

**Ecole Centrale de Lyon**

THESE

Pour obtenir le grade de  
Docteur de l'Ecole Centrale de Lyon

Spécialité : Informatique

Présentée et soutenue publiquement par

**René CHALON**

Le 16 décembre 2004

---

**Réalité Mixte et Travail Collaboratif :  
IRVO, un modèle de  
l'Interaction Homme – Machine**

---

Directeur de thèse : M. Bertrand DAVID

Ecole Doctorale Informatique et Information pour la Société (EDIIS)

**Jury**

M. Michel BEAUDOUIN-LAFON	Président
M. Bertrand DAVID	Examineur
M. Alain DERYCKE	Rapporteur
M. Philip GRAY	Examineur
Melle Laurence NIGAY	Rapporteur
M. Dominique SCAPIN	Examineur



# Remerciements

Cette thèse n'aurait pu aboutir sans l'aide et le soutien de nombreuses personnes.

Je tiens tout d'abord à remercier mon Directeur de Thèse, Bertrand David, pour avoir bien voulu encadrer cette thèse portant sur un thème nouveau au laboratoire ICTT et de m'avoir soutenu tout au long du travail.

Je voudrais aussi remercier l'ensemble des membres du laboratoire ICTT pour leur aide et leur soutien. J'exprime plus particulièrement ma gratitude à Christian Bessière, Marie-France Kouloumdjian et Jacqueline Vacherand-Revel, qui m'ont fait découvrir les Sciences Humaines et leurs méthodes dans les temps glorieux de l'IRPEACS.

Je tiens également à remercier pour leur participation à ces travaux les étudiants de l'Ecole Centrale de projets d'Etude : Nadji Cherief, Julien Lasalle et Johan Moinard (promo 2002, projet Eye-tracking), Aurélien Dadou, Vincent Dorval, Emiliano Gambaretto, Mathieu Guingo et Pierre Munier (promo 2005, projet Polyphème), Jean-Philippe Bianchi, Thibaut Brusseau, Romain Clédat, Anna Michailovsky et Michel Sato (promo 2005, projet TableGate), Thomas Abbé, Vincent Dupuis, Yibo Fang, Pierre-Olivier Grégoire, Benoît Mathern (promo 2006, projet TableGate 2). Merci également à Magali Beldame pour son travail sur les modèles de tâches.

Je voudrais tout particulièrement remercier Monsieur le professeur Alain Derycke et Mademoiselle Laurence Nigay, pour avoir accepté d'être les rapporteurs de cette thèse, et qui, par les échanges que nous avons eu, m'ont permis d'améliorer le manuscrit.

Enfin, je remercie tous les autres membres du jury, MM. Michel Beaudouin-Lafon, Philip Gray et Dominique Scapin pour avoir accepté de participer à la soutenance et d'évaluer ce travail.



