne.

### 4.1.3 Dispositif technique

Le matériel utilisé se compose d'un PC auquel est connecté un localisateur 6D : le *Flock of Bird* [FOB], basé sur la détection de champs magnétiques. Un émetteur est fixé sur la table et crée un champ magnétique hémisphérique. Le récepteur, accolé à une représentation grossière d'une aiguille chirurgicale, est manipulé par l'utilisateur et capte à chaque instant les caractéristiques du champ magnétique, qui varient selon la position et l'orientation du récepteur. Les données captées sont alors transmises au PC via un port série classique (type RS-232). De ces données, notre application de test, une version adaptée de *CASPER*, déduit la position de la pseudo-aiguille en 6D (position et orientation) et met à jour l'affichage pour permettre à l'utilisateur de comparer la position de sa pseudo-aiguille avec la trajectoire à reproduire. Cette trajectoire a été fixée logiciellement dans notre application de test, pour les besoins de l'expérimentation. Les instructions et les informations fournies aux utilisateurs sont jointes en Annexe B.

### 4.1.4 Mesures

Chaque utilisateur a huit scénarii à réaliser. Ces scénarii sont décrits dans le paragraphe suivant. Pour chaque personne, huit fichiers de données sont donc générés par le système informatique. Les données contenues dans chacun de ces fichiers correspondent à la distance (en millimètres) en  $x$ ,  $y$  et  $z$  de l'extrémité de l'aiguille à la

trajectoire à reproduire. La fréquence de rafraîchissement de l'image étant de l'ordre de 20 Hz, 20 lignes de données sont générées par seconde pour un scénario et un utilisateur.

## 4.2 Méthodologie expérimentale

### 4.2.1 Facteurs ou variables indépendantes

Dans cette étude, cinq facteurs, appelés aussi variables indépendantes, sont étudiés. Le premier d'entre eux est un **facteur comportemental** propre à chaque individu et caractérisant la manière avec laquelle l'individu repère dans un espace tridimensionnel. Cette typologie des individus [Ohlmann 90], réalisée grâce à un dispositif appelé "*Rod and Frame Test*" incliné (RFT) [Oltman 68], met en évidence trois groupes :

- **les double-dépendants (2D)** : ce premier groupe est celui des sujets référencés sur le couple géométrie/cinématique. Pour se repérer ces individus se basent principalement sur les informations visuelles statiques ou en mouvement provenant du monde extérieur.
- **les double-indépendants (2I)** : ce second groupe, à l'inverse, est composé de sujets se basant sur les informations liées aux forces gravito-inertielles, émises et perçues par l'individu dans le monde réel. Les individus de ce groupe sont aussi appelés "gravitaires".
- **les mixtes (M)** : ce troisième groupe est composé de sujets se basant tantôt sur les informations visuelles, tantôt sur les informations liées aux forces gravito-inertielles.

Dans cette expérimentation, nous n'avons choisi que des sujets 2D et 2I qui sont les extrêmes les plus intéressants dans le cadre d'une étude des référentiels spatiaux. Parmi nos 12 utilisateurs, 7 sont des double-dépendants et 5 sont des double-indépendants.

Le second facteur étudié est le **support d'affichage** utilisé pour rendre perceptibles les données de guidage par le chirurgien. Deux dispositifs sont testés : il s'agit de l'écran, qui est le dispositif utilisé dans la première version de *CASPER*, et du casque semi-transparent de *CASPER-V2*.

Un troisième facteur est la **représentation des données** fournies aux utilisateurs. Comme pour le dispositif, deux représentations sont testées. La première est celle utilisée dans la première version de *CASPER* : il s'agit d'un système de viseur basé sur trois croix, dont le fonctionnement est décrit plus précisément dans le chapitre VI. La seconde représentation faisant l'objet de l'évaluation est le cône 3D de *CASPER-V2*. La conception de cette solution de présentation a été développée au paragraphe 1.2.

Le facteur suivant a également été évoqué précédemment : il s'agit du **référentiel de présentation** adopté par le système pour rendre perceptible les données. Nous avons mentionné au paragraphe 1.2 trois référentiels de présentation possibles pour présenter le cône à l'utilisateur : le référentiel de présentation lié à la trajectoire (Figure 61), le référentiel de présentation lié à l'aiguille (Figure 62) et le référentiel de présentation lié à l'utilisateur. Dans cette expérimentation, nous n'avons étudié que les deux premiers référentiels de présentation envisagés. Au moment de l'expérimentation, l'utilisation du troisième référentiel de présentation n'était pas encore envisageable. En effet la procédure de calibrage du casque, décrite au paragraphe 3.3 n'était pas encore établie et il n'était



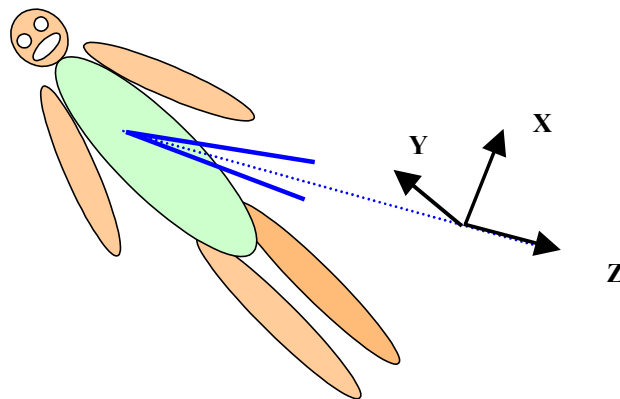
alors pas possible d'aligner convenablement les données virtuelles au monde réel. Pourtant ce dernier référentiel de présentation est sans doute le mieux adapté à la tâche considérée, compte tenu des résultats de l'analyse ergonomique faite au paragraphe 2. Ensuite, quelle que soit la représentation adoptée, deux points de vue sont testés :

- le référentiel de présentation lié à la trajectoire, la représentation de la trajectoire affichée demeurant alors immobile (paragraphe 3.2.1),
- le référentiel lié à l'aiguille, la représentation de l'aiguille affichée à l'écran restant immobile ou étant même supprimée (paragraphe 3.2.2).

Enfin, l'étude des mouvements de la pseudo-aiguille se fait suivant les trois axes  $x$ ,  $y$  et  $z$ , formant un repère orthogonal où l'axe  $z$  est confondu avec la trajectoire à reproduire.

#### 4.2.2 Mesures ou variables dépendantes

Les variables dépendantes correspondent aux données dont les valeurs produites par les sujets peuvent être influencées par les facteurs ou variables indépendantes précédemment décrits. Il s'agit donc d'identifier les données dont nous souhaitons étudier l'évolution en fonction des facteurs. Trois variables constituent ce que nous appellerons dans la suite de ce paragraphe le signal. Concrètement le signal traduit les mouvements provoqués par l'utilisateur sur la pseudo-aiguille. Il est décrit par trois composantes que nous dénotons par  $x$ ,  $y$  et  $z$ , et qui représentent les mouvements de l'extrémité de la pseudo-aiguille dans trois plans orthogonaux. La Figure 79 illustre la position de ces trois composantes par rapport à la trajectoire définie. Ces trois variables dépendantes représentent donc la distance à chaque instant entre l'extrémité de l'aiguille et chacun des axes  $x$ ,  $y$  et  $z$  du repère associé à la trajectoire.



**Figure 79** Représentation des composantes  $x$ ,  $y$ ,  $z$  par rapport à la trajectoire prédéfinie.

Les données enregistrées selon l'axe  $x$  correspondent à des mouvements latéraux par rapport à la trajectoire. Les données enregistrées selon l'axe  $y$  correspondent à des mouvements de haut en bas par rapport à la trajectoire et enfin les données enregistrées selon l'axe  $z$  correspondent à des mouvements de translation le long de la trajectoire.

Pour l'analyse des données, les trois composantes du signal global sont converties par transformées de Fourier en spectres de puissance. Les variables dépendantes étudiées sont alors l'intégrale de chaque spectre de puissance. Cette variable représente alors l'énergie dépensée par l'utilisateur pour accomplir sa tâche. Nous avons choisi d'adopter un modèle multiplicatif pour rendre compte des effets de mouvement [Droulez 89]. Cette modélisation sous-entend que les effets causés par les différents facteurs se multiplient



entre eux pour agir sur le comportement de l'utilisateur. Cette hypothèse se justifie du fait que la composition des forces auxquelles est soumis l'utilisateur ne correspond pas uniquement à des forces linéaires. Des forces de Coriolis entrent en jeu, impliquant de ce fait des effets multiplicatifs. Les trois composantes du signal ( $x$ ,  $y$  ou  $z$ ) pourront donc être notées  $X_{ijkl}$ ,  $Y_{ijkl}$  ou  $Z_{ijkl}$ . Chaque indice représentant un des quatre facteurs identifiés au 4.2.1. On a alors par exemple :

$$\begin{aligned} X_{ijkl} = & C_i \times S_j \times R_k \times P_l \times \\ & C_i S_j \times C_i R_k \times C_i P_l \times S_j R_k \times S_j P_l \times R_k P_l \times \\ & C_i S_j R_k \times C_i S_j P_l \times C_i R_k P_l \times S_j R_k P_l \times C_i S_j R_k P_l \times \mathcal{E}. \end{aligned}$$

- où  $X_{ijkl}$  ..... représente la composante du mouvement selon  $x$ ,  
 $C_i$  ..... est l'effet dû au comportement des différents types d'individu sur le mouvement ( $C_1$  pour les dépendants,  $C_2$  pour les indépendants),  
 $S_j$  ..... l'effet dû aux différents supports sur le mouvement ( $S_1$  pour l'écran,  $S_2$  pour le casque),  
 $R_k$  ..... l'effet dû aux différentes représentations sur le mouvement ( $R_1$  pour le cône,  $R_2$  pour les croix),  
 $P_l$  ..... l'effet dû aux différents référentiels de présentation sur le mouvement ( $P_1$  pour le référentiel de présentation lié à l'aiguille,  $P_2$  pour le référentiel de présentation lié à la trajectoire),  
 $C_i S_j$ , etc. .... l'interaction entre les facteurs,  
 $\mathcal{E}$  ..... la part due à l'erreur de mesure.

Cependant, la méthode d'analyse statistique proposée par nos collègues du Laboratoire de Psychologie Expérimentale de Grenoble, et que nous détaillons plus précisément dans le paragraphe 4.3 repose sur une décomposition linéaire, c'est-à-dire additive, de la variance des facteurs. Il est donc nécessaire de passer d'un modèle multiplicatif à un modèle additif entre les effets des facteurs. Une solution consiste à effectuer une transformation logarithmique en base 10 des données [Winer 62]. Le choix se porte sur cette base afin d'exprimer en décibels les gains dus à la modification d'un facteur. Chaque composante du mouvement est donc maintenant exprimée par :

$$\begin{aligned} \log_{10} X_{ijkl} = & \log_{10} C_i + \log_{10} S_j + \log_{10} R_k + \log_{10} P_l + \\ & \log_{10} C_i S_j + \log_{10} C_i R_k + \log_{10} C_i P_l + \log_{10} S_j R_k + \log_{10} S_j P_l + \log_{10} R_k P_l + \\ & \log_{10} C_i S_j R_k + \log_{10} C_i S_j P_l + \log_{10} C_i R_k P_l + \log_{10} S_j R_k P_l + \\ & \log_{10} C_i S_j R_k P_l + \log_{10} \mathcal{E}. \end{aligned}$$

### 4.2.3 Hypothèses concernant le signal

Nous présentons dans ce paragraphe les hypothèses émises concernant chacun des facteurs présentés au paragraphe 4.2.1. L'analyse des résultats de l'expérimentation doit confirmer ou infirmer ces hypothèses.

Considérons tout d'abord la **typologie** des sujets : dépendants (ou double-dépendants) et indépendants (ou double-indépendants). Les dépendants sont particulièrement influencés par les informations cinématiques, dans notre cas l'affichage des données. Or la tâche qu'ils doivent accomplir consiste à déplacer la pseudo-aiguille en fonction des données de guidage affichées. Ils sont donc amenés à transformer des informations cinématiques en informations haptiques, ce qui nécessitera pour eux un effort cognitif élevé. En effet ils doivent traduire des informations exprimées dans "une langue" qui leur est familière (langage cinématique, la vue), vers "une langue" qui leur est moins favorable (langage cinétique, les mouvements). Au contraire, les indépendants ont une plus grande aptitude à manipuler des informations de type cinétique, par exemple le déplacement de la pseudo-aiguille, que celles de type cinématique, par exemple les informations visuelles. Ils se trouvent donc dans une situation inverse des précédents : ils doivent interpréter une information codée dans "une langue" peu familière (langage cinématique, la vue) pour les retranscrire dans une langue qu'ils maîtrisent mieux (langage cinétique, les mouvements). Cette différence est à rapprocher des exercices littéraires que sont le thème, traduction d'un texte en langue maternelle vers une langue étrangère, et la version, traduction d'un texte en langue étrangère vers la langue maternelle). Le thème est toujours plus difficile à réaliser que la version. Par analogie, on peut donc s'attendre à ce que les dépendants aient plus de mal à retranscrire les informations de guidage affichées par le système en mouvements physiques de la pseudo-aiguille.

Au niveau du **support d'affichage** des données, l'écran ou le casque semi-transparent, l'hypothèse est que le casque devrait induire moins de variations dans le signal enregistré. En effet, le casque étant transparent, l'utilisateur peut régulièrement visualiser la pseudo-aiguille au travers du casque sans avoir à déplacer la tête. En revanche, avec l'écran, le sujet doit tourner la tête, d'où une augmentation des perturbations induisant probablement un accroissement de la difficulté à réaliser la tâche. Cette hypothèse est à mettre en relation avec notre analyse prédictive de *CASPER-V2* (paragraphe 2) où nous avons souligné que le casque semi-transparent résout les problèmes d'incompatibilité et de discontinuité perceptuelles identifiés dans *CASPER*.

Quant à la **représentation** utilisée pour présenter les données de guidage, les croix ou le cône, l'hypothèse est que le cône devrait être plus facilement interprétable et par conséquent diminuer l'énergie dépensée par l'utilisateur pour reproduire la trajectoire. En effet, il s'agit d'une analogie à la trajectoire que les sujets ont à reproduire. Les croix sont plus abstraites, et nécessitent d'avoir une représentation cognitive stable pour comprendre et interpréter la représentation : quelle couleur de croix est associée à telle partie de la pseudo-aiguille ? Dans quel sens doit-on déplacer tel point de la pseudo-aiguille pour corriger ? De ce fait, le cône devrait induire moins de mouvements de la part des sujets puisqu'il s'avère être plus représentatif. Cette hypothèse est à mettre en relation avec notre analyse prédictive de *CASPER-V2* (paragraphe 2) où nous avons souligné que le cône réduit les facteurs d'incompatibilité et de discontinuité cognitives.

Concernant le **référentiel de présentation** des données de guidage, référentiel de présentation lié à l'aiguille ou référentiel de présentation lié à la trajectoire, notre hypothèse est que le référentiel aura un impact lié à la typologie des utilisateurs. Les

dépendants devraient obtenir de meilleurs résultats avec le référentiel de présentation lié à la trajectoire. En effet, se basant sur les données visuelles perçues dans le monde réel pour se repérer dans l'espace, les utilisateurs dépendants peuvent être favorisés par la présentation de données directement liés à l'orientation et la position de la pseudo-aiguille dans l'espace. Il semble donc que ce soit dans le référentiel de présentation lié à la trajectoire que les données présentées correspondent le mieux à leurs besoins en terme d'orientation dans l'espace. Au contraire, les indépendants se basant sur les forces gravito-inertielles qu'ils émettent et perçoivent, il est envisageable que le référentiel de présentation lié à l'aiguille, dans lequel les données liées à la position et l'orientation de la trajectoire sont affichées par rapport à la pseudo-aiguille, reflète mieux les implications des forces qu'ils ont effectivement appliquées sur cette pseudo-aiguille.

Enfin terminons avec les **axes**. L'analyse de l'axe  $z$  ne représente que très peu d'intérêt. En effet cet axe est confondu avec la trajectoire à reproduire. Des variations selon cet axe ne représentent donc que des avancées ou des retraits de la pseudo-aiguille. Par conséquent, quelque soit la représentation choisie l'interprétation de cette dimension est évidente. Par contre les variations selon les axes  $x$  et  $y$  devraient être fortement corrélées. En effet tout déplacement selon l'un de ces deux axes est représenté de la même manière. Les difficultés d'interprétation et de correction de la position de la pseudo-aiguille selon ces deux axes devraient donc être semblables.

### 4.3 Analyse statistique du signal global

Pour la variable étudiée, et ses trois composantes, nous avons procédé à une étude statistique inférentielle : ANOVA [Abdi 87]. ANOVA est un outil d'analyse statistique de référence qui se base sur l'analyse de la variance des résultats obtenus. Cette analyse inférentielle vise à établir l'existence d'effets sur la réalisation de la tâche, liés à des facteurs ou à des combinaisons de facteurs, parmi ceux identifiés au paragraphe 4.2.1. Pour une situation donnée, une différence entre la moyenne des résultats obtenus dans cette situation et la moyenne globale des résultats, peut signifier l'existence d'un effet du facteur caractérisant alors la situation. Cet effet est considéré comme étant significatif si la statistique du test d'analyse des variances, notée  $p$ , est inférieure ou égale au seuil fixé. Habituellement ce seuil est fixé à 0.05. Dans ce cas, la probabilité d'annoncer un effet alors qu'il n'y en a pas (effet du hasard) est de 5%. Nous présentons dans le paragraphe suivant les méthodes d'analyse d'un effet principal et d'un effet d'interaction.

#### 4.3.1 Effets principaux et effets d'interaction

Considérons deux facteurs  $F1$  et  $F2$ , ayant chacun deux modalités notées  $a$  et  $b$ , soit  $F1a$ ,  $F1b$  et  $F2a$ ,  $F2b$ . Dans notre expérimentation,  $F1$  et  $F2$  représentent par exemple le support d'affichage et le référentiel de présentation. Les modalités de  $F1$ ,  $F1a$  et  $F1b$  sont respectivement l'écran et le casque. Les modalités de  $F2$ ,  $F2a$  et  $F2b$  sont respectivement le référentiel de présentation lié à l'aiguille et le référentiel de présentation lié à la trajectoire.

Tester l'influence de ces facteurs sur l'utilisateur induit quatre situations à faire réaliser à l'utilisateur, résultant de la combinaison des modalités des facteurs. Les situations obtenues sont donc :

- $F1a - F2a$ ,  $F1a - F2b$ ,
- $F1b - F2a$ ,  $F1b - F2b$ .

Chaque situation est testée par tous les utilisateurs et pour chacune d'elle il est possible de calculer la moyenne des valeurs prises par la variable dépendante sur l'ensemble des utilisateurs. Dans la Figure 80 ces moyennes sont notées  $m_{XY}$ , où X est une modalité du facteur F1 (F1a ou F1b) et Y une modalité du facteur F2 (F2a ou F2b). La notation  $m.Y$  (respectivement  $m.X$ .) représente la moyenne des résultats obtenus pour une modalité Y (respectivement X) du facteur F2 (respectivement F1) indépendamment de la modalité du facteur F1 (respectivement F2) utilisée. On note  $m..$  la moyenne des valeurs prises par la variable dépendante sur l'ensemble des situations possibles, tous utilisateurs confondus.

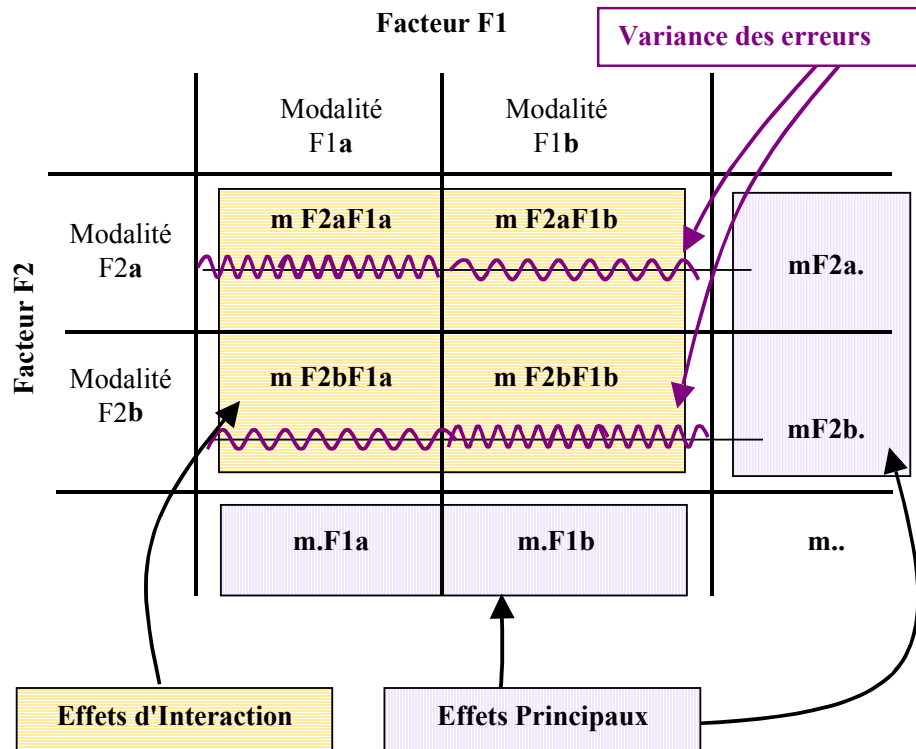


Figure 80 Schéma de principe de l'analyse de la variance.