# Table des matières

<b>CHAPITRE I</b>	<b>INTRODUCTION .....</b>	<b>13</b>
<b>CHAPITRE II</b>	<b>LA REALITE MIXTE : UN PARADIGME POUR L'INTERACTION HOMME-MACHINE .....</b>	<b>17</b>
II.1	DEFINITIONS DE LA REALITE MIXTE.....	17
II.1.1	<i>La réalité augmentée – premières définitions.....</i>	17
II.1.2	<i>Objets réels et virtuels.....</i>	18
II.1.2.1	Définitions.....	18
II.1.2.2	La Réalité Virtuelle versus la Réalité Augmentée .....	19
II.1.3	<i>Continuum réel – virtuel.....</i>	20
II.1.4	<i>De la Réalité Augmentée à la Réalité Mixte.....</i>	21
II.1.5	<i>L'informatique ubiquitaire ou pervasive.....</i>	21
II.1.6	<i>Les interfaces tangibles : une forme de Réalité Mixte.....</i>	22
II.1.7	<i>Synthèse.....</i>	22
II.2	EXEMPLES DE SYSTEMES DE REALITE MIXTE .....	23
II.2.1	<i>Applications bureautiques.....</i>	23
II.2.1.1	Charade [Baudel 1993].....	24
II.2.1.2	DigitalDesk [Wellner 1993b].....	24
II.2.1.3	DoubleDigitalDesk [Wellner 1993b] .....	25
II.2.1.4	Chameleon [Fitzmaurice 1993] .....	25
II.2.1.5	Applications basées sur Navicam [Rekimoto 1995].....	26
II.2.1.6	Paper-based audio notebook [Stifelman 1996] .....	26
II.2.1.7	LivePaper [Robinson 2001] .....	27
II.2.2	<i>Applications pour la conception.....</i>	27
II.2.2.1	Video Mosaic [MacKay 1993, Mackay 1998a] .....	27
II.2.2.2	Build-IT [Fjeld 1999, Fjeld 2002a].....	28
II.2.2.3	SeamlessDesign [Kiyokawa 1999] .....	28
II.2.2.4	Réalité Mixte pour les premières étapes de la conception de produit [Santos 2003] .....	29
II.2.3	<i>Applications pour l'architecture et l'urbanisme .....</i>	29
II.2.3.1	Ariel [MacKay 1998a].....	29
II.2.3.2	Collaborative Interior Design [Ahlers 1995, Klinker 1997] .....	29
II.2.3.3	Architectural Anatomy [Webster 1996].....	30
II.2.3.4	Augmented Urban Planning Workbench [Ishii 2002].....	30
II.2.3.5	Collaborative City-Planning System [Kato 2003] .....	30
II.2.3.6	Bird's-eye View System [Kim 1999].....	31
II.2.4	<i>Applications pour la maintenance .....</i>	31
II.2.4.1	Karma [Feiner 1993] .....	31
II.2.4.2	SEAR [Goose 2003a, Goose 2003b] .....	32
II.2.4.3	Collaborative Mechanical Repair [Klinker 1997].....	32
II.2.5	<i>Applications pour l'enseignement, la formation et l'apprentissage.....</i>	32
II.2.5.1	SharedView et GestureCam [Kuzuoka 1994] .....	32
II.2.5.2	Training Tool for Assembly Tasks [Boud 1999] .....	33
II.2.5.3	MagicBook [Billinghurst 2001a, Billinghurst 2001b] .....	33
II.2.5.4	Augmented Chemistry [Fjeld 2002b].....	34
II.2.5.5	Caretta [Hosoi 2003] .....	35
II.2.6	<i>Applications dans le domaine médical.....</i>	35
II.2.6.1	CASPER v2 [Dubois 2001a, Dubois 2001b, Dubois 2003].....	35
II.2.6.2	Virtual Mirror Box [O'Neill 2003].....	36
II.2.6.3	Autres applications .....	36
II.2.7	<i>Applications pour l'archéologie, la paléontologie .....</i>	36
II.2.7.1	MAGIC [Renevier 2001, Nigay 2002, Renevier 2004a].....	36
II.2.7.2	Augmented Paleontology [Bimber 2002].....	37
II.2.8	<i>Applications pour les musées, les expositions et les visites de sites .....</i>	38
II.2.8.1	Musée Augmenté [Rekimoto 1995].....	38
II.2.8.2	LISTEN [Terrenghi 2004] .....	39
II.2.8.3	Touring Machine et MARS [Feiner 1997, Höllerer 1999a, Höllerer 1999b].....	39
II.2.9	<i>Jeux.....</i>	39

II.2.9.1	Mah-Jongg game [Szalavári 1998].....	39
II.2.9.2	Rasende Roboter [Demeure 2002] .....	40
II.2.9.3	TROC [Bouchet 2002, Renevier 2004a].....	40
II.2.10	<i>Autres domaines d'application</i> .....	41
II.2.10.1	Augmented library [Rekimoto 1995].....	41
II.2.10.2	Caméléon [Mackay 1997, Mackay 1998a, MacKay 1998b].....	41
II.2.10.3	Rasa, NISMap et NISChart [Cohen 2004].....	42
II.2.11	<i>Synthèse et discussion</i> .....	43
II.3	TAXONOMIES.....	46
II.3.1	<i>Taxonomie des types d'augmentation de Mackay</i> .....	46
II.3.2	<i>Taxonomie de Paul Milgram</i> .....	47
II.3.3	<i>Taxonomie de Bowskill</i> .....	48
II.3.4	<i>Taxonomies de Robinson &amp; Robertson</i> .....	49
II.3.5	<i>Taxonomie de Dubois</i> .....	50
II.3.6	<i>Taxonomie des systèmes interactifs mixtes de Nigay et al.</i> .....	50
II.3.7	<i>Taxonomies de Renevier</i> .....	50
II.3.7.1	Caractéristiques des augmentations.....	51
II.3.7.2	Taxonomie des systèmes mixtes collaboratifs et mobiles.....	51
II.3.8	<i>Synthèse des taxonomies</i> .....	52
II.4	BILAN.....	53

## CHAPITRE III DISPOSITIFS D'ENTREE/SORTIE POUR LA REALITE MIXTE ..... 55

III.1	INTRODUCTION.....	55
III.2	DISPOSITIFS EN SORTIE.....	56
III.2.1	<i>Dispositifs d'affichage</i> .....	56
III.2.1.1	Casques de visualisation ou HMD (Head-Mounted Display).....	56
III.2.1.2	Vidéoprojecteurs.....	57
III.2.1.3	Mini-écrans.....	58
III.2.1.4	Classes des systèmes de Réalité Mixte de Paul Milgram.....	58
III.2.2	<i>Restitution du son</i> .....	59
III.2.3	<i>Interfaces haptiques et kinesthésiques</i> .....	60
III.3	DISPOSITIFS EN ENTREE.....	60
III.3.1	<i>Dispositifs liés à la position, au mouvement, à la force</i> .....	60
III.3.1.1	Capteurs de positions – localisateurs (trackers).....	61
III.3.1.2	Vision artificielle.....	62
III.3.1.3	Gants numériques.....	63
III.3.1.4	Tablettes et tableaux tactiles ou numériques.....	63
III.3.2	<i>Périphériques d'entrée audio</i> .....	64
III.3.3	<i>Eye-trackers ou oculomètres</i> .....	64
III.3.4	<i>Interfaces cerveau-ordinateur</i> .....	65
III.4	SYNTHESE.....	66
III.4.1	<i>Influence des dispositifs sur l'interaction</i> .....	66
III.4.2	<i>Problème du choix des dispositifs par les concepteurs</i> .....	66
III.5	BILAN DU CHAPITRE.....	67

## CHAPITRE IV MODELES POUR L'INTERACTION HOMME-MACHINE..... 69

IV.1	MODELES D'INTERACTION.....	69
IV.1.1	<i>Théorie de l'action et manipulation directe</i> .....	70
IV.1.2	<i>Interaction instrumentale</i> .....	72
IV.1.3	<i>Modèle DPI (Document, Presentations, Instruments)</i> .....	73
IV.1.4	<i>Théorie de l'activité médiatisée par l'instrument</i> .....	74
IV.1.5	<i>Modèle MCRpd</i> .....	76
IV.1.6	<i>Modèles de l'Interaction physique de Sato</i> .....	76
IV.1.7	<i>Modèle spécifique pour les interfaces de Réalité Mixte</i> .....	77
IV.1.7.1	Formalisme ASUR.....	77
IV.1.7.2	ASUR++.....	79
IV.1.7.3	ASUR 2004.....	79
IV.1.7.4	Critique du modèle ASUR.....	80
IV.2	MODELES D'ARCHITECTURE.....	81
IV.2.1	<i>Modèles en couche</i> .....	81
IV.2.1.1	Modèle « Seheim ».....	81

IV.2.1.2	Modèle « Arch » .....	82
IV.2.2	Modèles <i>multi-agents</i> .....	82
IV.2.2.1	MVC.....	83
IV.2.2.2	PAC.....	83
IV.2.2.3	AMF.....	84
IV.2.3	Modèles Hybrides.....	86
IV.2.3.1	PAC-Amodeus.....	86
IV.2.3.2	AMF et Arch : AMF comme contrôleur de dialogue.....	87
IV.2.3.3	PAC-Amodeus modifié pour les systèmes mixtes.....	87
IV.2.3.4	Synthèse sur la portée des modèles d'architecture logicielle pour la Réalité Mixte.....	88
IV.3	BILAN DU CHAPITRE.....	89