Pour un facteur, un effet principal est mis en évidence par l'analyse des variances, alors l'analyse des moyennes obtenues dans les cases de la Figure 80 hachurées verticalement, c'est-à-dire pour chaque modalité de ce facteur, permet de déduire la modalité favorable à la réalisation de la tâche. Dans le cas où le facteur a seulement deux modalités, la comparaison de ces moyennes est une comparaison simple qui nous permet d'identifier la modalité du facteur pour lequel la tâche est la mieux réussie. Dans notre cas, la variable mesurée étant l'énergie dépensée par l'utilisateur, la modalité obtenant la moyenne la plus faible sera considérée comme étant la plus favorable à la réalisation de la tâche.

Si le facteur a plus de deux modalités, il convient de mener une analyse comparée planifiée, afin de déterminer l'aspect significatif ou non de la différence des résultats observés sur les modalités du facteur. L'analyse comparée planifiée met alors en évidence quelle modalité est responsable de l'effet principal observé sur le facteur.

Si aucun effet principal d'un facteur n'est identifié, il se peut toutefois que la combinaison de deux facteurs ait une influence sur la réalisation de la tâche. Aussi, il est nécessaire de réaliser un test d'analyse de la variance, c'est-à-dire de calculer p, pour

toutes les combinaisons de facteurs envisageables. Chaque fois que  $p \leq 0.05$ , on dira qu'il existe un effet significatif dû à l'interaction entre les facteurs de la combinaison étudiée. Une analyse comparée planifiée est alors nécessaire pour déterminer pour chaque modalité d'un facteur, la combinaison optimale des modalités des autres facteurs impliqués. Dans le cas de la combinaison des facteurs F1 et F2, ceci consiste à mener une analyse de la variance pour F1 avec F2a fixée, puis avec F2b fixée. On ne considère donc que les mesures d'une ligne du tableau de la Figure 80 au lieu de toutes les mesures. A l'issue du test d'analyse de la variance, un contraste est donc établi entre la combinaison de F2a avec F1a et F1b, ainsi qu'un contraste entre la combinaison de F2b avec F1a et F1b. L'analyse comparée planifiée selon F2a fournit une valeur de  $p$  pour la première ligne qui, si elle est inférieure ou égale à 0.05, permet d'identifier la modalité de F1 la plus favorable à la réalisation de la tâche quand F2a est présent. De la même façon, on pourra identifier la modalité de F1 la plus favorable à la réalisation de la tâche quand F2b est présent. Dans ce cas, l'étude s'appuie sur les moyennes des cases du tableau de la Figure 80 qui sont hachurées horizontalement.

Pour toute modalité identifiée comme étant plus favorable à la réalisation de la tâche qu'une autre, le rapport de Fischer, noté  $F$ , permet alors d'établir dans quelle mesure la modalité identifiée est préférable à l'autre. En effet, le rapport de Fischer est égal au rapport de la variance des moyennes sur la variance des erreurs. Ce rapport représente donc l'apport moyen d'une modalité, F1a par exemple, par rapport à l'autre, F1b, compte tenu des erreurs, c'est-à-dire des variations constatées lors de l'utilisation de la modalité F1a par l'ensemble de la population testée.

#### 4.3.2 Identification des effets

Nous avons, dans un premier temps, évalué les effets des cinq facteurs, typologie des utilisateurs, support d'affichage, type de représentation, référentiel de présentation et axes, ainsi que ceux de toutes les interactions possibles entre ces cinq facteurs. Il en résulte donc 31 cas. Parmi ces 31 situations, nous n'avons ensuite retenu que celles pour lesquelles l'effet observé était significatif, c'est-à-dire produisant un  $p \leq 0.05$ , ainsi que celles présentant une tendance, c'est-à-dire dont l'effet est presque significatif ( $0.05 \leq p \leq 0.1$ ).

Les résultats obtenus montrent que parmi les cinq facteurs pris en compte dans cette analyse statistique, quatre ont un effet significatif. Ce sont des effets principaux sur le comportement de l'utilisateur lors de la réalisation de la tâche. En effet, les résultats de cette étude ne font pas apparaître clairement d'effet liés à la typologie des utilisateurs.

L'analyse des interactions montrent que deux combinaisons ont des effets significatifs sur le comportement de l'utilisateur. Il s'agit d'une part de la combinaison de la représentation (croix ou cône) avec les déplacements de l'aiguille selon les axes  $x$ ,  $y$  ou  $z$ , et d'autre part de la combinaison de ces deux facteurs (représentation et axe) avec le référentiel de présentation dont les modalités sont aiguille et trajectoire.

Enfin, deux interactions présentent une tendance, c'est-à-dire ont un effet quasi-significatif sur le comportement de l'utilisateur. Il s'agit de la combinaison de la représentation et du référentiel de présentation, ainsi que de la combinaison de la

typologie des utilisateurs, du support d'affichage, de la représentation et des mouvements selon les axes. Notons toutefois que, pour cette interaction à 4 facteurs, la valeur de  $p$  peut paraître élevée ( $p = 0.0926$ ). Etant donné que nous avons fixé comme limite représentative d'une tendance  $p = 0.1$ , ce résultat pourrait être interprété comme étant une tendance faible. Cependant, étudier l'impact de l'utilisation d'un dispositif de réalité augmentée sur le comportement de l'utilisateur n'avait encore jamais été réalisé. La limite supérieure que nous nous sommes fixée aurait donc pu être plus souple afin de nous permettre d'évaluer plus largement les effets sur le comportement de l'utilisateur. C'est pourquoi l'effet à tendance significative identifié pour l'interaction entre les quatre facteurs doit être considérée et prise en compte au même titre que les autres tendances.

Le tableau de la Figure 81 rassemble pour chacune des situations retenues la moyenne obtenue sur les carrés de l'effet (MC), la probabilité que l'effet observé soit lié uniquement au hasard ( $p$ ), ainsi que le rapport de Fischer (F).

	MC	F	niveau p
<b>Effets significatifs</b>			
<b>Des facteurs seuls :</b>			
Support	2.48	20.19	0.0012
Type de présentation	13.75	46.35	< 0.0001
Référentiel de présentation	2.05	32.29	0.0002
Axe	1.01	23.50	< 0.0001
<b>Des interactions entre :</b>			
Représentation et Axe	3.59	36.90	< 0.0001
Représentation, Référentiel de présentation et Axe	2.67	19.04	< 0.0001
<b>Tendances (Effets presque significatifs)</b>			
<b>Des interactions entre :</b>			
Représentation et Référentiel de présentation	0.74	4.41	0.0621
Typologie, Support, Représentation et Axe	0.19	2.69	0.0926

**Figure 81** Résultats obtenus pour les facteurs et les combinaisons de facteurs présentant un effet significatif ( $p \leq 0.05$ ) ou une tendance ( $0,05 \leq p \leq 0,1$ ).

Chacune de ces situations met en jeu un ou plusieurs facteurs, qui ont chacun au moins deux modalités. Par exemple, le support a un effet principal, mais ce facteur a deux modalités possibles : le casque semi-transparent et l'écran. Il convient donc de préciser les effets identifiés lors de l'analyse globale. Dans le cas d'un effet identifié pour une interaction entre deux facteurs, il convient de mener une analyse comparée planifiée comme expliqué au paragraphe 4.3.1. Par exemple, le facteur représentation a deux modalités : croix et cône ; le facteur référentiel de présentation en a deux également : aiguille et trajectoire. L'analyse comparée planifiée consiste à étudier :

- le contraste entre la combinaison du cône avec le référentiel de présentation lié à l'aiguille et le référentiel de présentation lié à la trajectoire,
- le contraste entre la combinaison de la croix avec le référentiel de présentation lié à l'aiguille et le référentiel de présentation lié à la trajectoire.

Si les effets mesurés dans ces deux comparaisons sont significatifs ( $p \leq 0.05$ ), il sera possible d'identifier un référentiel de présentation optimal pour la représentation basée sur un cône et un référentiel de présentation optimal pour la représentation basée sur les croix.

Nous présentons donc, dans les paragraphes suivants, l'analyse détaillée des quatre effets principaux identifiés ci-dessus, puis la comparaison planifiée des deux interactions ayant un effet significatif, ainsi que de l'interaction entre représentation et référentiel de présentation. Cette dernière interaction est particulièrement intéressante pour permettre d'évaluer l'impact de l'utilisation du casque et le choix de la représentation à adopter. Par contre, nous n'approfondirons pas l'analyse de l'interaction à tendance significative entre les quatre facteurs typologie des utilisateurs, support d'affichage, représentation de la trajectoire et axe. En effet, trop de facteurs sont en jeu dans cette interaction pour nous permettre de conduire une analyse : les résultats seraient beaucoup trop complexes à analyser et par conséquent les conclusions ne pourraient pas être totalement objectives.

### 4.3.3 Analyse des effets principaux

#### 4.3.3.1 *Le support*

L'analyse statistique révèle un effet significatif du support avec un  $p < 0.0011$ . La Figure 82 montre que la moyenne des logarithmes des résultats obtenus dans les situations utilisant l'écran comme support d'affichage est supérieure à celle obtenue dans les situations utilisant le casque. Nous rappelons qu'un résultat correspond à la dépense d'énergie de l'utilisateur pour réaliser la tâche (cf. paragraphe 4.2.2).

L'effet principal du support nous permet donc de conclure que l'énergie dépensée par l'utilisateur lorsqu'il visualise les données dans le casque est moins importante que celle dépensée pour réaliser la même tâche en visualisant les données sur l'écran.

Par conséquent, l'utilisation du casque permet de limiter les mouvements effectués par l'utilisateur pour reproduire la trajectoire avec la pseudo-aiguille. Le rapport de Fischer s'élève à 20.19 ce qui confirme l'avantage du casque par rapport à l'écran. Ce résultat est en accord avec nos hypothèses (cf. paragraphe 4.2.3) ainsi qu'avec les résultats de notre analyse prédictive menée sur *CASPER-V2* (paragraphe 2) établissant que l'utilisation du casque semi-transparent élude l'incompatibilité et la discontinuité perceptuelles. Comme mentionné précédemment, l'effet lié au support et établi en faveur du casque, est certainement dû au fait que le casque est transparent, ce qui évite les mouvements involontaires de l'utilisateur chaque fois qu'il doit tourner la tête en direction de la pseudo-aiguille ou de l'écran.

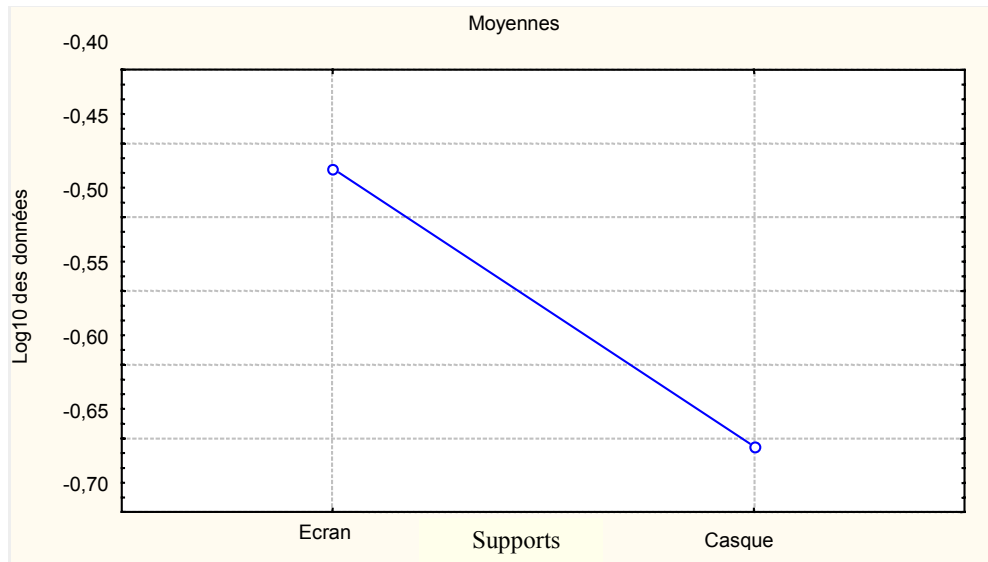


Figure 82 *Effet du support.*

#### 4.3.3.2 Le type de présentation

L'analyse statistique du facteur type de présentation, ou représentation, montre un effet significatif de ce dernier avec  $p < 0.0001$ . En reprenant le même raisonnement qu'au paragraphe précédent, la Figure 83 permet d'établir que le cône engendre moins de mouvements chez les sujets : il est donc plus efficace. Le rapport de Fischer s'élève à 46.35. La représentation basée sur un cône a donc un effet qui permet d'améliorer plus de 46 fois la réalisation de la tâche par rapport aux situations mettant en œuvre la représentation basée sur les croix. Ceci est également en accord avec nos attentes formulées dans le paragraphe 4.2.3 ainsi qu'avec le résultat de notre analyse prédictive (paragraphe 2) qui établit que les facteurs d'incompatibilité et de discontinuité cognitives sont réduits lors de l'utilisation d'une représentation basée sur un cône. L'effet lié au type de présentation, et établi en faveur du cône, est dû au fait que les croix constituent une représentation abstraite, qui nécessite un modèle mental persistant pour comprendre et interpréter la représentation. Le cône est, quant à lui, une représentation plus simple à manipuler par les utilisateurs.

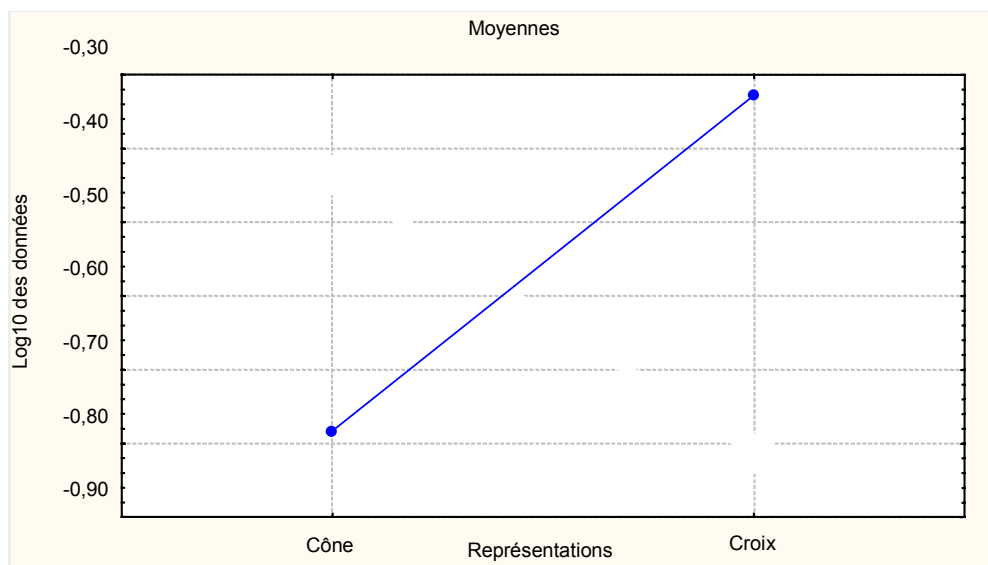


Figure 83 *Effet de la représentation.*



#### 4.3.3.3 Le référentiel de présentation

Le référentiel de présentation montre également un effet significatif avec  $p < 0.0002$ . La Figure 84 ci-dessous montre que le référentiel de présentation lié à l'aiguille induit moins de mouvements que le référentiel de présentation lié à la trajectoire. Le rapport F est là encore très élevé ( $F=32.29$ ), ce qui atteste de l'importance de cet effet principal sur la réalisation de la tâche.

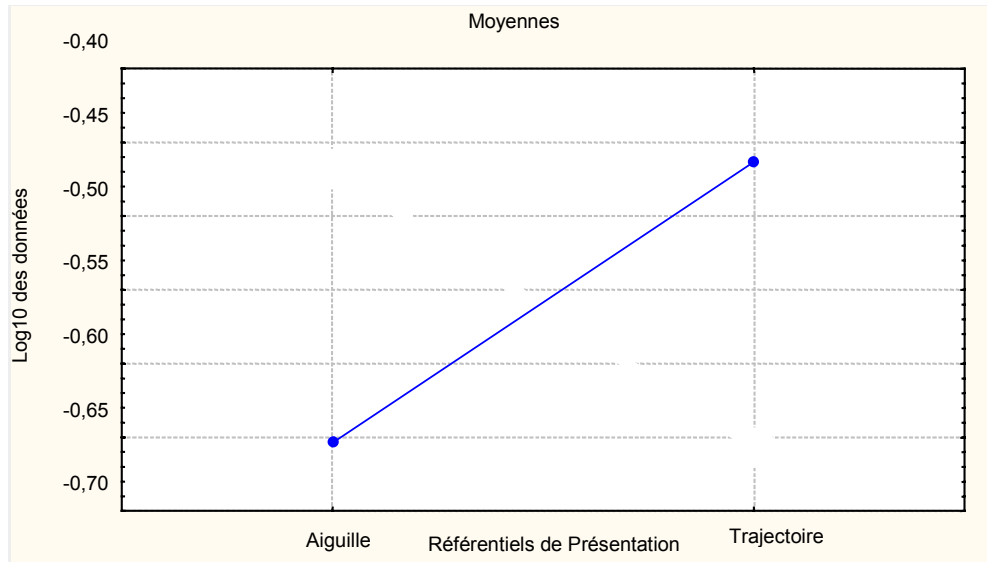


Figure 84 Effet du référentiel de présentation.

#### 4.3.3.4 Les mouvements selon les trois axes

Un effet de l'axe est également révélé par l'analyse statistique avec un  $p < 0.0001$ . Il ressort de la Figure 85 que l'axe  $x$  est l'axe pour lequel les mouvements sont les moindres, suivi de l'axe  $y$  puis de l'axe  $z$ .

Etant donné qu'il y a trois modalités pour le facteur étudié ici, il est nécessaire de mener une comparaison planifiée afin de confronter les modalités du facteur entre elles. Cette analyse révèle que la différence entre les moyennes obtenues pour les axes  $x$  et  $y$  n'est pas significative ( $p = 0.2277$ ). Par contre les différences observées entre les moyennes obtenues selon les axes  $x$  et  $z$  d'une part et  $y$  et  $z$  d'autre part, sont significatives ( $p = 0.0001$  pour  $x / z$  et  $p = 0.0006$  pour  $y / z$ ).

Le résultat de l'effet des axes sur la réalisation de la tâche confirme notre hypothèse de départ (paragraphe 4.2.3) : les problèmes d'interprétation des données de guidage se posent de la même manière en  $x$  et en  $y$ , par contre en  $z$  il s'agit d'avancer ou de reculer en suivant la trajectoire. Nous concluons donc que, d'une part, la différence entre  $x$  et  $y$  n'est pas significative et que, d'autre part, l'énergie dépensée pour accomplir le mouvement selon l'axe  $z$  est supérieure à celle dépensée selon les deux autres axes. En effet, c'est selon  $z$  que le déplacement attendu de la part de l'utilisateur est le plus grand : la portée de la trajectoire à reproduire était de l'ordre de 50 cm, alors que les marges de tolérances autour des croix ou dans le cône étaient au maximum de 5 cm.

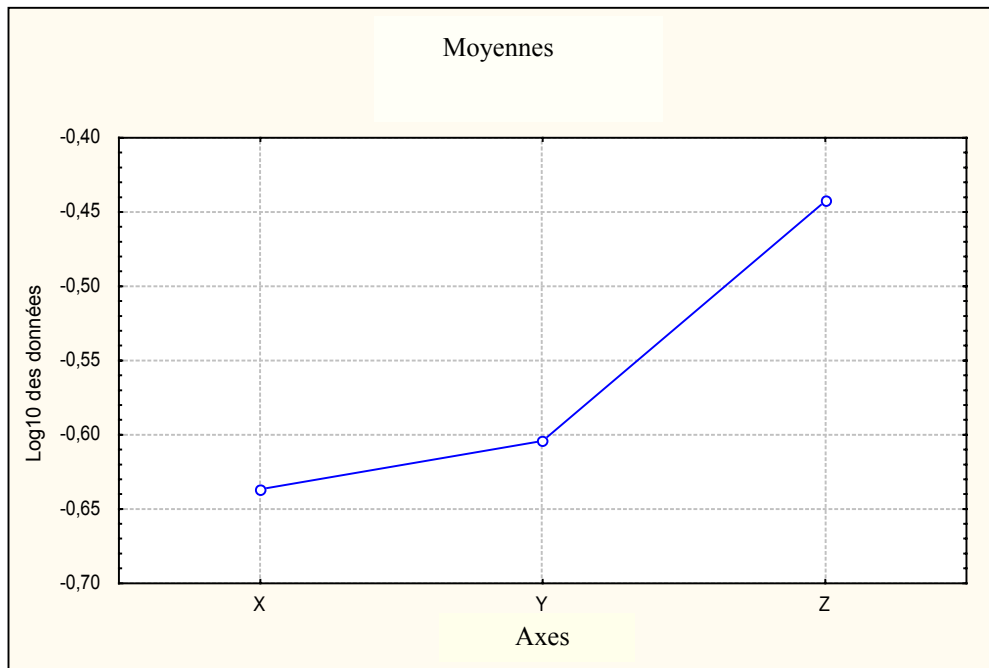


Figure 85 Effet de l'axe.

#### 4.3.4 Analyse des interactions

##### 4.3.4.1 L'interaction entre représentation et axes

L'analyse statistique a également mis en évidence un effet significatif de l'interaction entre la représentation (croix ou cône) et les axes ( $p < 0.0001$ ), illustré à la Figure 86. L'analyse comparée planifiée selon les axes, révèle :

- une différence significative selon l'axe  $x$  entre les deux types de représentations. Cette différence est nettement en faveur du cône ( $F = 16.39$ ).
- une différence significative également, selon l'axe  $z$ . Cette différence est très nettement en faveur du cône ( $F = 138.36$ ).
- aucun effet significatif ni même une tendance selon l'axe  $y$  ( $p = 0.1385$ ).

Les deux premières conclusions sont donc également en accord avec nos hypothèses du paragraphe 4.2.3. La grande différence, à l'avantage du cône, soulignée pour l'axe  $z$  et largement confortée par la valeur du rapport de Fischer ( $F = 138.36$ ) est simple à justifier : dans une représentation basée sur les croix, l'utilisateur doit regarder alternativement les croix, qui guident le geste en  $x$  et  $y$ , la pseudo-aiguille et la jauge située à droite des croix. Cette jauge représente l'avancée de la pseudo-aiguille le long de l'axe  $z$ . La perception de cette donnée est donc beaucoup plus contraignante qu'avec une représentation basée sur un cône, dans laquelle l'avancée de la pseudo-aiguille le long de l'axe  $z$  est perçue au sein de la même représentation que celle fournissant le guidage. Ce résultat expérimental est à mettre en relation avec notre analyse prédictive de *CASPER-V2* (paragraphe 2) où nous avons souligné que l'utilisation de langages basés sur une même dimension réduit les facteurs d'incompatibilité et de discontinuité cognitives.

En ce qui concerne l'axe  $y$ , aucune tendance n'est mise en évidence. Néanmoins avec  $p = 0.138$ , nous considérons tout de même que le résultat est très voisin de celui obtenu pour  $x$ , d'autant plus qu'au paragraphe précédent nous avons montré qu'aucune différence

significative ne pouvait être dégagée entre les axes  $x$  et  $y$  du point de vue de la quantité d'énergie dépensée par l'utilisateur pour réaliser son geste.

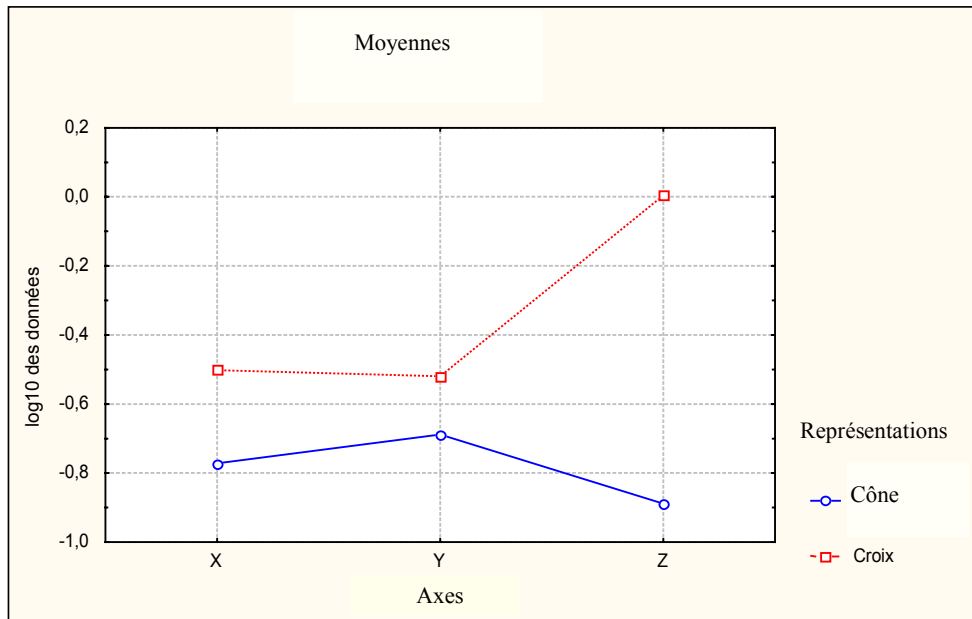


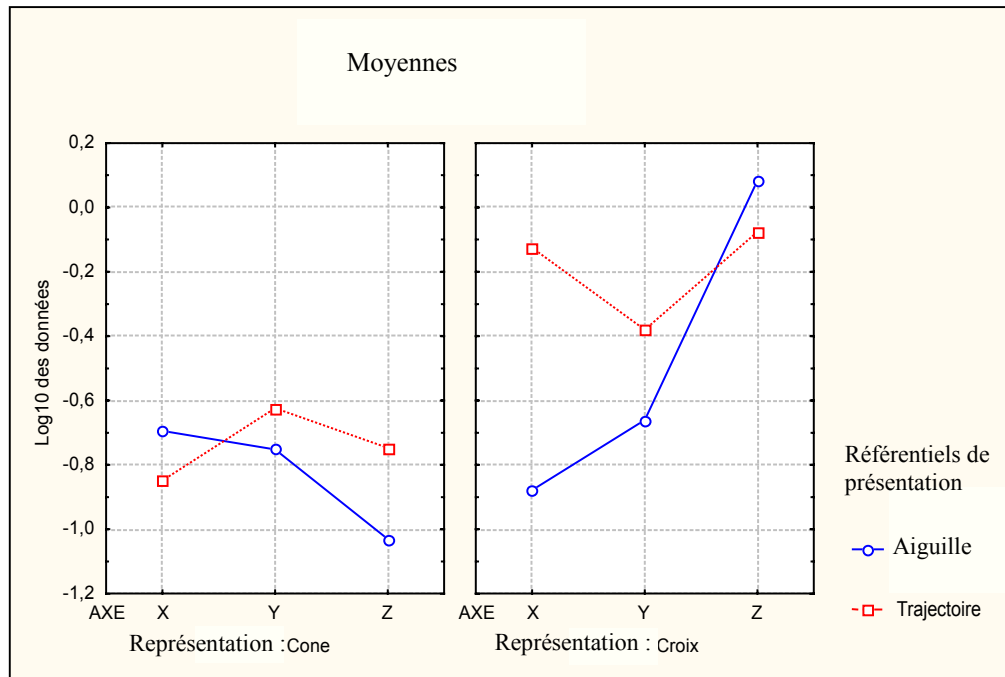
Figure 86 Interaction entre représentation et axe.

#### 4.3.4.2 L'interaction entre représentation, axes et référentiel de présentation

Le même phénomène que pour la précédente interaction est observé ici : l'interaction entre les trois facteurs, illustrée à la Figure 87, est significative ( $p < 0.00002$ ) et l'analyse comparée planifiée, nous amène à constater l'absence d'une tendance ou d'effet significatif selon l'axe  $y$  : pour la représentation basée sur un cône, la différence entre les résultats obtenus avec les deux référentiels de présentation donne  $p = 0.1542$ , et pour la représentation basée sur les croix, nous obtenons  $p = 0.1862$ . La même remarque qu'au paragraphe précédent concernant l'axe  $y$  peut être faite ici. Nous considérerons donc également les résultats selon  $y$  voisins des résultats selon  $x$ .

Selon les deux autres axes, les différences sont significatives. Les résultats obtenus selon l'axe  $x$  et  $z$  sont opposés :

- pour une représentation basée sur un cône, le référentiel de présentation lié à la trajectoire est préférable pour l'axe  $x$  ( $p = 0.034$ ,  $F = 5.99$ ), alors que c'est le référentiel de présentation lié à l'aiguille pour l'axe  $z$  ( $p = 0.0091$ ,  $F = 2.37$ ).
- pour la représentation basée sur les croix, le référentiel de présentation lié à la trajectoire est préférable pour l'axe  $z$  ( $p = 0.0395$ ,  $F = 5.60$ ), alors que c'est le référentiel de présentation lié à l'aiguille pour l'axe  $x$  ( $p = 0.00001$ ,  $F = 61.75$ ).



**Figure 87** Interaction entre la représentation, le référentiel de présentation et l'axe.

Les résultats sont donc contradictoires du point de vue des axes. Pour tirer des conclusions, appuyons nous sur les résultats précédents.

Considérons tout d'abord la représentation basée sur un cône. L'analyse de l'interaction entre la représentation et l'axe, menée au paragraphe 4.3.4.1 conclut que cette représentation est de manière indéniable ( $F = 138$ ) significativement plus favorable pour l'axe z. Or la Figure 87 tend à privilégier le référentiel de présentation lié à l'aiguille pour l'axe z avec une représentation basée sur un cône. La combinaison de ces deux résultats nous amène à conclure que le référentiel de présentation lié à l'aiguille devrait être privilégié lorsque la représentation basée sur un cône est utilisée, et ce afin de pouvoir bénéficier de l'effet de la combinaison de l'axe z et de la représentation basée sur un cône.

Pour la représentation basée sur les croix, l'analyse de l'interaction entre la représentation et les axes, menée au paragraphe 4.3.4.1 ne permet pas de mieux conclure. Il convient donc de choisir le référentiel de présentation favorisant le mieux l'axe sur lequel sont exprimées les données les plus importantes. La combinaison des modalités résultant de cette analyse réduira l'énergie dépensée pour se conformer aux indications relatives à l'axe le plus important. Dans *CASPER-V2*, l'axe z est le plus critique car si l'extrémité de l'aiguille dépasse le point cible, c'est le cœur qui est touché ! L'hémorragie ainsi provoquée est une cause de décès du patient. Dans *CASPER-V2*, nous opterions donc pour une combinaison des solutions visant à privilégier l'axe z. Il s'agit en l'occurrence d'adopter le référentiel de présentation lié à la trajectoire si la représentation est basée sur les croix (avec  $F = 5.60$ ).

Au paragraphe suivant, nous vérifions que ces hypothèses, concernant les configurations à adopter, se concrétisent également lors de l'analyse de l'interaction entre la représentation et le référentiel de présentation, sans prendre en compte l'axe.

#### 4.3.4.3 L'interaction entre représentation et référentiel de présentation

Rappelons que cette interaction, illustrée à la Figure 88, ne présente qu'une tendance à un effet significatif ( $p = 0.0621$ ), faiblement pondérée ( $F = 4.40$ ). Le référentiel de présentation lié à la trajectoire semble induire plus de mouvements que le référentiel de présentation lié à l'aiguille et cette différence semble s'amplifier avec l'utilisation d'une représentation basée sur les croix au lieu d'une représentation basée sur un cône.

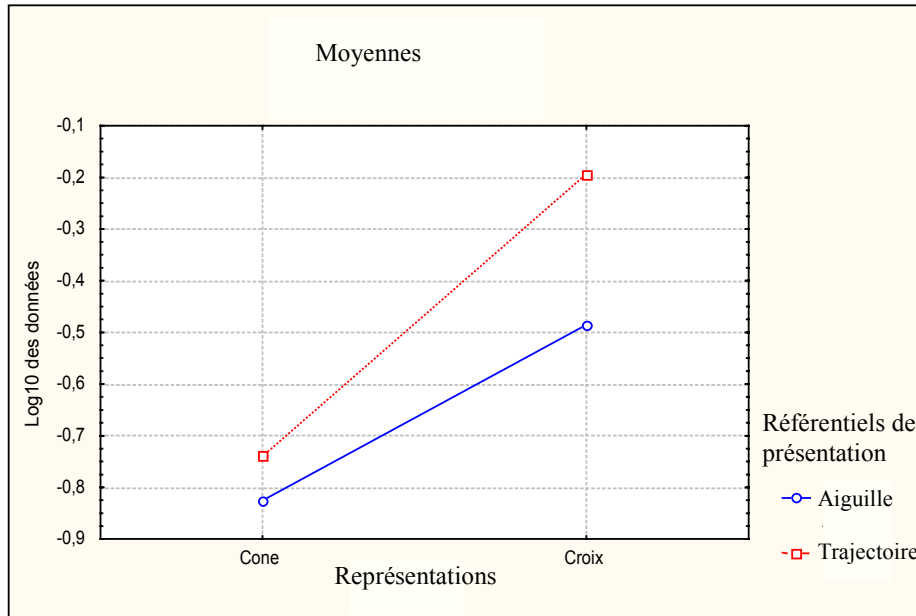


Figure 88 Interaction entre la représentation et le référentiel de présentation

L'analyse comparée planifiée met en évidence que pour une représentation basée sur les croix, la différence entre le référentiel de présentation lié à la trajectoire et celui lié à l'aiguille est significative ( $p = 0.0019$ ). Cette différence tend à privilégier le référentiel de présentation lié à l'aiguille. Cette conclusion va donc à l'encontre du choix que nous avons privilégié au paragraphe précédent. Toutefois, l'interaction entre la représentation et le référentiel de présentation présentée ici ne présente qu'une tendance significative et non un effet significatif. Ce résultat est donc à prendre avec réserve et doit être mis en relation avec les autres conclusions établies.

Pour une représentation basée sur le cône, la différence entre les deux référentiels de présentation fait apparaître une tendance significative ( $p = 0.0872$ ) également en faveur du référentiel de présentation lié à l'aiguille, puisque c'est celui qui engendre le moins de déplacements. Au paragraphe précédent, nous avons conclu de la même manière. Notons toutefois que ce résultat n'exprime qu'une tendance significative, au sein d'une interaction qui est elle-même à tendance significative. Ce résultat est donc également à considérer avec prudence.

#### 4.3.5 Bilan de l'analyse du signal global

L'analyse statistique du signal global met en évidence trois facteurs principaux permettant d'améliorer grandement la qualité de réalisation de la tâche du point de vue de l'énergie dépensée par l'utilisateur : le support d'affichage, la représentation et le référentiel de présentation.

Ces trois facteurs principaux permettent d'identifier chacun une modalité favorisant la réalisation de la tâche. Ainsi, la première conclusion est la suivante :

**L'utilisation du casque, de la représentation basée sur un cône et du référentiel de présentation lié à l'aiguille, considérés indépendamment les uns des autres, concourent à l'amélioration de la qualité de réalisation de la tâche.**

Nous avons ensuite complété l'étude de ces effets principaux par l'analyse statistique des interactions. La première conclusion issue de l'analyse de l'interaction entre la représentation et l'axe est la suivante :

**Les axes  $x$  et  $z$  sont très nettement avantagés par l'utilisation d'une représentation basée sur un cône.**

Rappelons que des déplacements selon l'axe  $x$  ou  $y$  correspondent à des déplacements de la pseudo-aiguille dans un plan perpendiculaire à la trajectoire à reproduire. Des déplacements selon l'axe  $z$  se situent quant à eux le long de la trajectoire.

Une différence entre les axes  $x$  et  $z$  apparaît lors de l'analyse de l'interaction significative entre les axes, le référentiel de présentation et la présentation. Les résultats mènent à des conclusions inverses pour les axes  $x$  et  $y$ . En effet, pour la représentation basée sur un cône, le référentiel de présentation le plus adapté selon l'axe  $x$  est celui lié à la trajectoire, alors qu'il s'agit de celui lié à l'aiguille selon l'axe  $z$ . Pour une représentation basée sur les croix, les résultats sont inversés. L'axe  $y$  n'exprime aucune tendance pour l'une et l'autre des deux représentations, mais les résultats restent très voisins de ceux de l'axe  $x$ .

Les résultats obtenus selon  $x$  et  $z$  dans cette interaction étant significatifs, mais opposés l'un par rapport à l'autre, nous nous sommes appuyés sur les résultats de l'analyse de l'interaction entre l'axe et la représentation, d'où il ressort que l'axe  $z$  est largement favorisé par une représentation basée sur un cône. Notre conclusion est donc la suivante :

**Le référentiel de présentation lié à l'aiguille lors de l'utilisation d'une représentation basée sur un cône semble favoriser la réalisation de la tâche.**

Cette conclusion est également étayée par l'analyse de l'interaction à tendance significative entre le référentiel de présentation et la présentation.

En revanche, concernant la représentation basée sur les croix, aucun résultat plus concret n'a pu être obtenu. Il faut donc établir le choix du référentiel en fonction de l'axe que l'on souhaite particulièrement privilégier.

Pour affiner ces résultats de l'analyse statistique du signal global, nous avons conduit une analyse statistique selon chaque composante du signal,  $x$ ,  $y$  et  $z$ , que nous présentons au paragraphe suivant.

#### **4.4 Analyse statistique de chaque composante du signal**

La méthode suivie est similaire à celle appliquée pour l'analyse du signal global. De l'analyse de la composante  $x$  du signal, il résulte trois effets significatifs :

- l'effet principal lié au support d'affichage ( $p = 0.0045$ ) qui prouve l'intérêt du casque sur l'écran ( $F = 13.20$ ),
- l'effet principal lié à la représentation ( $p = 0.0008$ ), qui met en évidence que le cône est moins bien adapté que la représentation basée sur les croix ( $F = 21.92$ )
- l'effet de l'interaction entre la représentation et le référentiel de présentation ( $p = 0.0374$ ).

Pour ce dernier effet significatif, l'analyse comparée planifiée a révélé que pour le référentiel de présentation lié à l'aiguille, la représentation la plus adaptée est celle basée sur les croix ( $p = 0.0001$  et  $F = 37.58$ ). Ceci renforce donc la conclusion établie lors de l'étude de l'interaction à tendance significative entre la représentation et le référentiel de présentation dans le signal global. Par contre aucun effet significatif n'est mis en évidence pour le référentiel de présentation lié à la trajectoire.

L'analyse de la composante  $y$  du signal a mis en évidence deux effets significatifs.

Le premier est lié à l'interaction entre le type de représentation et le référentiel de présentation ( $p = 0.0379$ ). L'analyse comparée planifiée ne permet pas de conclure plus précisément puisque quelque soit le référentiel de présentation considéré, l'effet n'est pas significatif.

Le deuxième effet significatif résulte de l'interaction entre quatre facteurs : typologie des utilisateurs, support d'affichage, représentation et référentiel de présentation. Nous n'avons pas approfondi plus avant cet effet étant donné qu'il n'est pas ressenti selon  $x$ . Or la force de corrélation entre ces deux axes est de 0.95 sur les données brutes et 0.81 après transformation logarithmique en base 10 des données. Cette forte corrélation nous a donc incité à n'explorer que les effets communs à  $x$  et  $y$ . De plus, une tentative d'interprétation sur une interaction à quatre facteurs serait beaucoup trop discutable.

Enfin l'analyse de la composante  $z$  fait apparaître les mêmes effets significatifs que l'analyse selon la composante  $x$  du signal.

Le casque est toujours plus favorable ( $p = 0.0074$ ,  $F = 11.15$ ), la représentation basée sur les croix également mais dans une moindre mesure ( $p = 0.0113$ ,  $F = 9.56$ ).

En ce qui concerne l'interaction entre la représentation et le référentiel de présentation, les différences sont significatives. Selon l'axe  $z$ , il apparaît que pour la représentation basée sur les croix, le référentiel de présentation lié à l'aiguille est mieux adapté ( $p = 0.0556$ ,  $F = 4.68$ ). Ceci renforce également la conclusion établie lors de l'étude de l'interaction à tendance significative entre la représentation et le référentiel de présentation, dans le signal global. Pour la représentation basée sur un cône, le référentiel de présentation lié à la trajectoire est le plus adapté ( $p = 0.0042$ ,  $F = 13.55$ ).

## 4.5 Bilan de l'évaluation expérimentale

Avant de conclure l'évaluation expérimentale, il est important de noter que les effets observés lors de l'analyse d'une composante du signal ne sont pas en accord avec ceux observés lors de l'analyse du signal global. En effet les résultats de l'analyse statistique du signal selon  $z$ , en accord avec l'analyse du signal selon  $x$ , vont à l'encontre de ceux obtenus lors de l'analyse du signal global, en particulier pour l'interaction entre représentation, référentiel de présentation et axe du signal global (paragraphe 4.3.4.2).

Cependant lors de l'analyse d'une composante du signal global, les effets observés peuvent légitimement ne pas apparaître de la même façon que lors d'une analyse globale du signal. En effet, le signal global est composé d'un ensemble d'effets qui peuvent se compenser mutuellement ou s'annuler, pour finalement rendre des effets plus modérés ou différents de ceux observés sur chaque composante du signal. Pour notre étude, il est primordial de considérer les effets observés lors de l'analyse du signal global, puisque c'est l'intégralité du signal qui est produit par l'utilisateur lors de la réalisation de sa tâche. Or la qualité de réalisation de la tâche dans des configurations données constitue l'un des objectifs motivant l'expérimentation.

Cette expérimentation établit fermement l'adaptation du casque à la réalisation de cette tâche de reproduction d'une trajectoire. En effet l'analyse du signal global ainsi que celle de deux de ses trois composantes établissent l'apport du casque dans la réalisation de la tâche. L'analyse de la troisième composante, l'axe  $y$ , ne permet pas de déterminer un effet ou une tendance ( $p = 0.163$ ).

Comme pour le casque, la représentation basée sur un cône est largement mise en avant par l'analyse du signal global, et des composantes  $x$  et  $z$ . Là encore, la composante  $y$  du signal ne permet pas d'établir de différence significative, la valeur obtenue pour  $p$  étant 0.208.

Ainsi cette expérimentation permet de conclure que :

**L'utilisation d'un casque semi-transparent et  
l'utilisation d'une représentation de la trajectoire basée sur un cône,  
considérés indépendamment l'un de l'autre, indépendamment  
du référentiel de présentation et indépendamment des axes,  
concourent nettement à l'amélioration de la qualité de réalisation de la tâche.**

Par ailleurs, en croisant les résultats obtenus lors de l'analyse des interactions entre facteurs, nous pouvons établir que pour la tâche de chirurgie réelle considérée :

**La combinaison du référentiel de présentation lié à l'aiguille,  
avec la représentation des données basée sur un cône  
semble favoriser la qualité de réalisation de la tâche.**

Les deux conclusions ci-dessus prouvent expérimentalement nos résultats conceptuels : la continuité de l'interaction (perceptuelle et cognitive) est un paramètre important pour réduire les efforts de l'utilisateur lors de la réalisation de la tâche.