## CHAPITRE V LE TRAVAIL COLLABORATIF ASSISTÉ PAR ORDINATEUR .....91

V.1	PRESENTATION DU TCAO ET DEFINITIONS.....	91
V.1.1	<i>Taxonomies</i> .....	92
V.1.2	<i>Le trèfle fonctionnel des collecticiels</i> .....	93
V.1.3	<i>IHM et les collecticiels</i> .....	94
V.2	MODELES D'INTERACTION.....	95
V.2.1	<i>Formalisme de Renevier pour les systèmes mixtes collaboratifs et mobile</i> .....	95
V.3	MODELES D'ARCHITECTURE.....	97
V.3.1	<i>Modèle de Dewan</i> .....	97
V.3.2	<i>Modèle PAC*</i> .....	98
V.3.3	<i>Modèle « Clover »</i> .....	99
V.3.4	<i>Modèle AMF-C</i> .....	100
V.3.4.1	AMF-C fragmenté.....	100
V.3.4.2	AMF-C répliqué.....	101
V.3.5	<i>PAC-Amodeus modifié pour les systèmes mixtes collaboratifs et mobiles</i> .....	102
V.3.6	<i>Synthèse sur les modèles d'architecture des collecticiels et la Réalité Mixte</i> .....	103
V.4	BILAN DU CHAPITRE.....	104

## CHAPITRE VI IRVO : NOTRE PROPOSITION POUR LA MODELISATION DE L'IHM DANS LE CONTEXTE DE LA REALITE MIXTE .....105

VI.1	PRESENTATION DU MODELE IRVO.....	105
VI.1.1	<i>Sources du modèle</i> .....	105
VI.1.2	<i>Objectifs</i> .....	107
VI.1.3	<i>Principales entités et relations d'IRVO</i> .....	107
VI.2	FORMALISME GRAPHIQUE D'IRVO.....	109
VI.2.1	<i>Frontières</i> .....	109
VI.2.1.1	Frontière entre les mondes réel et virtuel.....	109
VI.2.1.2	Frontières entre différents lieux du monde réel.....	110
VI.2.1.3	Frontières dans le monde virtuel.....	110
VI.2.2	<i>Caractéristiques communes aux entités</i> .....	110
VI.2.3	<i>Utilisateurs</i> .....	111
VI.2.4	<i>Objets</i> .....	112
VI.2.4.1	Objets réels et virtuels.....	112
VI.2.4.2	Objets mixtes.....	113
VI.2.5	<i>Modèle interne</i> .....	114
VI.2.6	<i>Transducteurs</i> .....	115
VI.2.6.1	Caractéristiques des transducteurs.....	115
VI.2.6.2	Distinction entre transducteur et dispositif d'entrée/sortie :.....	116
VI.2.7	<i>Relations</i> .....	117
VI.2.7.1	Relations partant ou aboutissant à un utilisateur vers un outil ou un objet.....	117
VI.2.7.2	Relations entre utilisateurs.....	118
VI.2.7.3	Relations entre les outils et les objets.....	119
VI.2.7.4	Relations entre le modèle interne et les outils et les objets.....	119
VI.2.7.5	Relations principales et secondaires.....	119
VI.2.7.6	Statut des transducteurs vis-à-vis des relations.....	120
VI.2.7.7	Fusions de relations.....	120
VI.2.7.8	Caractéristiques détaillées des relations.....	120
VI.2.8	<i>Propriété de mobilité des entités</i> .....	121
VI.2.8.1	Contraintes absolues.....	121