De plus, le couplage du référentiel de présentation lié à la trajectoire avec la représentation arbitraire basée sur des croix semble constituer une combinaison qui favorise également la réalisation de la tâche. Néanmoins, l'analyse des composantes  $x$  et  $y$  du signal tend à donner une conclusion contraire. Aussi, malgré notre remarque sur l'importance de l'analyse du signal global par rapport à celle des composantes du signal, nous pensons que ce résultat reste à approfondir.



Enfin, pour terminer cette expérimentation, il est important de noter qu'avec un nombre plus grand de sujets, des effets complémentaires ou accentués, en relation avec la typologie des sujets, auraient pu être dégagés. L'influence, sur le comportement des utilisateurs, de dispositifs de visualisation de réalité augmentée, tel que le casque semi-transparent, reste donc à approfondir. Nous envisageons d'analyser en détail cet aspect dans une prochaine étude, qui intégrerait également le troisième référentiel de présentation, référentiel de présentation utilisateur.

## 5 Conclusion

Ce chapitre consacré à *CASPER-V2* couvre toutes les étapes du cycle de vie logiciel, de la conception à l'évaluation expérimentale.

Nous avons tout d'abord présenté les alternatives à la conception de la première version de *CASPER*, en vue de pallier aux problèmes ergonomiques de l'interaction identifiés au chapitre précédent. Parmi ces solutions, nous avons retenu celles représentant les données de guidage au moyen d'un cône 3D affiché dans un casque semi-transparent, au travers duquel le clinicien perçoit directement le champ opératoire. Plusieurs référentiels de présentation pour les données de guidage complètent cette solution de conception de *CASPER-V2*.

En appliquant notre démarche d'analyse ergonomique prédictive à *CASPER-V2*, nous avons montré que la nouvelle conception permet de résoudre les principaux points négatifs et en particulier la discontinuité perceptuelle et cognitive.

La conception menée à terme et validée de façon prédictive, nous avons expliqué la réalisation logicielle de *CASPER-V2*. Nous avons souligné les apports du modèle d'architecture PAC-Amodeus adapté à la réalisation des systèmes mixtes. Plusieurs aspects logiciels ont ensuite été détaillés, en particulier la superposition des profondeurs des champs réel et virtuel et la localisation des yeux du chirurgien, c'est-à-dire le calibrage du casque semi-transparent.

*CASPER-V2* développé, les premiers tests ont été conduits avec le chirurgien. Nous avons ensuite effectué une évaluation expérimentale, visant à établir l'impact de l'utilisation d'un casque semi-transparent sur le comportement de l'utilisateur. Cette expérimentation nous a permis de valider l'utilisabilité de *CASPER-V2*, et donc nos résultats conceptuels liés à nos propriétés ergonomiques : la compatibilité et la continuité perceptuelles et cognitives. En effet, les résultats de l'expérimentation soulignent un effet nettement positif lié à l'utilisation du casque semi-transparent, c'est-à-dire lié au respect de notre propriété de continuité perceptuelle. Nous avons également établi l'existence d'un effet significatif favorisant l'interaction, lié à la mise en œuvre d'une représentation 3D de la trajectoire basée sur un cône. Cette représentation 3D, comme nous l'avons expliqué, favorise la propriété de continuité cognitive.

L'expérimentation conduite est pionnière par le fait qu'elle visait à établir l'utilisabilité d'un système mixte, combinant les mondes virtuel et réel. Néanmoins, la prise en compte de paramètres psychologiques propres à chaque utilisateur et caractérisant les principes mis en œuvre par l'utilisateur pour se repérer dans l'espace, n'a pas permis d'établir des conclusions définitives. Ce manque de résultat est en partie lié au nombre de participants (12) jugé trop faible par nos collègues du Laboratoire de Psychologie Expérimentale de Grenoble pour permettre d'établir des conclusions. Nous envisageons donc de conduire

une nouvelle expérimentation. L'objectif sera aussi de comparer la qualité de réalisation de la tâche en présentant les données de guidage :

- selon le référentiel de présentation utilisateur,
- selon les points de vue testés dans la première expérimentation.

Outre une deuxième expérimentation, de nombreuses améliorations et extensions logicielles à *CASPER-V2* sont en cours.

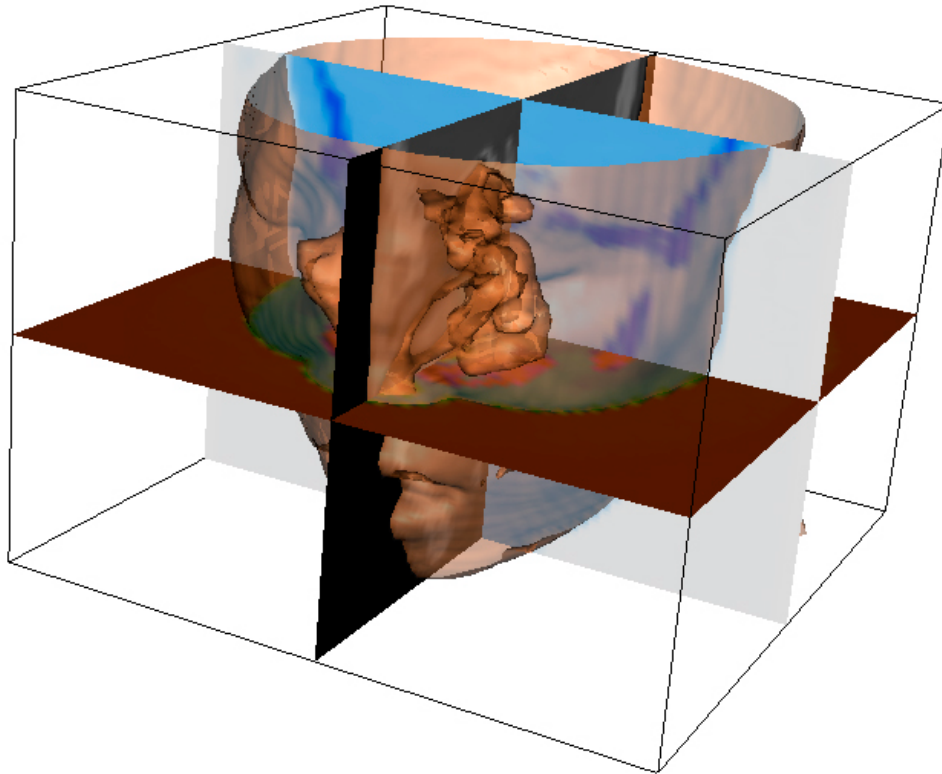
Nous travaillons actuellement à accroître la vitesse de calcul et d'affichage des données de guidage 3D dans le casque. Pour une vision stéréoscopique, la vitesse de rafraîchissement moyenne est légèrement inférieure à 20 Hz. Il conviendrait de parvenir à 25-30 Hz pour un confort accru de visualisation. L'optimisation en cours concerne essentiellement les phases de calcul des coordonnées des éléments constituant la scène 3D.

Parallèlement, nous affinons la précision du recalage des données virtuelles sur le patient réel. La précision de *CASPER-V2* est actuellement de l'ordre du centimètre. Ceci est insuffisant pour une intervention chirurgicale réelle pour laquelle une précision sub-millimétrique est requise. La construction des scènes 3D affichées se basant sur les données issues du calibrage du casque, l'alignement des données virtuelles sur le monde réel sera grandement amélioré par une précision accrue du calibrage du casque. Augmenter cette précision est tout à fait possible, comme l'attestent les travaux sur le calibrage de casques semi-transparents présentés dans [Azuma 95] et [Salb 01]. Nous envisageons d'appliquer la technique utilisée à l'Université de Karlsruhe [Salb 01] : en effet la précision de localisation du casque obtenue est inférieure à 0.3 millimètres. Cette technique repose sur la conception et l'utilisation d'un corps-rigide dédié au casque semi-transparent et fixé sur la partie rigide du casque contenant les "écrans" semi-transparents. Or, dans *CASPER-V2*, nous avons utilisé un corps-rigide classique, fixé sur l'armature du casque. Le corps-rigide n'est donc pas spécialement adapté au casque. De plus l'armature, sur laquelle il est fixé, est légèrement déformable par rapport à la partie rigide du casque à laquelle sont fixés les écrans. Les contacts déjà établis avec l'Université de Karlsruhe, dont les travaux sont menés dans les mêmes conditions matérielles que *CASPER-V2*, nous permettent d'envisager à court terme un meilleur calibrage du casque et donc un alignement amélioré des données virtuelles sur le monde réel.

A plus long terme, il conviendra d'améliorer la qualité de la scène calculée dans le cas d'une vision stéréoscopique. Dans la version actuelle de *CASPER-V2*, nous avons mis en œuvre un modèle relativement simpliste de la stéréoscopie. Nous envisageons donc de l'enrichir afin de mieux rendre perceptible les profondeurs dans le champ virtuel, en particulier en essayant d'introduire des effets de flou proposés par la librairie OpenGL. Cette amélioration devrait également permettre de mieux superposer les profondeurs des champs réel et virtuel.

La combinaison d'un meilleur calibrage avec une vision stéréoscopique de qualité nous permettra d'afficher des données de guidage plus complexes. En particulier, nous souhaitons modifier le principe de visualisation utilisé lorsque trois plans de coupe dans une structure anatomique sont nécessaires pour guider efficacement le clinicien. A chacun de ces plans correspond une image médicale frontale, distale ou sagittale de la structure anatomique traitée, sur laquelle sont ajoutées des données de guidage, comme la position

de l'aiguille réelle par rapport à la trajectoire planifiée. Classiquement, ces plans sont affichés côte à côte sur un écran. En utilisant le casque semi-transparent, nous envisageons une solution tridimensionnelle basée sur un modèle 3D de la structure anatomique à traiter. Aligné avec la structure anatomique réelle, ce modèle sera enrichi par l'ajout des trois plans de coupe et de leurs données de guidage. La Figure 89 illustre ce principe pour le modèle 3D d'un visage, sur lequel sont ajoutés trois plans de coupe, porteurs d'images scanner du crâne du patient. Pour cela, l'utilisation de la bibliothèque de visualisation graphique 3D Visualisation ToolKit (VTK) est pressentie.



**Figure 89** *Superposition de 3 plans de coupe porteurs d'images scanner et d'un modèle 3D du visage d'un patient.  
Photographie issue du manuel de programmation de Visual ToolKit (VTK).*

Il conviendra aussi d'étudier en détail les moyens d'interaction avec ces données additionnelles : par exemple le clinicien doit pouvoir spécifier les plans à afficher, les données à afficher sur chacun des plans, ou encore la transparence à appliquer entre les plans. Le développement complet de cette solution constituera alors une nouvelle modalité de guidage pour les gestes médico-chirurgicaux assistés par ordinateur.

De nombreuses autres possibilités qui ne reposent pas uniquement sur des données de guidage visuelles sont envisageables. D'ailleurs, comme nous l'avons annoncé, une version de *CASPER* basée sur un retour d'effort est en cours de réalisation [Schneider 99].

Dans une perspective encore plus vaste, la combinaison d'un casque avec un dispositif à retour d'effort comme le *Phantom* de SensAble Technologies [Phantom] devrait permettre de réaliser des situations de simulation et d'enseignement. Nous développons ces perspectives à nos travaux dans le chapitre suivant, la conclusion de ce mémoire.

## 6 Références

[Abdi 87]

Abdy, H., *"Introduction au Traitement Statistique des Données Expérimentales"*, Presses Universitaires de Grenoble, (1987), 420 pages.

[Azuma 95]

Azuma, R., T., *"Predictive Tracking for Augmented Reality"*, Thèse de l'Université de Caroline du Nord, Chapel Hill - USA, (1995), 242 pages.

[Droulez 89]

Droulez, J., Darlot, C., *"The geometric and dynamic implications of the coherence constraints in three-dimentionnal sensorimotor interactions"*, dans M. Jeannerod (Ed.), *Attention and Performance XIII*, (1989), p. 495-526.

[FOB]

Documentations techniques et Téléchargement de drivers, URL : <http://www.ascension-tech.com/products/flockofbirds/>

[Grimson 96]

Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells, W.M., Kikinis R., *"An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization"*, IEEE Transactions on Medical Imaging, 15(2), (1996), p. 129-140.

[Ohlmann 90]

Ohlmann, T., *"Les systèmes Perceptifs Spatiaux Vicariants"*, dans Cognition : l'Individuel et l'Universel, Psychologie d'Aujourd'hui, PUF, (1990), p. 21-58.

[Oltman 68]

Oltman, T., *"A Portable Od and Frame Apparatus, Perceptual and Motor Skills"*, 26, (1968), p. 503-506.

[Peuchot 95]

Peuchot, B., Tanguy, A., Eude, M., *"Virtual Reality as an Operative Tool During Scoliosis Surgery"*, Acte de la conférence CVRMed'95, Nice, France, (1995), p. 549-554.

[Phantom]

Phantom de Sensable Technologies, URL : [www.sensable.com/haptics/products/phantom.html](http://www.sensable.com/haptics/products/phantom.html)

[Radermacher 97]

Radermacher, K., Portheine, F., Zimolong, A., Eichhorn, C., Staudte, H., W., Rau, G., *"Image Guided Orthopedic Surgery Using Individual Template"*, Acte de la conférence CVRMed-MRCAS'97, Grenoble, France, (1997), p. 606-615.

[Salb 01]

Salb, T., Brief, J., Burgert, O., Gockel, T., Giessler, B., Dillmann, R., *"Comparison of Tracking Presentation of Medical Data Using a See-Through Head-Mounted Display"*, Acte de la Conférence MMVR'2001, Newport Beach - CA, USA, (2001).

[Schneider 99]

Schneider, O., Troccaz, J., Chavanon, O., Blin, D., *"Synergistic robotic assistance in cardiac procedure"*, Acte de la conférence CAR'99, (1999), p. 803-807.

[Winer 62]

Winer, B., J., *"Statistical Principles in Experimental Design"*, Mc Graw-Hill Book Company. Perdue University, (1962).



# **CHAPITRE VIII**

## **CONCLUSION**

# Sommaire

<b>Chapitre VIII : Conclusion</b>	<b>235</b>
<b>1 Résumé de notre contribution</b>	<b>235</b>
<b>2 Perspectives</b>	<b>238</b>
2.1 Extensions à court terme	238
2.1.1 Réalisation logicielle	238
2.1.2 Evaluation	239
2.2 Prolongements	239
2.2.1 Identification du niveau d'abstraction des composants	240
2.2.2 Approche logicielle et composants réutilisables	242
2.2.3 Ouverture à d'autres domaines	242

## **Chapitre VIII : Conclusion**

Ces travaux sont résolument pluridisciplinaires et contribuent aux domaines de recherche que sont les Gestes Médico Chirurgicaux Assistés par Ordinateur (GMCAO) et l'Interaction Homme-Machine (IHM). Notre étude jette un regard nouveau sur la conception et la réalisation des systèmes de GMCAO, en adoptant une approche centrée sur l'interaction homme-machine. Nos résultats de conception et de réalisation contribuent plus généralement à l'étude de systèmes mixtes, terme que nous avons adopté pour dénoter tout système qui allie les mondes réel et virtuel. Dans cet axe de recherche de l'IHM, les avancées sont brutales et se manifestent par de nombreux systèmes présentés isolément, ainsi que par une prolifération de termes employés le plus souvent sans rigueur. L'établissement d'un cadre de classification et d'éléments de conception ergonomique et de réalisation logicielle, constitue une avancée nécessaire et fédératrice pour appréhender de façon systématique la fusion des deux mondes, réel et virtuel.

Nous proposons ici, en conclusion, un résumé de nos travaux organisé selon les étapes du cycle de vie du logiciel. Puis dans le paragraphe 2, nous exposons les extensions à court terme et les perspectives à plus long terme devant faire suite à ces travaux.

### **1 Résumé de notre contribution**

Un bilan analytique des approches de caractérisation des systèmes de GMCAO (Chapitre I) et plus généralement des systèmes mixtes (Chapitre II) nous conduit à présenter notre cadre taxinomique : celui-ci est centré sur l'interaction de l'utilisateur avec le système (Chapitre III). Cette taxonomie distingue deux classes de systèmes :

- la réalité augmentée, où l'objet de la tâche est une entité de l'environnement physique de l'utilisateur, c'est-à-dire du monde réel,
- la virtualité augmentée où l'objet de la tâche est une entité virtuelle dont la manipulation ou la perception s'appuient sur des outils du monde réel.

Cette distinction sur l'objet de la tâche établit explicitement le monde modifié suite à l'accomplissement de la tâche : le monde réel pour la réalité augmentée et le monde virtuel pour la virtualité augmentée. Par exemple, un système de GMCAO est un système de Réalité Augmentée, puisque l'intervention chirurgicale constitue la tâche à réaliser. De plus, notre taxonomie s'appuie sur le type d'augmentation, en se référant aux deux phases de réalisation de la tâche, l'évaluation ou l'exécution : quelle phase de la réalisation de la tâche (évaluation ou exécution) est enrichie par le système mixte.

Outre la comparaison de systèmes existants, cette taxonomie est aussi un outil pour le concepteur, car elle l'aide à cerner ses besoins, c'est-à-dire agir sur le monde réel ou le monde virtuel, ainsi que les possibilités favorisant une meilleure intégration des mondes réel et virtuel, c'est-à-dire l'amélioration de l'évaluation et/ou de l'exécution lors de la réalisation de la tâche.

L'identification des composants et relations ASUR ainsi que leurs caractérisations permettent de décrire les différentes facettes de l'interaction offertes par un système mixte. Notre seconde contribution, la notation ASUR (chapitre IV), permet de décrire et



de caractériser les deux types d'interaction coexistant dans un système mixte : l'interaction avec l'environnement réel et l'interaction avec le monde informatique, virtuel. Or c'est au sein de cette dualité de l'interaction que réside la spécificité des systèmes mixtes. Aussi ASUR est un outil de conception qui aide le concepteur à se poser les bonnes questions sur la dualité de l'interaction à mettre en œuvre, un aspect non traité par les méthodes de conception existantes. De plus les caractéristiques ASUR des composants et relations contiennent et étendent celles des approches existantes. En cela ASUR enrichit et fédère les approches existantes de caractérisation des systèmes mixtes.

Pour compléter la notation de description ASUR, support à la conception, nous dépeignons l'utilisabilité en termes de quatre propriétés ergonomiques : la continuité et la compatibilité, au niveau perceptuel et cognitif (Chapitre V). Ces propriétés prennent particulièrement en compte la dualité de l'interaction Utilisateur-Système Mixte. Cette dualité n'étant pas présente dans les systèmes interactifs classiques, nos quatre propriétés complètent donc les propriétés ergonomiques existantes.

Leur traduction en termes de caractéristiques de composants et relations ASUR sert de fondement à notre démarche d'analyse ergonomique pour les systèmes mixtes. Cette démarche d'analyse constitue un outil d'aide à l'évaluation prédictive de nos propriétés ergonomiques, lors de la conception.

Nous établissons des liens entre nos éléments de conception de l'interaction et la réalisation logicielle de systèmes mixtes (Chapitre V). Pour ce faire, nous définissons une solution générique de réalisation logicielle selon le modèle d'architecture logicielle PAC-Amodeus adapté aux systèmes mixtes. La décomposition en composants logiciels véhiculée par le modèle adapté souligne là encore la dualité de l'interaction, avec des branches de composants dédiées à l'interaction avec le monde réel et d'autres à l'interaction avec le monde informatique, virtuel. Le modèle adapté préconise aussi que le Contrôleur de Dialogue, composant principal du modèle, soit indépendant des adaptateurs mis en évidence dans la description ASUR. Cet indépendance permet une forte modifiabilité du code vis-à-vis des adaptateurs. Ces derniers, établissant la passerelle entre les deux mondes, réel et virtuel, sont centraux dans la conception d'un système mixte et une forte modifiabilité du code vis-à-vis de ces adaptateurs est requis dans le cadre d'une conception itérative.

Nous examinons la validité des concepts développés en les mettant en œuvre dans des réalisations logicielles. Le système conçu et développé vise à assister par ordinateur la réalisation d'une ponction péricardique. Ce système se nomme CASPER (Chapitre VI).

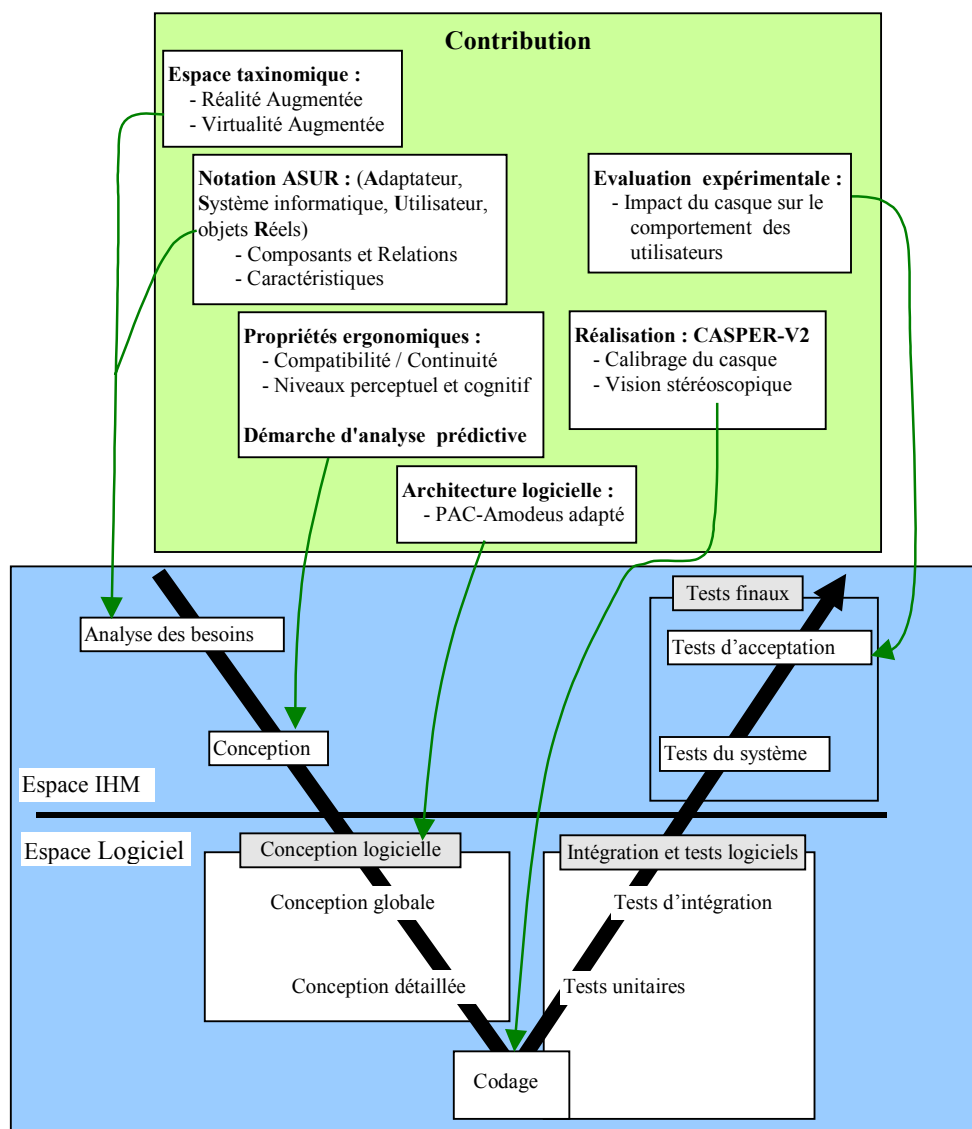
Nous soulignons l'apport de notre démarche de conception et d'analyse prédictive ergonomique à la re-conception du système (Chapitre VII). Par sa simplicité et l'ensemble des caractéristiques considérées, notre démarche d'analyse permet de répondre aux exigences du domaine tout en privilégiant le clinicien lors de la conception de l'interaction Chirurgien-Système de GMCAO.

La réalisation logicielle de CASPER-V2 (Chapitre VII) a permis de valider le modèle d'architecture PAC-Amodeus adapté. De plus un processus de calibrage du casque semi-transparent a été développé afin de permettre un alignement des données virtuelles

affichées dans le casque avec l'environnement réel de l'utilisateur. Enfin la mise en œuvre de la visualisation 3D stéréoscopique dans le casque a aussi été réalisée.

Enfin, nous validons une partie de nos résultats par une évaluation expérimentale de CASPER-V2, menée en collaboration avec des psychologues. Pionnière de par l'un de ses objectifs, évaluer l'impact de l'utilisation d'un dispositif d'affichage de systèmes mixtes sur le comportement de l'utilisateur, cette expérimentation valide l'importance de nos propriétés ergonomiques pour l'utilisabilité. Elle confirme également que les caractéristiques ASUR que nous considérons pour dépeindre ces propriétés ont un impact sur la qualité de réalisation de la tâche.

Nous concluons la synthèse de notre contribution, en présentant les résultats obtenus selon les étapes du cycle de vie logicielle en V à la Figure 90. Nous y retrouvons la dualité de nos travaux, conceptuels et techniques.



**Figure 90** Synthèse de nos résultats localisés dans un cycle de vie logiciel en V.

## 2 Perspectives

Les deux facettes de nos travaux, conception et réalisation, sont également apparentes dans les perspectives que nous donnons à ces travaux. Dans ce paragraphe nous présentons tout d'abord des perspectives à court terme avant de présenter nos perspectives à plus long terme.

### 2.1 Extensions à court terme

A court terme, les extensions concernent principalement la réalisation logicielle et les évaluations à mener sur *CASPER-V2*.

#### 2.1.1 Réalisation logicielle

Les extensions logicielles que nous envisageons d'apporter à *CASPER-V2*, ont été décrites dans le chapitre VII. Celles-ci consistent tout d'abord en l'amélioration de la vitesse de rafraîchissement des données de guidage dans le casque. Cette amélioration est notamment nécessaire lors de l'affichage simultané du cône 3D et de l'image échographique de contrôle. Cette première extension nécessite une optimisation du processus de calcul de la scène à afficher ainsi que du stockage de l'image échographique.

Nous avons également mentionné, au chapitre VII, une précision moyenne de l'alignement des données virtuelles affichées dans le casque par rapport aux données réelles correspondantes. Cette optimisation nécessite inévitablement une amélioration du processus de calibrage du casque. Depuis le moment où nous avons conçu cette méthode de calibrage, plusieurs équipes de recherche ont été confrontées au même problème et une solution permettant un calibrage précis du casque sera développée prochainement en collaboration avec l'Université de Karlsruhe.

Outre ces premières extensions visant à améliorer la qualité et la vitesse du rendu graphique dans le casque, nous envisageons également de concevoir une scène 3D pour les données de guidage correspondant mieux aux deux phases de réalisation de la tâche. Cette représentation comporte deux parties : l'une est composée d'un cône tronqué plus évasé que dans la version actuelle et correspond aux données de guidage nécessaires à la phase d'approche, c'est-à-dire tant que l'aiguille est totalement hors du corps du patient. La seconde partie de la scène se compose d'un cône dont le diamètre est beaucoup plus étroit que celui actuellement utilisé et qui représente effectivement la trajectoire que l'aiguille doit suivre à l'intérieur du patient.

Nous envisageons enfin de réaliser une "vue rétroviseur" dans le casque. Dans cette vue, les données sont de nature différente de celles affichées dans le casque. Il s'agit par exemple d'une image échographique fixe, ou de données de guidage présentées sous une forme analogue à celle mise en œuvre dans la première version de *CASPER* (les croix et la jauge). Pour percevoir cette "vue rétroviseur", le clinicien doit tourner la tête au-delà d'un seuil restant à définir. Ainsi, lorsque le clinicien regarde le champ opératoire, il perçoit automatiquement les données l'aidant à réaliser sa tâche clinique ; par contre en détournant le regard du champ opératoire, ou plus généralement de l'objet de sa tâche, le clinicien perçoit d'autres données. L'intérêt de la "vue rétroviseur" est de fournir au clinicien, dans un environnement contraint, les fonctionnalités d'un écran multiple. Ainsi plusieurs données pertinentes pour la réalisation de la tâche, mais exprimées par une

relation ASUR de poids faible ou périphérique, pourront être perçues par le clinicien au moyen d'une tâche articulaire la moins contraignante possible.

### 2.1.2 Evaluation

Pour valider expérimentalement l'intégralité de nos résultats de conception, nous envisageons de conduire de nouvelles expérimentations.

D'autres situations d'interaction sont au centre de ces futurs tests utilisateurs. Il s'agit de présenter différents types de données, tels que des données anatomiques seules, ou des données de stratégie seules, ou encore des deux combinées. Pour chacune de ces situations, l'impact des trois référentiels de présentation identifiés, le référentiel lié à l'outil chirurgical, le référentiel lié à la trajectoire et le référentiel utilisateur, est à tester. Ce dernier référentiel de présentation n'avait pas pu être testé lors de notre première expérimentation, du fait d'un manque évident de précision. Depuis, le référentiel utilisateur procure des données de guidage très satisfaisantes et devra faire l'objet d'une évaluation expérimentale au même titre que les deux autres référentiels de présentation.

Enfin, l'étude des liens entre d'une part, le support d'affichage et/ou la forme des données, et d'autre part, des caractéristiques cognitives propres aux utilisateurs, constitue un objectif pour nos expérimentations. Il s'agit d'identifier d'éventuelles implications sur le comportement de l'utilisateur lors de l'utilisation d'un système mixte. En collaboration avec le Laboratoire de Psychologie Expérimentale de Grenoble, l'étude porte sur la typologie des utilisateurs vis-à-vis de la prise de repère dans l'espace. Ces résultats n'ont pas pu être observés lors de notre première expérimentation, et nous estimons que le faible nombre de participants en est une raison.

Concernant l'application clinique *CASPER-V2*, une évaluation clinique portant sur l'utilisation du casque semi-transparent doit être menée afin de permettre son utilisation en bloc opératoire sur des patients réels. Un pré-requis est toutefois l'amélioration de la précision de l'alignement des données virtuelles affichées dans le casque avec le monde réel, comme mentionné au paragraphe précédent : la précision requise en chirurgie cardiaque est sub-millimétrique. L'écart entre les données réelles et virtuelles doit donc être inférieur au millimètre.

## 2.2 Prolongements

Dans les systèmes mixtes, et en particulier dans les systèmes de GMCAO, de nombreuses fonctionnalités sont communes, comme par exemple les procédures de recalage, c'est-à-dire d'alignement, des différentes données manipulées, les différentes formes d'affichage des données virtuelles, les processus de suivi des entités manipulées, etc. Malgré ces points de conception communs, très peu d'outils génériques sont développés. La réalisation d'un système mixte s'appuie donc presque systématiquement sur un développement complet du noyau fonctionnel et des interfaces homme-machine. Face à la variété et la multiplicité des formes d'interaction qui visent à l'intégration des mondes réel et virtuel, il convient d'aider le développeur en fournissant des composants logiciels réutilisables. Notre objectif est donc de définir une boîte à outils qui fournisse un ensemble de composants logiciels et matériels réutilisables. L'enjeu est ici d'identifier la granularité de ces composants puis de les mettre en œuvre.

Le développement de composants réutilisables nécessite en premier lieu de définir le niveau d'abstraction requis par ces composants. En effet, si les composants produits

forment un agrégat de fonctionnalités que le concepteur souhaite pouvoir utiliser indépendamment les unes des autres, cela signifie que la granularité des composants logiciels fournis par la boîte à outils n'est pas adaptée aux besoins. L'identification du niveau d'abstraction constitue donc un premier axe dans nos perspectives, que nous détaillons au paragraphe 2.2.1. Dans un second temps, le développement de la boîte à outils doit s'appuyer sur une architecture logicielle adéquate. Ainsi, une abstraction des composants produits nous amènera à produire des composants génériques permettant de concevoir et développer des systèmes mixtes dans des domaines applicatifs variés. Cette seconde étape concrétise nos résultats de conception et nous la détaillons dans les paragraphes 2.2.2 et 2.2.3.

### 2.2.1 Identification du niveau d'abstraction des composants

Le développement de formes d'interaction complémentaires est l'un des soucis les plus importants. Les nouvelles formes d'interaction proposées viseront à intégrer de manière encore plus transparente les actions et données fournies par le système à l'environnement réel de l'utilisateur. Nous identifions trois axes de recherche complémentaires pour l'identification de nouvelles formes d'interaction : celles basées sur de nouvelles formes de données, celles exploitant de nouveaux types de dispositifs et celles exploitant le contexte physique de l'utilisateur. Nous revenons sur ces trois axes dans les paragraphes suivants.

Dans *CASPER-V2*, une donnée relativement simple est présentée : une zone de sécurité autour d'une trajectoire à suivre. Néanmoins, dans la majorité des applications de navigation, le praticien utilise simultanément trois coupes - frontale, latérale et sagittale - de la zone anatomique traitée, pour visualiser par exemple l'avance d'un outil par rapport à la position des structures anatomiques à risque qu'il ne peut percevoir. Une forme de visualisation adaptée à la présentation de ces trois plans est donc requise. C'est par exemple le cas d'un système de GMCAO très similaire à *CASPER*, avec lequel une vis doit être positionnée dans l'os d'un bassin fracturé. Avec la nouvelle forme de présentation, les données de guidage de la vis sont entièrement basées sur des données 3D représentant d'une part un modèle du bassin et de la position idéale de la vis et, d'autre part des données de guidage selon trois plans liés à la position du foret. Sur l'intersection de ces trois plans avec le modèle du bassin, la projection de la trajectoire idéale est rendue et une coupe scanner est superposée à ces données de stratégie. Les données de guidage et anatomiques sont alors alignées avec le bassin du patient réel, permettant ainsi au clinicien de comparer au sein d'une seule représentation graphique 3D la position de son aiguille réelle en regard de la trajectoire idéale exprimée.

Au niveau des dispositifs, nous avons privilégié l'utilisation d'un casque semi-transparent pour pallier aux problèmes ergonomiques identifiés dans l'interaction Utilisateur-*CASPER*. Une autre solution technologique envisagée est l'utilisation de mini-écrans positionnés directement dans l'environnement perceptif dominant de la tâche de l'utilisateur. De tels écrans pourront être dans un premier temps des écrans de *Personal Digital Assistant* (PDA), qui offrent désormais des très bonnes capacités de calcul et d'affichage. De plus, une communication entre le calculateur et le PDA à base d'infrarouges laisse aussi présager la possibilité de communication sans fil, souhaitable

étant données les conditions d'asepsie imposées dans un bloc opératoire, ainsi que pour limiter l'encombrement du bloc opératoire par la présence de fils. Un système exploitant la communication sans fil est en cours d'étude dans le cadre du projet MAGIC (France Telecom R&D) menée par l'équipe IIHM. Cette application, développée sur une tablette tactile dotée de capacités de calcul élevées, constituerait donc un point de départ.

Par ailleurs, la chirurgie étant une activité manuelle, l'utilisation d'informations tactiles semble tout à fait adaptée. Ainsi, l'utilisation de robots à retour d'effort est un second axe de recherche. Toutefois, cela impose de nouveaux pré-requis qu'il convient d'étudier et d'intégrer dans nos éléments de conception ergonomique. Des travaux sur le *Phantom* ont débuté et les travaux d'Olivier Schneider sur un prototype de robot passif à retour d'effort constituent une base solide pour le développement de nouvelles formes d'interaction pour les systèmes mixtes.

D'autre part, nous notons, dans l'évolution actuelle des systèmes interactifs, l'importance de la prise en compte des données fournies par ou captées dans le contexte environnant l'utilisateur. Le développement de systèmes mixtes exploitant cette dimension est un axe prometteur car il permet d'intégrer plus complètement l'environnement réel de l'utilisateur dans le système mixte. Nous les dénotons systèmes contextuels mixtes. Cependant, il n'existe aujourd'hui que quelques systèmes exploitant le contexte et le concept de contexte n'est lui-même pas clairement établi. Il n'existe d'ailleurs pas de cadre de conception permettant de structurer les axes de recherche concernant la conception des systèmes contextuels mixtes : il est donc actuellement très délicat de comparer des systèmes contextuels mixtes existants ou d'en envisager de nouveaux. Une situation similaire est à l'origine de nos travaux concernant les systèmes mixtes. Aussi il convient de reprendre les étapes de travail suivies lors de nos travaux concernant les systèmes mixtes afin d'enrichir nos résultats de conception avec des éléments intégrant le concept de contexte.

Concrètement, il s'agit tout d'abord d'identifier les caractéristiques et les propriétés du contexte environnant l'utilisateur pour établir des classes de dispositifs et de capteurs permettant de combler le gouffre entre l'environnement physique et le monde virtuel. Il convient également d'étudier les moyens permettant de rendre compte de ces modifications sans pour autant altérer l'utilisabilité du système. A ces fins, nous nous appuyons sur les résultats présentés dans cette thèse. En particulier, il convient d'étendre notre notation ASUR pour ajouter un composant de type "Contexte" à la liste des composants identifiés. De la même façon que pour les systèmes mixtes étudiés dans ces présents travaux, c'est dans l'analyse des relations entre les composants ASUR "Contexte" et les autres composants ASUR que réside l'identification d'éléments de conception pour les systèmes contextuels mixtes.

Enfin, les évolutions d'ASUR doivent permettre de prendre en considération les aspects collaboratifs de la réalisation d'une tâche à l'aide d'un système mixte. En effet, en milieu chirurgical par exemple, le chirurgien réalise une intervention chirurgicale au sein d'une équipe chirurgicale. La communication entre ces personnes, mais surtout la capacité de chacun d'elle d'apprécier l'état d'avancement de l'intervention est un élément très important pour garantir une bonne coordination entre les différents membres de l'équipe chirurgicale. Pour cela, les travaux menés autour du concept de "*grounding*" constituent

un point de départ intéressant. Le "*grounding*" représente les ressources communes de base dont disposent deux interlocuteurs, leur permettant de se comprendre lors d'une conversation orale [Clark 89]. Le "*grounding*" constitue d'ailleurs l'un des concepts sur lequel s'appuie la conception d'applications pour la formation à distance et la communication médiatisée [Traum 96]. Une approche complémentaire est celle adoptée dans le cadre des environnements collaboratifs virtuels. Ainsi, le "*Spatial Model*" identifie des caractéristiques ("*medium*", "*aura*", "*awareness*", "*focus*", "*nimbus*" et "*adapter*") pour les objets impliqués dans l'interaction au sein d'un environnement virtuel [Benford 93]. Sur cette base, des mécanismes de gestion de l'interaction entre utilisateurs ou avec d'autres objets sont décrits et constituent un aspect complémentaire à notre approche. En effet, ces travaux concernent des objets virtuels, alors que nous sommes confrontés à des objets réels, mais ils abordent le problème du partage des objets dans des situations collaboratives, problème qui existe nécessairement avec des systèmes mixtes.

En conclusion, outre l'identification des éléments de base de notre boîte à outils, nous prévoyons que la conception de ces nouvelles formes d'interaction, basées sur de nouveaux langages, de nouveaux dispositifs ou l'intégration du contexte, impliquera d'étendre nos résultats de conception ergonomique. De plus *CASPER* constitue une plateforme de test intéressante. Plus généralement, la phase de guidage des systèmes de GMCAO offre un environnement pour développer ces nouvelles formes d'interaction, car il est concret et des besoins précis, en terme d'interaction, de coopération et de partage d'entités réelles sont identifiés.

### **2.2.2 Approche logicielle et composants réutilisables**

Une fois les éléments identifiés, nous visons à la concrétisation de nos résultats de conception par un ensemble de composants logiciels et matériels réutilisables et composables : une boîte à outils pour la réalisation logicielle de systèmes contextuels mixtes, c'est-à-dire des systèmes mixtes prenant en compte le contexte. La composition de ces éléments doit reposer sur une plateforme logicielle à l'architecture adaptée. Nous envisageons ici l'extension de nos résultats de conception logicielle et en particulier l'extension de notre modèle d'architecture logicielle.

Dans une première étape, la boîte à outils sera dédiée au domaine applicatif que constituent les gestes médico-chirurgicaux assistés par ordinateur. Notre objectif est ici de développer des composants logiciels et matériels génériques répondant aux besoins et aux caractéristiques ergonomiques identifiés dans notre démarche de conception ergonomique. Ainsi, la nécessité de réponse propre à un besoin technologique ponctuel sera remplacée par une réutilisation intensive d'éléments conçus et validés spécifiquement pour le domaine.

### **2.2.3 Ouverture à d'autres domaines**

L'étape suivante est naturellement la généralisation de la boîte à outils à d'autres domaines d'application.

Parce que les besoins sont différents d'un domaine d'application à l'autre, la généralisation de notre boîte à outils nécessite de caractériser de façon générique les besoins pour chaque classe de systèmes contextuels mixtes identifiée dans notre taxonomie. Nous envisageons une approche ascendante, en commençant par l'étude

d'autres domaines avant de viser à cette généralité. En premier lieu, nous nous intéresserons aux domaines de la simulation et de l'éducation dans le milieu médical. Des simulateurs, permettant l'apprentissage par la répétition d'un geste, ont déjà été développés dans ce domaine. Cependant, coupler ces simulateurs avec des formes d'interaction permettant de visualiser des structures anatomiques proches ajouterait une dimension pédagogique non négligeable au simple simulateur haptique : présentation des structures anatomiques avoisinantes, mise en évidence des structures à risque, etc.

Cette étape constitue ensuite un tremplin pour l'adaptation et le développement de ces mêmes techniques dans d'autres domaines tels que les jeux éducatifs et les Technologies de l'Information et de la Communication dans l'Enseignement (TICE), qui me tiennent particulièrement à cœur. En effet, la préoccupation actuelle des TICE est plus de cerner les besoins, les moyens à mettre en œuvre, les apports. Mais rapidement, les TICE donneront lieu au développement de nombreuses applications concrètes, nécessitant des formes d'interaction ayant un réel apport en regard des techniques pédagogiques utilisées actuellement.