VI.2.8.2	Contraintes relatives .....	122
VI.2.8.3	Contraintes implicites.....	122
VI.2.9	<i>Résumé du formalisme d'IRVO</i> .....	123
VI.3	UTILISATION D'IRVO POUR LA DESCRIPTION DES SYSTEMES MIXTES .....	123
VI.3.1	<i>Exemples de modélisation avec IRVO</i> .....	124
VI.3.1.1	Modélisation de CASPER-v2.....	124
VI.3.1.2	Modélisation de Audio Notebook.....	125
VI.3.1.3	Modélisation du DoubleDigitalDesk.....	126
VI.3.1.4	Modélisation du jeu de Mah-Jongg.....	127
VI.3.1.5	Un autre exemple de modélisation.....	129
VI.3.2	<i>Notre Taxonomie des systèmes de Réalité Mixte</i> .....	130
VI.3.2.1	Taxonomie basée sur l'interaction d'un utilisateur.....	130
VI.3.2.2	Taxonomie pour les applications collaboratives et/ou complexes.....	132
VI.3.2.3	Application de notre taxonomie aux applications de la littérature.....	133
VI.3.3	<i>Comparaison d'IRVO avec ASUR et le formalisme de Renevier</i> .....	134
VI.3.3.1	Comparaison IRVO / ASUR 2004.....	134
VI.3.3.2	Comparaison IRVO / Formalisme de Renevier .....	136
VI.3.3.3	Bilan de ces comparaisons .....	137
VI.3.4	<i>Bilan sur les capacités descriptives d'IRVO</i> .....	137
VI.4	UTILISATION D'IRVO A DES FINS D'ANALYSES PREDICTIVES .....	138
VI.4.1	<i>Vérification de la complétude de l'interaction</i> .....	138
VI.4.2	<i>Vérification de l'observabilité</i> .....	139
VI.4.3	<i>Vérification de la cohérence des transducteurs vis-à-vis des relations</i> .....	139
VI.4.4	<i>Vérification de la continuité</i> .....	140
VI.4.5	<i>Vérifications liées au travail collaboratif</i> .....	141
VI.4.6	<i>Automatisation des vérifications</i> .....	142
VI.4.7	<i>Bilan sur les capacités prédictives d'IRVO</i> .....	142
VI.5	EVALUATION DU FORMALISME D'IRVO .....	142
VI.5.1	<i>Les « Principes de Namur »</i> .....	142
VI.5.2	<i>Evaluation d'IRVO selon les principes de Namur</i> .....	143
VI.5.3	<i>Critères d'évaluation de Brun</i> .....	143
VI.5.4	<i>Evaluation du formalisme d'IRVO selon les critères de Brun</i> .....	145
VI.5.5	<i>Bilan sur l'évaluation d'IRVO</i> .....	146
VI.6	BILAN DU CHAPITRE .....	146

## CHAPITRE VII CONCEPTION DE SYSTEMES DE REALITE MIXTE ..... 149

VII.1	APPROCHE BASEE SUR LES MODELES.....	149
VII.1.1	<i>Liste des modèles dans une approche basée sur les modèles</i> .....	149
VII.1.2	<i>Les modèles par rapport à l'architecture du logiciel</i> .....	151
VII.1.3	<i>Positionnement de IRVO dans une approche basée sur les modèles</i> .....	152
VII.2	ARBRE DE TACHES ET MODELE IRVO .....	153
VII.2.1	<i>Définitions</i> .....	153
VII.2.2	<i>Présentation de quelques formalismes</i> .....	153
VII.2.2.1	Formalisme MAD.....	154
VII.2.2.2	Formalisme CTT.....	154
VII.2.3	<i>Lien entre IRVO et l'arbre de tâche</i> .....	156
VII.2.3.1	Fusion de schémas IRVO et relations temporelles.....	156
VII.2.3.2	Schémas de synthèse d'une application.....	157
VII.2.4	<i>Bilan de la liaison entre IRVO et l'arbre de tâches</i> .....	158
VII.3	VERS UNE METHODOLOGIE DE CONCEPTION DE SYSTEMES DE REALITE MIXTE ET COLLABORATIFS.....	158
VII.3.1	<i>Présentation de la méthodologie CoC.Sys</i> .....	158
VII.3.2	<i>Insertion du modèle IRVO dans le modèle comportemental</i> .....	160
VII.3.2.1	Phase de recueil des scénarios.....	160
VII.3.2.2	Elaboration du modèle comportemental (CAB) .....	161
VII.3.2.3	Bilan.....	162
VII.3.3	<i>Conséquences de l'introduction de IRVO sur le processus de contextualisation, adaptation et spécialisation</i> ..	163
VII.3.3.1	Liens conceptuels entre IRVO et AMF-C.....	163
VII.3.3.2	Correspondances entre agents AMF-C et objets IRVO .....	164
VII.3.3.3	Correspondances entre agents AMF-C et outils IRVO .....	165
VII.3.3.4	Exemple de couplage de modèles IRVO et AMF-C .....	166
VII.3.3.5	Intérêt de ce couplage AMF-C / IRVO.....	167
VII.3.3.6	Bilan sur les modifications de l'architecture du logicielle .....	167