# ANNEXES

---



# **ANNEXE A**

## **SPECIFICATIONS TECHNIQUES DU CASQUE SONY LDI-D100/BE**



# Annexe A : Specifications Technique du Casque SONY LDI-D100/BE

## SPECIFICATIONS

	LDI-50B/BE	LDI-D50B/BE	LDI-100B/BE	LDI-D100B/BE
LCD	0.7", 180K dot	0.7", 180K dot	0.7", 1.55M dot	0.7", 1.55M dot
Dot Number/Panel (H/V)	800(H) x 225(V) dot	800(H) x 225(V) dot	832(H) x 3 x 624(V) dot	832(H) x 3 x 624(V) dot
Virtual Image Size	52" screen at 2m	52" screen at 2m	30" screen at 1.2m	30" screen at 1.2m
Viewing Angle (H)	30 °	30 °	28 °	28 °
See-through Function	Mechanical shutter sys.	Mechanical shutter sys.	LCD shutter sys.	LCD shutter sys.
Control Functions	Sound volume	Sound volume	Sound volume	Sound volume
	Mega bass	Mega bass	Mega bass	Mega bass
	Surround	Surround	—	—
	AVLS (*1)	AVLS (*1)	AVLS (*1)	AVLS (*1)
	Brightness	Brightness	Brightness	Brightness
	Hue (NTSC)	Hue (NTSC)	Hue (NTSC (Video) mode only)	Hue (NTSC (Video) mode only)
	Color	Color	Color (Video mode)	Color (Video mode)
	—	—	Contrast (Video mode)	Contrast (Video mode)
	—	—	Tracking (PC mode)	Tracking (PC mode)
	—	—	H Center (PC mode)	H Center (PC mode)
	—	—	V Center (PC mode)	V Center (PC mode)
	—	—	Sync (PC mode)	Sync (PC mode)
	—	—	Power save	Power save
Input Video Signal System	NTSC/PAL	NTSC/PAL	NTSC/PAL	NTSC/PAL
Input Video Signal	Composite	Composite	Composite, Y/C	Y/C
Input PC Signal	No	No	Analog RGB	Analog RGB
Stereo 3D Signal Input	PC Mode	No	No	Yes
	Video Mode	No	Yes	Yes
Stereo 3D System	—	Field-sequential	—	Field-sequential
	FLD Pulse (*2)	—	No need to input	—
Connectors	Special mini jack (x1) for Audio/Video	Special mini jack (x1) for Audio/Video	Special mini jack (x1) for Audio/Video	Special mini jack (x1) for Audio/Video
	Special mini jack (x1) for Headphone	Special mini jack (x1) for Headphone	Stereo mini jack (x1) for PC Audio	Stereo mini jack (x1) for PC Audio
Power Requirement	DC 7.5V-9V (3.8W)	DC 7.5V-9V (3.8W)	DC7.5V-10V(10W approx. at PC mode)	DC7.5V-10V(10W approx. at PC mode)
	—	—	DC7.5V-10V(12W approx. at Video mode)	DC7.5V-10V(12W approx. at Video mode)
	—	—	—	—
Operating Temperature	5 to 35 °C	5 to 35 °C	5 to 35 °C	5 to 35 °C
Storage Temperature	-10 to 60 °C	-10 to 60 °C	-10 to 60 °C	-10 to 60 °C
Dimensions (W x H x L mm)	Display	approx. 165 x 56 x 110 mm	approx. 165 x 56 x 110 mm	approx. 149 x 48 x 86 mm
	Power Supply Box	approx. 53 x 30 x 149 mm	approx. 53 x 30 x 149 mm	approx. 54 x 47 x 160 mm
Weight (*3)	Display Unit	approx. 150g	approx. 150g	approx. 120g
	Power Supply Box	approx. 110g	approx. 110g	approx. 260g
Supplied Accessories	AC power adaptor (AC-PLM2) (x1)	AC power adaptor (AC-PLM2) (x1)	AC power adaptor (AC-PLM3) (x1)	AC power adaptor (AC-PLM3) (x1)
	Special mini plug-plug (3m) (x1)	Special mini plug-plug (3m) (x1)	Special mini plug-plug (3m) (x1)	Special mini plug-plug (3m) (x1)
	Side-Visor (x1 pair)	Side-Visor (x1 pair)	Side-Visor (x1 pair)	Side-Visor (x1 pair)
Supplied AC Power Adaptor	AC power code (x2)	AC power code (x2)	AC power code (x2)	AC power code (x2)
	AC-PLM2	AC-PLM2	AC-PLM3	AC-PLM3
Output voltage 9V, 1.3A (DC 7.5V - 9V)	Output voltage 9V, 1.3A (DC 7.5V - 9V)	Output voltage 8.4V (DC 7.5V - 10V)	Output voltage 8.4V (DC 7.5V - 10V)	

\*1 Auto Volume Limiter System - To keep the maximum volume down to protect your hearing. The volume does not become louder than the level which is set even if you try to adjust the VOL control.

\*2 FLD pulse must be input when Stereo 3D PC mode

\*3 Weight - not including cable

\*Caution - D series of LDI series reproduces 3D video signal by field-sequential mode

### Remarks

- CAUTION: This product should not be used by children age 15 or younger. See owners manual for additional safety/health considerations.
- Customization may be available depending upon quantity commitment.
- There are also Block Models available.
- VGA cable not supplied with LDI-100B, LDI-D100B.

### WARNING

Your failure to follow these operating instructions and the supplied important safeguards may result in eye fatigue, eye impairment, or other eye injury, property damage or death. This product should not be used by children age 15 or younger. The eyes of children are still developing and may be adversely affected from use of this product.

### ● Sony Electronics Inc. (USA) HQ

1 Sony Drive, Park Ridge, NJ 07656

(TEL:+1-800-886-7669)

### ● Sony of Canada Ltd. (CANADA)

115 Gordon Baker Rd, Toronto, Ontario M2H 3R6

http://www.sony.com/professional

### ● Sony Broadcast & Professional Europe HQ

15, rue Floreal 75831 Paris Cedex 17, France

(TEL:+33-1-40-87-35-11) (FAX:+33-1-40-87-35-17)

Germany

Hugo-Eckener-Str. 20, 50829 Köln

(TEL:+49-221-5966-322) (FAX:+49-221-5966-491)

France

15, rue Floreal 75831 Paris Cedex 17

(TEL:+33-1-49-45-41-62) (FAX:+33-1-47-31-13-57)

UK

The Heights, Brooklands, Weybridge, Surrey KT13 0XW

(TEL:+44-690-331122) (FAX:+44-1932-817011)

Nordic

Per Albin Hanssons väg 20 S-214 32 Malmö Sweden

(TEL:+46-40-190-800) (FAX:+46-40-190-450)

Italy

Via Galileo Galilei 40 I-20092 Cinisello Balsamo, Milano

(TEL:+39-2-618-38-431) (FAX:+39-2-618-38-402)

### ● Sony Corp. B&P Systems Co. ISP Dpt. (JAPAN)

4-16-1 Okata, Atsugi-shi, Kanagawa-ken, 243-0221

(TEL:+81-462-27-2345) (FAX:+81-462-27-2347)

Sony online <http://www.sony.co.jp/ISP/>

©1999 Sony Corporation. All rights reserved.

Reproduction in whole or in part without written permission is prohibited.

Design and specifications are subject to change without notice.

99B



# **ANNEXE B**

## **INSTRUCTIONS FOURNIES AUX SUJETS DE L'EXPERIMENTATION MENEES SUR CASPER-V2**





# Annexe B : Instructions Fournies aux Sujets de L'Expérimentation Menée sur CASPER-V2

## Cadre général

L'application que vous allez tester permet de présenter des informations relatives à une trajectoire tridimensionnelle à suivre avec un objet réel. La trajectoire est définie dans l'ordinateur et l'objet que vous allez manipuler capte des ondes et en déduit sa position et son orientation par rapport à une référence posée sur la table. L'ordinateur affiche une représentation de la trajectoire pré-planifiée ainsi qu'une représentation de la position et de l'orientation de l'objet que vous manipulez.

La représentation des informations se fait par le biais de croix ou d'une scène 3D que nous vous présenterons au cours de l'expérimentation. Ces informations peuvent être visualisées du point de vue de l'aiguille ou de la trajectoire. Nous vous les présenterons également au cours de l'expérimentation.

Au cours de ce test, vous aurez à utiliser deux supports différents pour l'affichage des données. Il s'agit d'une part d'un écran d'ordinateur classique et d'autre part, d'un casque de réalité augmentée. Le casque, semi-transparent, vous permet de visualiser les informations produites par l'ordinateur et de voir, au travers, le monde réel.

## Scénario :

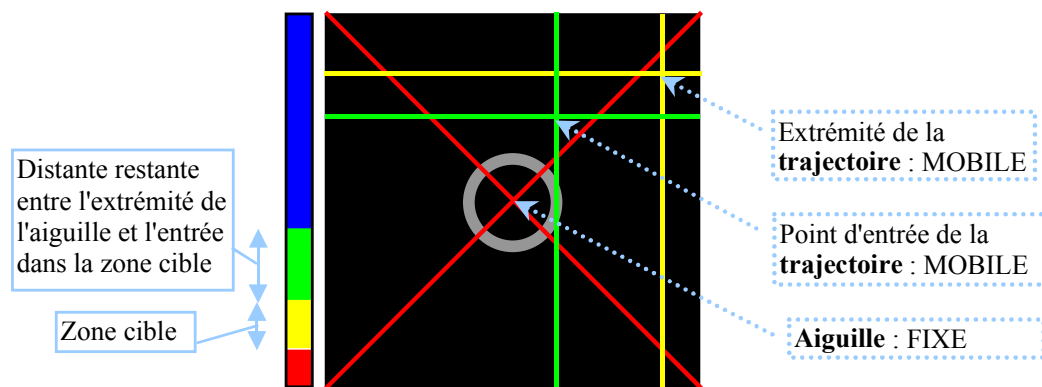
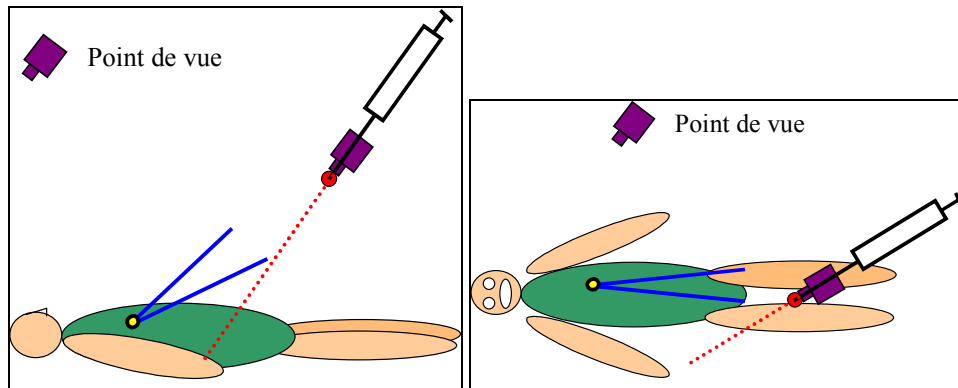
Avec chaque combinaison d'un point de vue (aiguille ou trajectoire), d'une représentation (croix ou cône) et d'un dispositif d'affichage (écran ou casque) que l'opérateur vous proposera, vous ajusterez l'objet réel jusqu'à ce que les informations affichées montrent qu'il est aligné avec la trajectoire prédéfinie. Puis en conservant au mieux cette orientation, vous avancerez l'objet jusqu'à entrer dans la zone cible. Veillez à ne pas passer au delà de la zone de cible.

Pour des raisons de vraisemblance avec la situation médicale, dans laquelle il s'agit de manipuler une aiguille et de vérifier sa rigidité, vous devrez, à chaque fois que vous entendrez le signal sonore, regarder l'objet réel et nous confirmer oralement par un "OUI" ou un "NON" que l'objet est toujours formé des deux parties le composant initialement : la partie grise et le crayon vert.

## Remarques

- Le temps mis à suivre la trajectoire n'est pas mesuré. Aussi appliquez vous à suivre la trajectoire proposée par l'ordinateur.
- L'objet réel doit être tenu avec la main au dessus de la partie grise.
- La position initiale de l'objet n'est jamais en dehors des limites de la table.
- Les déplacements du corps sont permis dans la mesure où il reste perpendiculaire à l'écran.
- Dans la mesure du possible, le corps ne doit pas être appuyé contre la table.
- Ne pas oublier de regarder l'aiguille à chaque beep, même lors de la phase d'alignement avec la trajectoire.

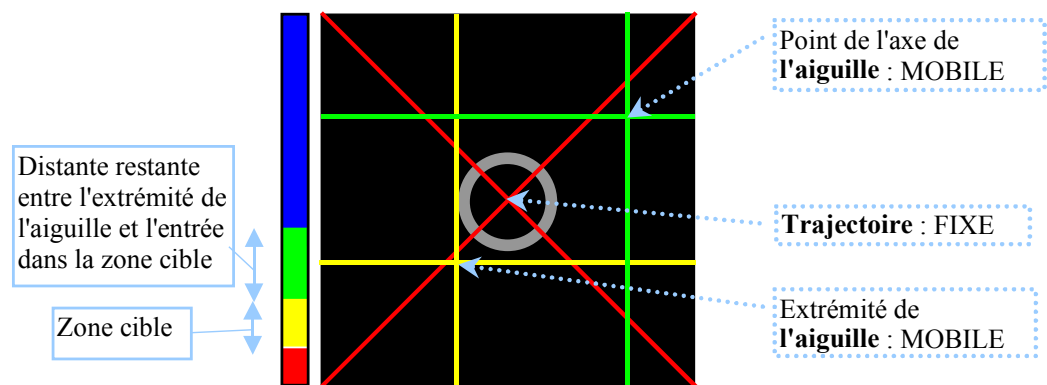
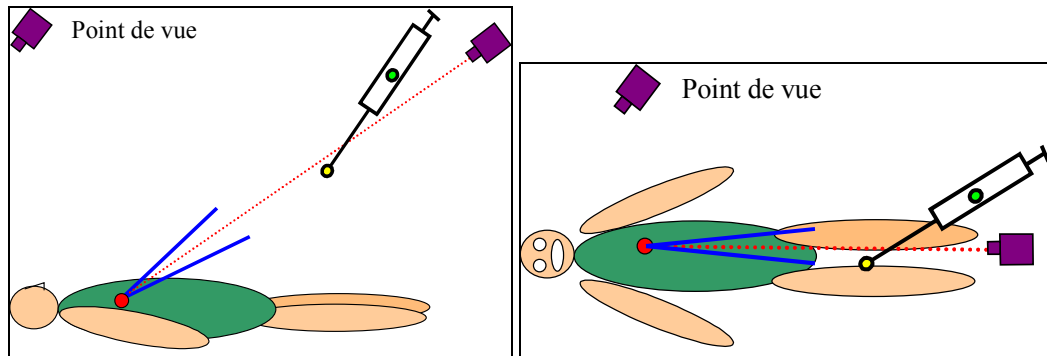
## Signification des croix : Point de vue = AIGUILLE



Chaque croix, représente la position d'un point de la trajectoire par rapport à l'objet réel que vous manipulez. Le disque gris représente une zone de laquelle les points représentés par les croix ne doivent pas sortir. Le curseur à gauche montre l'avancée de l'objet le long de la trajectoire ainsi que la distance restant à parcourir pour atteindre la zone cible.

En suivant les indications du scénario données ci-dessus, vous devez amener le curseur bleu de la jauge (partie gauche) au milieu de la zone cible indiquée en jaune. De plus, le point défini par chaque croix ne doit pas sortir des limites du disque gris.

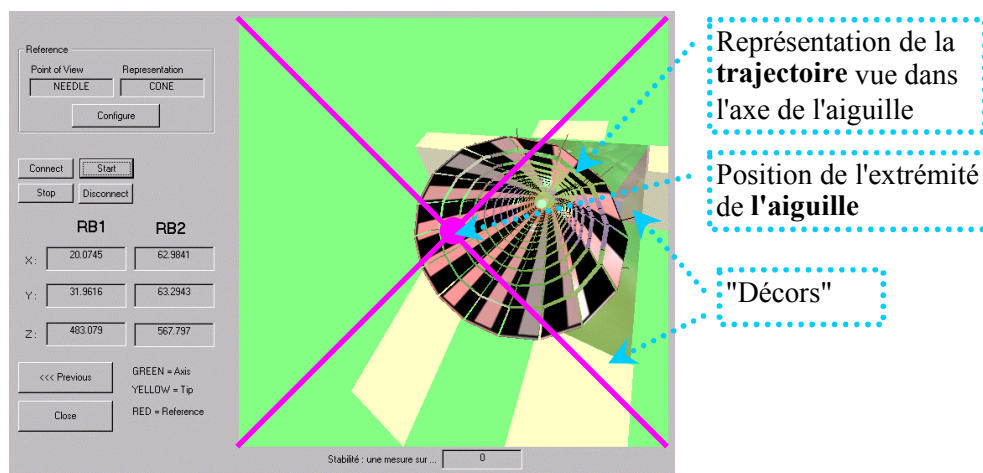
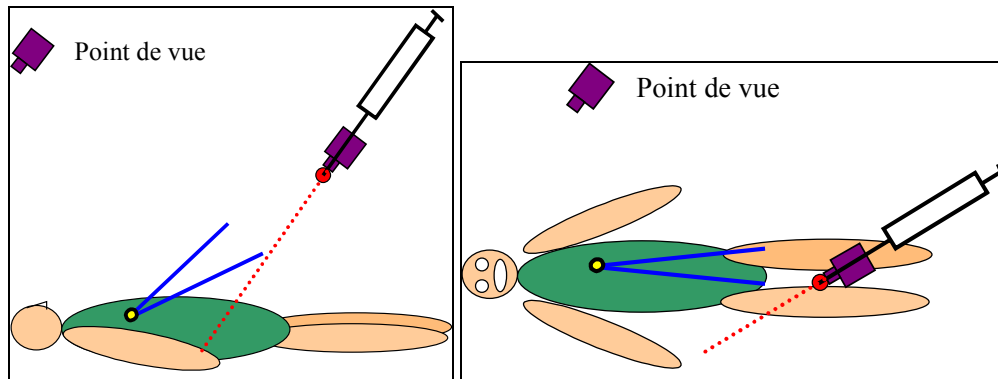
## Signification des croix : Point de vue = TRAJECTOIRE



Chaque croix, représente la position d'un point de l'objet réel que vous manipulez par rapport à la trajectoire prédéfinie. Le disque gris représente une zone de laquelle les points représentés par les croix ne doivent pas sortir. Le curseur à gauche montre l'avancée de l'objet le long de la trajectoire ainsi que la distance restant à parcourir pour atteindre la zone cible.

En suivant les indications du scénario données ci-dessus, vous devez réussir à amener le curseur bleu de la jauge (partie gauche) au milieu de la zone cible indiquée en jaune. De plus, le point défini par chaque croix ne doit pas sortir des limites du disque gris.

## Signification du cône: Point de vue = AIGUILLE

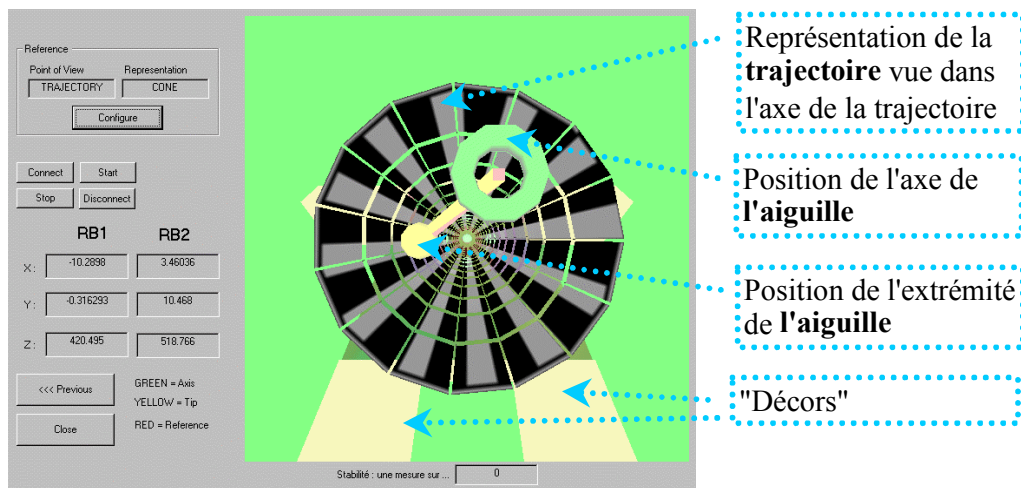
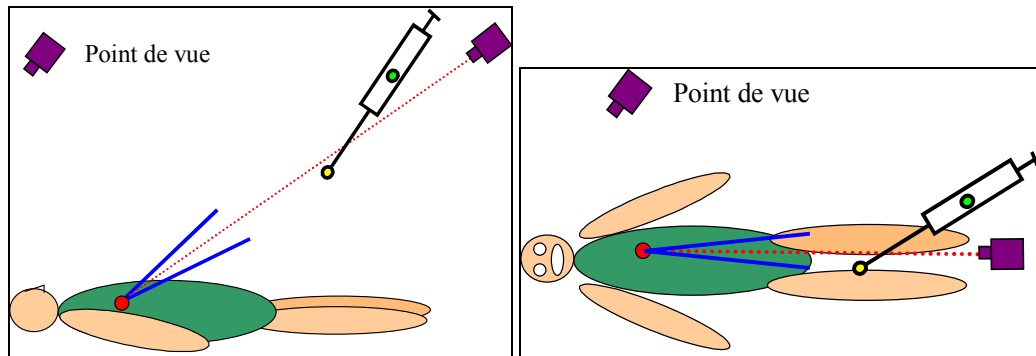


La scène 3D affichée se compose d'un cône représentant la trajectoire dans laquelle l'objet réel doit se déplacer, ainsi qu'un "décor" représentant très schématiquement la position d'un patient tel que la situation réelle le présenterait. La vue que vous obtenez de ce cône et du décor dépend de l'orientation et de la position de l'objet réel que vous manipulez.

En suivant les indications du scénario données ci-dessus, vous devez entrer complètement dans la zone cible située à l'intérieur du cône. L'entrée dans la zone est matérialisée par une surface rouge transparente. De plus vous ne devez pas sortir des limites du cône.

## Signification du cône :

### Point de vue = TRAJECTOIRE



La scène 3D affichée se compose d'un cône représentant la trajectoire dans laquelle l'objet réel doit se déplacer, d'un "décor" représentant très schématiquement la position d'un patient tel que la situation réelle le présenterait, et d'une représentation de l'objet réel que vous manipulez. La vue que vous obtenez de ce cône et du décor reste toujours la même. Par contre la position de la représentation de l'objet réel dépend de l'orientation et de la position de l'objet réel que vous manipulez.

En suivant les indications du scénario données ci-dessus, vous devez faire entrer entièrement la représentation de l'aiguille dans la zone cible située à l'intérieur du cône. L'entrée dans la zone est matérialisée par une surface rouge transparente. De plus vous ne devez pas faire sortir la représentation de l'aiguille en dehors des limites du cône.



# TABLES DES FIGURES

---



## Table des Figures

## Table des Figures

<b>Figure 1</b>	Schématisation de la relation entre le chirurgien et le monde de l'information.	22
<b>Figure 2</b>	Principes des systèmes de GMCAO : "Perception – Raisonement – Action".	23
<b>Figure 3</b>	Cadre d'illustration des systèmes de GMCAO.	26
<b>Figure 4</b>	Éléments d'une hanche artificielle (gauche) et interface de guidage de HipNav (droite). Photos issues de [Simon 97].	27
<b>Figure 5</b>	Schématisation du fonctionnement du système de Peuchot (droite) par rapport à un système de GMCAO basée sur un écran. Schémas issus de [Peuchot 95].	28
<b>Figure 6</b>	Informations affichées sur l'écran du système du MIT, optimisant la réalisation d'une craniotomie. Photo issue de [Grimson 96].	29
<b>Figure 7</b>	Vue dans le casque utilisé pour la laparoscopie. Photo issue de <a href="http://www.cs.unc.edu/~us/Laparo">www.cs.unc.edu/~us/Laparo</a> .	30
<b>Figure 8</b>	Données affichées par le système EchoSim.	32
<b>Figure 9</b>	Informations de guidage pour la visée pédiculaire.	33
<b>Figure 10</b>	Schéma de principe du dispositif utilisé pour l'"Image Overlay" de Blackwell. Photo issue de <a href="http://www.ri.cmu.edu/projects">www.ri.cmu.edu/projects</a> .	34
<b>Figure 11</b>	Schéma du principe de fonctionnement des templates. Schéma issu de <a href="http://sttl.hia.rwth-aachen.de">sttl.hia.rwth-aachen.de</a> .	37
<b>Figure 12</b>	Types de sources d'informations des systèmes de GMCAO.	40
<b>Figure 13</b>	Différence entre une cavité usinée dans le fémur manuellement (haut) et par Robodoc (bas).	41
<b>Figure 14</b>	Exemple de système de télé-opération en fonctionnement. (a) : partie locale au chirurgien, (b) partie locale au patient.	42
<b>Figure 15</b>	Résumé des solutions exploitées par les systèmes de GMCAO pour rendre perceptible de l'information lors de la phase d'action.	43
<b>Figure 16</b>	Illustration par des systèmes de GMCAO existants des différentes sources d'information possibles lors de la réalisation du geste.	44
<b>Figure 17</b>	Illustration du réfrigérateur augmenté : l'informatique classique (a) vs. le principe de la RA (b).	56
<b>Figure 18</b>	Triangle de la réalité virtuelle.	57
<b>Figure 19</b>	Triangles de la Réalité Augmentée (RA) et de la Réalité Virtuelle (RV).	58
<b>Figure 20</b>	Présentation des données réelles et virtuelles par Interior Design.	59
<b>Figure 21</b>	Illuminating Light en fonctionnement. Photo issue de <a href="http://tangible.media.mit.edu">tangible.media.mit.edu</a> .	60
<b>Figure 22</b>	Tangible Geospace en fonctionnement. Photo issue de <a href="http://tangible.media.mit.edu">tangible.media.mit.edu</a> .	60
<b>Figure 23</b>	Vue d'un utilisateur au travers du casque du Musée Augmenté. Photo issue de [Rekimoto 95].	62
<b>Figure 24</b>	Vue d'un utilisateur du système ARC au travers du casque semi-transparent. Photo issue de <a href="http://www.cs.columbia.edu">www.cs.columbia.edu</a> .	63
<b>Figure 25</b>	Dispositif du Paper-Based Audio Notebook. Photo issue de [Stifelman 96].	63
<b>Figure 26</b>	Éléments composant le Digital Desk.	65
<b>Figure 27</b>	Vue au travers du casque de l'un des joueurs de Mah-Jongg.	67
<b>Figure 28</b>	Systèmes mixtes : deux continua pour distinguer "Réalité Augmentée" et "Virtualité Augmentée".	84
<b>Figure 29</b>	Continuum de "Réalité Augmentée" et différentes façons d'enrichir la réalisation d'une tâche dont l'objet est une entité réelle.	85
<b>Figure 30</b>	Continuum de "Virtualité Augmentée" et différents types d'interaction pour enrichir la réalisation d'une tâche dont l'objet est une entité virtuelle.	86
<b>Figure 31</b>	Positionnement de systèmes mixtes du domaine des GMCAO dans notre cadre de classification.	89
<b>Figure 32</b>	Positionnement de systèmes mixtes du domaine des IHM dans notre cadre de classification.	91
<b>Figure 33</b>	Quatre composantes de l'utilisabilité. Schéma issu de <a href="http://www.usabilitymdx.com">http://www.usabilitymdx.com</a> .	94
<b>Figure 34</b>	Description avec ASUR du Musée Augmenté.	105
<b>Figure 35</b>	Caractérisation des composants participant à l'interaction dans le Musée Augmenté.	108
<b>Figure 36</b>	Caractérisation des relations participant à l'interaction dans le Musée Augmenté.	112

<b>Figure 37</b>	Descriptions avec ASUR de la visée pédiculaire, d'ACRobot, ainsi que du système de craniotomie assistée par ordinateur du MIT. _____	114
<b>Figure 38</b>	Descriptions avec ASUR de KARMA et du Musée Augmenté. _____	118
<b>Figure 39</b>	Description avec ASUR de la tâche de "Drag'n Drop" avec le Digital Desk, et de la manipulation d'un plan avec le Tangible Geospace. _____	120
<b>Figure 40</b>	Mise en relation des étapes du cycle de vie en V du logiciel (cadre simple) avec les étapes de conception de l'interaction (cadre double). _____	132
<b>Figure 41</b>	Propriétés ergonomiques : observabilité, honnêteté, compatibilité et continuité. _____	135
<b>Figure 42</b>	Schématisation du modèle ICS. _____	135
<b>Figure 43</b>	Caractéristiques ASUR pour l'analyse prédictive des compatibilités et des continuités perceptuelles et cognitives de l'interaction en sortie d'un système mixte. 138	
<b>Figure 44</b>	Description avec ASUR de la Visée Pédiculaire et de KARMA. _____	139
<b>Figure 45</b>	Composants du modèle d'architecture PAC-Amodeus [Nigay 94]. _____	145
<b>Figure 46</b>	Adaptation du modèle d'architecture PAC-Amodeus aux systèmes mixtes. _____	147
<b>Figure 47</b>	Coupe du cœur. _____	157
<b>Figure 48</b>	Echographie présentant un épanchement. _____	158
<b>Figure 49</b>	Matériel de localisation utilisé dans CASPER : les caméras du Polaris (à gauche), un foret muni d'un corps-rigide doté de 3 pastilles réfléchissantes (à droite). _____	161
<b>Figure 50</b>	Repères et matrices de passage utilisés pour la localisation de l'image échographique, de l'aiguille, du patient et de la référence dans CASPER. _____	162
<b>Figure 51</b>	Affichage des contours (verts) et de la zone fixe (rouge), issus de la phase de segmentation. _____	163
<b>Figure 52</b>	Organisation du bloc opératoire lors de l'utilisation de CASPER en phase de guidage. _____	164
<b>Figure 53</b>	Représentation courante de la position de l'aiguille par rapport à la trajectoire planifiée dans CASPER. <b>Cas a</b> ) : trajectoires différentes, <b>cas b</b> ) : trajectoires parallèles, <b>cas c</b> ) : trajectoires identiques. _____	165
<b>Figure 54</b>	Données de guidage présentées par CASPER : le viseur (à gauche), la jauge de profondeur (à droite). _____	166
<b>Figure 55</b>	Arbre de tâches minimal d'une ponction assistée par ordinateur, décrit avec la notation HTA. _____	167
<b>Figure 56</b>	Arbre de tâches effectif d'une ponction assistée par ordinateur, décrit avec la notation HTA. _____	167
<b>Figure 57</b>	Description avec ASUR de la tâche de ponction réalisée avec CASPER. _____	170
<b>Figure 58</b>	Caractérisation des composants participant à l'interaction dans CASPER. _____	171
<b>Figure 59</b>	Caractérisation des relations participant à l'interaction dans CASPER. _____	172
<b>Figure 60</b>	Schématisation de CASPER où réel et virtuel seraient fusionnés au sein d'une représentation virtuelle. _____	183
<b>Figure 61</b>	Schématisation du référentiel de présentation trajectoire pour la représentation des données dans CASPER-V2. _____	185
<b>Figure 62</b>	Schématisation du référentiel de présentation aiguille pour la représentation des données dans CASPER-V2. _____	186
<b>Figure 63</b>	Description avec ASUR de CASPER-V2. _____	187
<b>Figure 64</b>	Caractérisation des composants (a) et relations (b) participant à l'interaction dans CASPER-V2. _____	188
<b>Figure 65</b>	Structure du CD de CASPER-V2 selon le modèle PAC-Amodeus. _____	192
<b>Figure 66</b>	Structure effective du CD de CASPER-V2, intégrant le code existant de CASPER. _____	193
<b>Figure 67</b>	Architecture logicielle du pilier réel de CASPER-V2. _____	196
<b>Figure 68</b>	Présentation des données dans le référentiel de présentation trajectoire. _____	198
<b>Figure 69</b>	Schématisation des points Axe, Extrémité et Axe' de l'aiguille de ponction. _____	199
<b>Figure 70</b>	Présentation des données dans le référentiel de présentation aiguille. _____	199
<b>Figure 71</b>	Problème de superposition des profondeurs de champs de vision des scènes réelle et virtuelle. _____	201
<b>Figure 72</b>	Visualisation des données de guidage dans le référentiel de présentation utilisateur. _____	202
<b>Figure 73</b>	Combinaisons du référentiel de présentation, de la représentation et du support d'affichage, utilisables pour présenter les données de guidage dans CASPER-V2. _____	203

## Table des Figures

<b>Figure 74</b>	Boîte de dialogue permettant le choix de la représentation et du référentiel de présentation (point de vue) dans CASPER-V2. _____	204
<b>Figure 75</b>	Casque SONY équipé d'un corps-rigide blanc. _____	204
<b>Figure 76</b>	Schématisation des étapes de calibrage du casque. _____	205
<b>Figure 77</b>	Repères et matrices de passage requis pour l'affichage de données alignées avec le monde réel dans le casque semi-transparent. _____	206
<b>Figure 78</b>	Vue au travers du casque : superposition des données de guidage virtuelles à la vue du monde réel (saladier et bloc métallique). _____	207
<b>Figure 79</b>	Représentation des composantes x, y, z par rapport à la trajectoire prédéfinie. _____	211
<b>Figure 80</b>	Schéma de principe de l'analyse de la variance. _____	215
<b>Figure 81</b>	Résultats obtenus pour les facteurs et les combinaisons de facteurs présentant un effet significatif ( $p \leq 0.05$ ) ou une tendance ( $0,05 \leq p \leq 0,1$ ). _____	217
<b>Figure 82</b>	Effet du support. _____	219
<b>Figure 83</b>	Effet de la représentation. _____	219
<b>Figure 84</b>	Effet du référentiel de présentation. _____	220
<b>Figure 85</b>	Effet de l'axe. _____	221
<b>Figure 86</b>	Interaction entre représentation et axe. _____	222
<b>Figure 87</b>	Interaction entre la représentation, le référentiel de présentation et l'axe. _____	223
<b>Figure 88</b>	Interaction entre la représentation et le référentiel de présentation _____	224
<b>Figure 89</b>	Superposition de 3 plans de coupe porteurs d'images scanner et d'un modèle 3D du visage d'un patient. Photographie issue du manuel de programmation de Visual ToolKit (VTK). _____	230
<b>Figure 90</b>	Synthèse de nos résultats localisés dans un cycle de vie logiciel en V. _____	237

## Table des Figures

# **BIBLIOGRAPHIE GENERALE**

---

## Bibliographie Générale

## Bibliographie Générale

[Abdi 87]

Abdy, H., *"Introduction au Traitement Statistique des Données Expérimentales"*, Presses Universitaires de Grenoble, (1987), 420 pages.

[Abowd 92]

Abowd, D., Coutaz, J., Nigay, L., *"Structuring the Space of Interactive Properties"*, Acte de la conférence EHCI'92, (1992), p. 113-128.

[Ahlers 97]

Klinker, G., Ahlers, K., H. Breen, D., Chevalier, P.-Y., Crampton, C., Greer, D., S., Koller, D., Kramer, A., Rose, E., Tuceryan, M., Whitaker, R., *"Confluence of Computer Vision and Interactive Graphics for Augmented Reality"*, dans la revue *Presence: Teleoperators and Virtual Environments* (Special issue on Augmented Reality), 6, 4, (1997), p. 433-451.

[Aukstakalnis 92]

Aukstakalnis, S., Blatner, D., Silicon, *"Mirage - The Art and Science of Virtual Reality"*, Berkeley, CA, Peachpit Press, (1992).

[Azuma 95]

Azuma, R. T., *"Predictive Tracking for Augmented Reality"*, Thèse de l'Université de Caroline du Nord, Chapel Hill - USA, (1995), 242 pages.

[Azuma 97]

Azuma, R. T., *"A survey of Augmented Reality."*, dans la revue *Presence: Teleoperators and Virtual Environments* 6, 4, 1997, p.355-385

[Babbitt 01]

Babbitt, B., *"Medical Device Usability Engineering"*, Workshop de la conférence MMVR'2001, Newport Beach - CA, USA, (2001).

[Barnard 85]

Barnard, P., J., *"Cognitive Resources and the Learning of Computer Dialogs"*, dans la revue *Interfacing Thought, Cognitive Aspects of Human Computer Interaction*, Carroll, (ed.) MIT Press, (1985), p. 112-158.

[Barbe 97]

Barbe-Zoppis, C., *"Gestes Médico-Chirurgicaux sur tissus mous : application à la ponction péricardique assistée par échographie 2.5D"*, Thèse de l'Institut National Polytechnique de Grenoble, France, (1997).

[Baudel 93]

Baudel, T. and Beaudouin-Lafon, M., *"CHARADE: Remote Control of Objects using Free-Hand Gestures"*, *Communications de l'ACM*. 36(7): (1993), p. 28-35.

[Benford 93]

Benford, S., Fahlén, L., *"A Spatial Model of Interaction in Large Virtual Environments"*, Acte de la conférence European Conference on Computer Supported Cooperative Work (ECSCW'93), Milan, Italie, (1993), 16 pages.

[Bérard 99]

Bérard, F., *"Vision par ordinateur pour l'interaction homme-machine fortement couplée"*, Thèse de Doctorat de l'Université Joseph Fourier, Grenoble - France, (1999), 201 pages.



[Berlage 96]

Berlage, T., Fox, T., Grunst, G., Quast, K., J., "Supporting Ultrasound Diagnosis Using an Animated 3D Model of the Heart", Acte de la conférence IEEE Multimedia Computing and Systems, Hiroshima, Japon, (1996), 6 pages.

[Bernsen 94]

Bernsen, O., "Foundations of multimodal representations. A taxonomy of representational modalities", dans la revue *Interacting with Computers*, Vol. 6, 4, (1994), p.347-371.

[Blackwell 98]

Blackwell, M., Nikou, C., DiGioia, A., Kanade, T., "An Image Overlay System for Medical Data Visualization", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p.232-240.

[Bruley 99]

Bruley, C., "Analyse des Représentations Graphiques de l'Information - Extension aux représentations tridimensionnelles", Thèse de l'Université Joseph Fourier, Grenoble, (1999).

[Burdea 94]

Burdea, G., Coiffet, P., "Virtual Reality Technology", New York: John Wiley & Sons, (1994).

[Chavanon 97]

Chavanon, O., Barbe, C., Troccaz, J., Carrat, L., Ribuot, C., Blin, D., "Computer Assisted PERicardial punctures : animal feasibility study", Acte de CVRMed/MRCAS'97, Grenoble, (1997), p. 285-291.

[Chavanon 98]

Chavanon, O., "Ponction péricardique assistée par ordinateur. Guidage informatisé de la ponction péricardique", Rapport soumis au Comité de Consultation pour la Protection des Personnes pour la Recherche Biomédicale (CCPPRB) en vue d'avoir l'autorisation de faire des tests sur des patients, (1998).

[Chavanon 00]

Chavanon, O., Carrat, L., Pasqualini, C., Dubois, E., Blin, D., Troccaz, J., "Computer Guided Pericardiocentesis: Experimental results and Clinical Perspectives", publiée par Urban & Vogel dans la revue *Herz* 25, 8, (2000), p. 761-768.

[Chen 98]

Chen, N., Yu, W., and Pretlove, J.R. "Haptic interfaced telemanipulation", Acte du Symposium SPIE on Intelligent Systems and Advanced Manufacturing, Telemanipulator and Telepresence Technologies V, vol. 3524, Boston, (1998).

[Clark 89]

Clark, H., H., Schaefer, E., F., "Contributing to Discourse", dans la revue *Cognitive Science*, 13, (1989), p. 259-294.

[Cinquin 93]

P. Cinquin, "Gestes Médico-Chirurgicaux Assistés par Ordinateur", *Annales de radiologie*, Vol. 36 - n° 6/7, Europacs 93, (1993), p. 386-406.

[Coutaz 98]

Coutaz, J., Berard, F., Carraux, E., Crowlay, J., "Early Experience with the Mediaspace CoMedi", Acte de la Conférence EHCI'98, (1998).

[Cruz-Neira 93]

Cruz-Neira, C., Sandin, D., J., De-Fanti, T., A., "Surround Screen Projection-Based Virtual Reality: The Design and Implementation of the Cave", Acte de la Conférence SIGGRAPH'93, (1993), p. 135-142.

[Dix 98]

Dix, A., Finlay, A., Abowd, G., Beale, R., "Human-Computer Interaction", 2nd Edition, Prentice Hall, (1998).

[Droulez 89]

Droulez, J., Darlot, C., "The geometric and dynamic implications of the coherence constraints in three-dimentionnal sensorimotor interactions", dans M. Jeannerod (Ed.). *Attention and Performance XIII*, (1989), p. 495-526.

[Dubois 97]

Dubois, E., "Interaction Personne-Système : le cas particulier des Gestes Médico-Chirurgicaux Assistés par Ordinateur", rapport de DEA de l'Université Joseph Fourier, Grenoble - France, (1997), 113 pages.

[Edwards 95]

Edwards, P., J., Hill, D., L., G., Hawkes, D., J., Spink, R., Colchester, A., C., F., Strong, A., Gleeson, M., "Neurosurgical Guidance Using the Stereo Microscope", Acte de la conférence CVRMed'95, Ayache, N. (ed.), Springer Verlag, Nice, France, (1995), p. 555-564.

[Ellis 91]

Ellis, Gibbs, Rein, "Groupware: Some Issues and Experiences", communication de l'ACM, 34(1), (1991), p. 35-58.

[Evers 98]

Evers, H., Glombitza G, Mayer A, Meinzer HP. "Virtual Surgery: Methods and Areas of Application", Acte du forum TelecomIT, Frankfurt-am-Main, Allemagne, (1998).

[Feiner 93]

Feiner, S., MacIntyre, B., Seligmann, D., "Knowledge-Based Augmented Reality", Communication de l'ACM n°7, (1993), p. 53-61.

[Finlay 95]

Finlay, P., A., "Orthosista™: An Active Surgical Localiser for Assisting Orthopaedic Fracture Fixation", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 203-207.

[Fitzmaurice 93]

Fitzmaurice, G., "Situated Information Spaces and PalmTop Computers", Communication de l'ACM n°7, (1993), p.39-51.

[FOB]

Documentations techniques et Téléchargement de drivers, URL : <http://www.ascension-tech.com/products/flockofbirds/>

[Fortin 95]

Fortin, T., Coudert, J., L., Champlébourg, G., Sautot, P., Lavalée, S., "Computer Assisted Dental Implant Surgery Using Computed Tomography", dans la revue Journal of Image Guided Surgery, n°1, (1995), p. 53-58.

[Frankenthaler 98]

Frankenthaler, R., Moharir, V., M., Kikinis, R., van Kipshagen, P., Jolesz, F., A., Umans, C., Fried, M., P., "Virtual Otoscopy", dans la revue Otolaryngologic Clinics of North America. 31(2), (1998), p. 383-391. Egalement dans SPL Technical Report #76, (1998).

[Fuchs 98]

Fuchs, H., Livingston, M., A., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J., R., Rademacher, P., Drake, S., H., Meyer, A., "Augmented Reality Visualization for Laparoscopic Surgery", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p.934-943.

[Gaver 93]

Gaver, ., W., Sellen, A., Heath, C., Luff, P., "*One is Not Enough: Multiple Views in a MediaSpace*", Acte de la conférence INTERCHI'93, (1993), p. 335-341.

[Glauser 95]

Glauser, D., Fankhauser, H., Epitoux, M., Hefti, J.,L., Jacottet, A., "*Neurosurgical Robot Minerva. First Results and Current developments*", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 24 -30.

[Graham 00]

Graham, T. C. N., Watts, L., Calvary, G., Coutaz, J., Dubois, E., Nigay, L., "*A Dimension Space for the Design of Interactive Systems within their Physical Environments*", Acte de la conférence DIS'2000, (2000), p. 406-416.

[Gram 96]

Gram, C., Cockton, G., et al., "*Design Principles for Interactive Software*", Chapman et Hall, (1996), 248 pages.

[Green 95]

Green, P., S., Jensen, J., F., Hill, J., W., Shah, A., "*Mobile Telepresence Surgery*", Acte de la conférence MRCAS'95, Baltimore, Maryland USA, (1995), p. 97-103.

[Grimson 96]

Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells, W.M., Kikinis R., "*An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization*", dans la revue IEEE Transactions on Medical Imaging, 15(2), (1996), p. 129-140.

[Ham 98]

Ham, Van, G., Denis, K., Sloten, J., V., Audekercke, Van, R., Perre, Van der, G., Schutter, De, J., Aertbelien, E., Demey, S., Bellemans, J., "*Machining and Accuracy Studies for a Tibial Knee Implant Using a Force-Controlled Robot*", dans la revue Journal of Computer Aided Surgery, n°3, (1998), p. 123-133.

[Hasman 85]

Hasman, A., Lei, J. van der, Kwa, H.Y., Waage, M., "*A prototype expert system for the diagnosis and treatment of epilepsy*", dans Medical decision making: diagnostic strategies and expertsystems. (Bemmel, J.H. van, Gremy, F., Zvarova, J. eds), North Holland Publ Comp, Amsterdam, (1985), pp. 106-111.

[Hinckley 94]

Hinckley, K., et al., "*Passive Real-World Interface Props for Neurosurgical Visualization*", Actes de la conférence CHI'94, (1994), p.452-458.

[Ho 95]

Ho, S., C., Hibberd, R., D., Davies, B., L., "*Robot Assisted Knee Surgery*", dans la revue IEEE Engineering in Medicine and Biology Magazine, Special Issue on Robotics in Surgery, (1995), p. 292-300.

[Ikuta 98]

Ikuta, K., Takeichi, M., Namiki, T., "*Virtual Endoscope System with Force Sensation*", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p. 293-304.

[Ishii 97]

Ishii, H., Ullmer, B., "*Tangible Bits : Towards Seamless Interfaces between People, Bits and Atoms*", Acte de la conférence CHI'97, Atlanta, (1997), p. 234-241.

[Jolesz 97]

Jolesz, F., A., Lorensen, W., E., Shinmoto, H., Atsumi, H., Nakajima, S., Kavanaugh, P., Saiviroonporn, P., Setzer, S., E., Silverman, S., G., Phillips, M., Kikinis, R., "*Interactive*

*Virtual Endoscopy*", AJR, 169:1229-1235, SPL Technical Report #42, posted March 1997, (1997), 12 pages.

[Julliard 98]

Julliard, R., Lavallée, S., Dessenne, V., "*Computer Assisted Reconstruction of the Anterior Cruciate Ligament*", dans la revue *Clinical Orthopaedics and Related Research*, n° 354, (1998), p. 57-64.

[Kuzuoka 94]

Kuzuoka, H., Kosuge, T., Tanaka, M., "*GestureCam: a Video Communication System for Sympathetic Remote Collaboration*", Acte de la conférence CSCW'94, Chapel Hill - NC, USA, (1994), p. 35-43.

[Lavallee 89]

Lavallee, S., "*Gestes Médico-Chirurgicaux Assistés par Ordinateur : Application à la Neurochirurgie Stéréotaxique*", Thèse de l'Université Joseph Fourier, Grenoble, (1989).

[Lavallee 92]

Lavallée, S., et al., "*Image guided Robot : a clinical application in stereotactic neurosurgery*", Acte de la conférence IEEE Robotics and Automation, Nice, (1992), p. 618-625.

[Mackay 96]

Mackay, W.E., "*Réalité Augmentée : le Meilleur des Deux Mondes*", La Recherche n°285, (1996), p. 80-84.

[Mackay 98]

Mackay, W.E., Fayard, A.-L., Frobert, L., Médini, L., "*Reinventing the Familiar : an Augmented Reality Design Space for Air Traffic Control*", Acte de la conférence CHI'98, Los Angeles, (1998), p558-565.

[May 97]

May, J., Scott, S., Barnard, P., "*Modelling Multimodal Interaction: A Theory-based Technique for Design Analysis and Support*", Support de travaux dirigés (Tutorial) de la Conférence INTERACT'97, <http://www.shef.ac.uk/~pc1jm/guide.html>, (1997), 74 pages.

[McGurk 76]

McGurk, H., MacDonald, J., "*Hearing lips and seeing voices*", dans la revue *Nature*, n°264, (1976), p. 746-748.

[Merloz 97]

Merloz, P., Tonnetti, J., Eid, A., Faure, C., Lavallee, S., Troccaz, J., Sautot, P., Hammadeh, A., Cinquin, P., "*Computer Assisted Spine Surgery*", dans la revue *Clinical Orthopaedics and Related Research*, n°337, (1997), p. 86-96.

[Merloz 00]

Merloz, P., Lavallee, S., Tonnetti, J., Pittet, L., "*Image-Guided Spinal Surgery: Technology, Operative Technique and Clinical Practice*", dans la revue *Operative Techniques in Orthopaedics*, Vol 10, n°1, (2000), p. 56-63.

[Milgram 94]

Milgram, P., Kishino, F., "*A Taxonomy of Mixed Reality Visual Displays*", *Transactions on Information Systems*, E77-D(12), (1994), p.1321-1329.

[Miller 85]

Miller RA, Pople HE, Myers JD. INTERNIST-1, "*An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine*", dans Reggia J and Tuhrim S (eds), *Computer-Assisted Medical Decision-Making*, New York: Springer-Verlag (1985).

[Mittelstadt 93]

Mittelstadt, B., Paul, H., Kazanzides, P., Zuhars, J., Williamson, B., Pettitt, R., Cain, P., Kloth, D., Rose, L., Musits, B., "*Development of a Surgical Robot for Cementless Total*

*Hip Replacement*", dans la revue *Robotica*, vol. 11, Cambridge University Press, (1993), p. 553-560.

[Mori 95]

Mori, K., Hasegawa, J., I., Toriwaki, J., I., Anno, H., Katada, K., "*Automated Extraction and Visualization of Bronchus from 3D CT Images of Lung*", Acte de la Conférence CVRMed'95, Nice, France, (1995), p. 542-548.

[NaviCam]

URL : <http://www.csl.sony.co.jp/person/rekimoto/navi.html>

[Nigay 94]

Nigay, L., "*Conception et Modélisation logicielles des systèmes interactifs : applications aux interfaces multimodales*", Thèse de l'Université Joseph Fourier, Grenoble, (1994), 350 pages.

[Nigay 96]

Nigay L. et Coutaz J., "*Espaces conceptuels pour l'interaction multimédia et multimodale*", dans la revue *TSI*, spécial Multimédia et Collecticiel, AFCET & Hermès Publ. 15(9), (1996), pp. 1195-1225.

[Nigay 97]

Nigay, L., Coutaz, J., "*Software Architecture Modelling: Bridging Two Worlds Using Ergonomics and Software Properties*", dans Palanque, P., Paterno, F. (eds.) *Formal Methods in Human-Computer Interaction*, Springer Verlag, (1997), p. 49-73.

[Noma 96]

Noma, H., Miyasato, T., Kishino, F., "*A palmtop display for dextrous manipulation with haptic sensation*", Acte de la conférence *Human factors in computing systems*, 1996, p.126-133.

[Norman 86]

Norman, D., "*Cognitive Engineering*", In *User Centered System Design, New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, (1986), p.31-61.

[Okada 94]

Okada, K.-I., Maeda, F., Ichikawa, Y., Matsushita, Y., "*Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design*", Acte de la conférence CSCW'94, (1994), p. 385-393.

[Olfacom]

OLFACOM : Système de perception olfactif . Développé par la société A.C.2.I., URL : <http://www.olfacom.com/anglais/pages/presenta.html>.

[Ohlmann 90]

Ohlmann, T., "*Les systèmes Perceptifs Spatiaux Vicariants*", dans *Cognition : l'Individuel et l'Universel*, Psychologie d'Aujourd'hui, PUF, (1990), p. 21-58.

[Oltman 68]

Oltman, T., "*A Portable Od and Frame Apparatus, Perceptual and Motor Skills*", 26, (1968), p. 503-506.

[Oppenheimer 01]

Oppenheimer, P., Gupta, A., Weghorst, S., Porter, J., R., Sweet, R., "*The Representation of blood flow in Endoscopic Surgical Simulations*", Acte de la conférence MMVR'01, Newport Beach - CA, USA, (2001), p. 365-371.

[Peuchot 95]

Peuchot, B., Tanguy, A., Eude, M., "*Virtual Reality as an Operative Tool During Scoliosis Surgery*", Acte de la conférence CVRMed'95, Nice, France, (1995), p. 549-554.

[Pfaff 85]

Pfaff, G., E., "*User Interface Management Systems*", Eurographics Seminars, Springer Verlag, (1985).

[Phantom]

Phantom de Sensable Technologies, URL : [www.sensable.com/haptics/products/phantom.html](http://www.sensable.com/haptics/products/phantom.html)

[Poulose 98]

Poulose, B., Kutka, M., Mendoza-Sagaon, M., Barnes, A., Yang, C., Taylor, R., Talamini, M., "*Human Versus Robotic Organ Retraction During Laparoscopic Nissen Fundoplication*", Acte de la conférence MICCAI'98, Cambridge MA, USA, (1998), p. 197-206.

[Radermacher 96]

K. Radermacher, A. Zimolong, M. Anton, B. Thull, "*Recommandations for Ergonomic Analysis*", Deliverable 04 du projet IGOS, Lehrstuhl für Biomedizinische Technik der RWTH Aachen, (1996), 91 pages.

[Radermacher 97]

Radermacher, K., Portheine, F., Zimolong, A., Eichhorn, C., Staudte, H., W., Rau, G., "*Image Guided Orthopedic Surgery Using Individual Template*", Acte de la conférence CVRMed-MRCAS'97, Grenoble, France, (1997), p. 606-615.

[Rau 96]

Rau, G., Radermacher, K., Thull, B., Pichler, C. v., "*Aspects of an Ergonomic System Design of a Medical Worksystem*", in: Taylor, R., Lavallée, St., Burdea, G., Moesges, R. (eds.): *Computer Integrated Surgery*, MIT-Press, Cambridge, MA, 1996, pp.203-221

[Reich 96]

Reich, Y., Konda, S., L., Levy, S., N., Monarch, I., A., Subrahmanian, E., "*Varieties and Issues of Participation and Design*", dans la revue Design Studies, 17, 2, (1996), p. 165-180.

[Rekimoto 95]

Rekimoto, J., Nagao, K., "*The World through the Computer : Computer Augmented Interaction with Real World Environments*", Acte du symposium UIST'95, (1995).

[Renevier 01]

Renevier, P., Nigay, L., "*Mobile Collaborative Augmented Reality: the Augmented Stroll*", à paraître dans les actes de la conférence EHCI'01, Toronto, Canada, (2001).

[Sackier 93]

Sackier, J., et al., "*Computer-Aided Guidance for Endoscopic Surgery*", American College of Surgeons, San Francisco, (1993).

[Salb 01]

Salb, T., Brief, J., Burgert, O., Gockel, T., Giessler, B., Dillmann, R., "*Comparison of Tracking Presentation of Medical Data Using a See-Through Head-Mounted Display*", Acte de la Conférence MMVR'2001, Newport Beach - CA, USA, (2001).

[Salber 95]

Salber, D., "*De l'interaction homme-machine individuelle aux systèmes multi-utilisateurs*", Thèse de doctorat de l'Université Joseph Fourier - Grenoble I, (1995), 305 pages.

[Schneider 99]

Schneider, O., Troccaz, J., Chavanon, O., Blin, D., "*Synergistic robotic assistance in cardiac procedure*", Acte de la conférence CAR'99, (1999), p. 803-807.

[Schneiderman 96]

Schneiderman, B., "*The Eyes Have It: A Task by Data Taxonomy for Information Visualizations*", Acte de la conférence Visual Languages'96, Boulder, USA, (1996), p. 336-343.

[Scholz 98]

Scholz, M., Konen, W., Tombrock, S., Fricke, B., Adams, L., von Düring, M., Hentsch, A., Heuser, L., Harders, A., G., "*Development of an endoscopic navigation system based on digital image processing*", dans la revue Journal of Computer Aided Surgery, Volume 3, Issue 3, (1998), p. 134-143.

[Shahidi 98]

Shahidi, R., Wang, B., Epitoux, M., Grzeszczuk, Adler, J., "*Volumetric Image Guidance Via a Stereotactic Endoscope*", Acte de la Conférence MICCAI'98, Cambridge MA, USA, (1998), p. 241-252.

[Simon 97]

Simon, D. A., B. Jaramaz, M. Blackwell, F. Morgan, A.M. DiGioia, E., Kischell, B. Colgan, T. Kanade, "*Development And Validation Of A Navigational Guidance System For Acetabular Implant Placement.*", Acte de la conférence conjointe CVRMed / MRCAS, (1997), 583-592.

[Stifelman 96]

Stifelman, L., J., "*Augmenting real-world objects: a paper-based audio notebook*", dans le Tome II des actes de la conférence CHI '96, 1996, p. 199 - 200.

[Streitz 94]

Streitz, N., A., Geißler, J., Haake, J., M., Hol, J., "*DOLPHIN: Integrated Meeting Support across Local and Remote Desktop Environments and LiveBoards*", Acte de la conférence CSCW'94, (1994), p. 345-358.

[Szalavari 98]

Szalavari, Z., Eckstein, E., Gervautz, M., "*Collaborative Gaming in Augmented Reality*", Acte du symposium Virtual Reality Software and Technology, (1998), p. 195-204.

[Traum 96]

David Traum and Pierre Dillenbourg , "*Miscommunication in Multi-modal Collaboration*", dans working notes of the AAAI Workshop on Detecting, Repairing, And Preventing Human-Machine Miscommunication, (1996), p. 37-46.

[Troccaz 96-a]

J. Troccaz, S. Lavallée, P. Cinquin, "*Computer augmented surgery*", Human Movement Science 15, (1996), p. 445-475.

[Troccaz 96-b]

Troccaz, J., Delnondedieu, Y., "*Semi-Active Guiding Systems in Surgery. A Two-DOF Prototype of the Passive Arm with Dynamic Constraints (PADyC)*", dans la revue Mechatronics, Vol 6, n°4, Elsevier Science Ltd., (1996), p. 399-421.

[UIMS 92]

The UIMS Workshop Tool Developers : A Metamodel for the Runtime Architecture of an Interactive Systems, SIGCHI Bulletin, 24, 1, (1992), p. 32-37.

[Underkoffler 98]

Underkoffler, J., Ishii, H., "*Illuminating Light : An Optical Design Tool with a Luminous-Tangible Interface*", Acte de la conférence CHI'98, Los Angeles, (1998), p. 542-549.

[Vernier 01]

Vernier, F., "*La Multimodalité en Sortie et son Application à la visualisation de grandes quantités d'Information*", Thèse de l'Université Joseph Fourier, Grenoble, (2001), 272 pages.

## Bibliographie Générale

[Want 92]

Want, R., Hopper, A., Falcao, V., Gibbons, J., "*The Active Badge Location System*", dans *ACM Transactions on Information Systems*, 10(1), (1992), p. 91-102.

[Ware 96]

Ware, C., "Multimedia output devices and techniques", dans la revue *ACM Computing Surveys* (28, 1), (1996), pp. 181-183.

[Webster 96]

Webster, A., Feiner, S., MacIntyre B., Massie, W., Krueger, T., "*Augmented Reality in Architectural Construction, Inspection, and Renovation*", présenté au congrès ASCE Computers in Civil Engineering, (1996), p. 913-919.

[Webster 97]

Webster, A., & al., "*A Touring Machine : Prototyping 3D Mobile Reality Systems for Exploring the Urban Environment*", Acte de la conférence ISWC'97, Cambridge, (1997), p. 74-81.

[Wegner 97]

Wegner, Kristen and Karron, D.B. "*Surgical Navigation Using Audio Feedback*." Acte de la conférence MMVR'97, San Diego, (1997).

[Wellner 93]

Wellner, P., "*Interacting with paper on DigitalDesk*", *Communication del'ACM*, (36), 7, (1993).

[Whirlpool 01]

URL : [http://www.whirlpool.com.au/WhatsNew/Internet\\_fridge.htm](http://www.whirlpool.com.au/WhatsNew/Internet_fridge.htm)

[Whitaker 95]

Whitaker, R., T., Crampton, C., Breen, D., E., Tuceryan, M., Rose, E., "*Object Calibration for Augmented Reality*", Acte de la conférence Eurographics'95, Vol. 14, 3, (1995), p. 15-28.

[Winer 62]

Winer, B., J., "*Statistical Principles in Experimental Design*", *Mc Graw-Hill Book Company*. Perdue University, (1962).







---

## **Chirurgie Augmentée : un Cas de Réalité Augmentée; Conception et Réalisation Centrées sur l'Utilisateur**

Les travaux pluridisciplinaires de cette thèse, contribuent à deux domaines : les Interfaces Homme-Machine (IHM), en particulier la Réalité Augmentée (RA), et les Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO), domaine concret de mise en œuvre de la RA. La contribution est double : d'une part elle enrichit et organise les concepts d'IHM pour la RA en adoptant un point de vue centré sur la tâche, et d'autre part elle améliore l'utilisabilité par les cliniciens des systèmes de GMCAO.

La fusion du monde réel et du monde virtuel ouvre des voies nouvelles pour la conception de systèmes de RA, comme les systèmes de GMCAO. Mais les limites sont mal cernées. Dans ce contexte, nous avons identifié des caractéristiques intrinsèques des systèmes de RA et établi un espace taxinomique. Afin de caractériser plus finement la dualité de l'interaction (Utilisateur-Monde Réel et Utilisateur-Monde Virtuel), nous avons défini la notation ASUR. Pour décrire l'interaction, ASUR identifie les entités intervenant dans le système, en particulier les adaptateurs, composants ASUR réalisant le lien entre les deux mondes. L'utilisation d'ASUR, par ces aspects descriptifs, taxinomiques et conceptuels constitue la base de notre démarche de conception ergonomique et logicielle de systèmes de RA. Cette démarche se fonde en outre sur l'étude de propriétés ergonomiques au sein de la description ASUR d'un système et sur l'adaptation du modèle d'architecture PAC-Amodeus aux systèmes de RA.

En collaboration avec le chirurgien, nous avons appliqué et testé nos résultats méthodologiques sur une application concrète de chirurgie cardiaque. Cette application, CASPER, assiste le chirurgien lors de la réalisation d'une ponction péricardique, en lui procurant des informations de guidage du geste clinique sur un écran. Une re-conception de CASPER nous a amené à implémenter un casque semi-transparent permettant de visualiser ces informations directement sur le patient.

**Mots-Clés :** Interface Homme-Machine, Réalité Augmentée, Chirurgie Assistée par Ordinateur, Méthode de Conception Ergonomique et Logicielle

---

## **Augmented Surgery: an Augmented Reality Case Study; User-Centered Design and Realization**

The multi-disciplinary work of this thesis results in its contributing to two research domains: Human-Computer Interaction (HCI), in particular Augmented Reality (AR), and Computer Assisted Medical Intervention (CAMI), a concrete application domain of AR. The contribution is twofold: on the one hand, it enriches the HCI design concepts and methods of AR by adopting a task centered point of view; on the other hand it improves the clinical usability of CAMI systems.

Merging real and virtual worlds opens a wide range of possibilities for the design of AR systems, such as CAMI systems. However, there is currently no consensus either on a precise definition of AR or on a design space. Within this context, we have identified intrinsic characteristics of AR systems and then developed a classification space. In order to gain an understanding of the duality of interaction, (users interacting with part of the real world and simultaneously with part of the virtual world) we have developed the ASUR notation that describes the interaction. ASUR identifies entities involved in the system, in particular the adapters. The adapters constitute one type of ASUR component, which establishes a bridge between the real and virtual worlds. The notation ASUR constitutes the basis of our ergonomic and software design method for AR systems. Our ergonomic design method is grounded on the translation of ergonomic properties in terms of ASUR, while our software design method is based on the specialization of the PAC-Amodeus software architecture model for AR systems.

In collaboration with a surgeon, we have applied and tested our methodological results on a cardiac surgery application, CASPER. It provides the surgeon with on-screen guidance information while performing a pericardial puncture. A new version of CASPER has been designed that implements a see-through head-mounted display, which enables the perception of the guidance information superimposed on the patient.

**Keywords :** Human-Computer Interaction, Augmented Reality, Computer Assisted Surgery, Ergonomic and Software Design Method