VII.3.4	<i>Bilan la méthodologie de conception</i> .....	168
VII.4	COMPARAISON AVEC D'AUTRES METHODOLOGIES DE CONCEPTION POUR LES SYSTEMES MIXTES.....	168
VII.4.1	<i>Approche de Trevisan et al.</i> .....	168
VII.4.2	<i>Approche de Renevier</i> .....	169
VII.5	BILAN DU CHAPITRE .....	170

**CHAPITRE VIII PROJET ACCOREA — AIDE A LA CONCEPTION COOPERATIVE  
SUPPORTEE PAR LA REALITE AUGMENTEE — ET PROTOTYPES  
ASSOCIES.....173**

VIII.1	DESCRIPTION DU PROJET ACCOREA.....	173
VIII.2	DEMARCHE.....	174
VIII.3	TRAVAUX PRELIMINAIRES.....	175
VIII.3.1	<i>Rédaction collaborative</i> .....	175
VIII.3.2	<i>Modélisation – arbre de tâches</i> .....	176
VIII.4	PROJET TABLEGATE ET PROTOTYPE WORDTAB.....	177
VIII.4.1	<i>Objectifs de WordTab</i> .....	177
VIII.4.2	<i>Conception de WordTab</i> .....	178
VIII.4.2.1	Scénario d'utilisation .....	178
VIII.4.2.2	Construction du modèle comportemental.....	178
VIII.4.3	<i>Réalisations</i> .....	180
VIII.4.3.1	Dispositif technique - Table interactive .....	180
VIII.4.3.2	Logiciel WordTab.....	182
VIII.4.4	<i>Expérimentation de WordTab</i> .....	185
VIII.4.4.1	Protocole expérimental.....	185
VIII.4.4.2	Déroulement des tests .....	187
VIII.4.4.3	Résultats issus des observations.....	188
VIII.4.4.4	Résultats issus des questionnaires.....	189
VIII.4.4.5	Bilan des tests .....	190
VIII.5	PROJET TABLEGATE II.....	191
VIII.6	BILAN DU CHAPITRE .....	192

**CHAPITRE IX CONCLUSION.....195**

IX.1	BILAN DE NOS CONTRIBUTIONS .....	195
IX.2	PERSPECTIVES .....	198

**BIBLIOGRAPHIE .....201**



# Table des figures

Figure 1 : Triangles de la Réalité Augmentée (RA) et de la Réalité Virtuelle (RV) (d'après [Dubois 2001])	20
Figure 2 : Le "Continuum réel – virtuel" (d'après [Milgram 1994a])	20
Figure 3 : Comparaison des styles d'interaction (d'après [Rekimoto 1995])	23
Figure 4 : Schéma du DigitalDesk (d'après [Wellner 1993b])	24
Figure 5 : Prototypage de l'Audio Notebook (d'après [Stifelman 1996])	27
Figure 6 : Vue générale du dispositif Build-IT (d'après [Fjeld 2002])	28
Figure 7 : Vue détaillée d'une interaction avec deux briques (d'après [Fjeld 2002])	28
Figure 8 : Karma (d'après [Feiner 1993])	31
Figure 9 : Système SEAR (d'après [Goose 2003a])	32
Figure 10 : les 3 modes d'utilisation du MagicBook : (a) réalité ; (b) Réalité Augmentée ; (c) immersion en Réalité Virtuelle (d'après [Billighurst 2001b])	33
Figure 11 : La paillasse augmentée de Augmented Chemistry (d'après [Fjeld 2002])	34
Figure 12 : Prototypage du système CASPER v2 (d'après [Dubois 2001b])	36
Figure 13 : Interface de MAGIC sur le Tablet PC ; la fenêtre en haut à droite est la vue qui est affichée sur le casque de visualisation semi-transparent (d'après [Renevier 2004a])	37
Figure 14 : Musée augmenté : vue prise à travers le casque semi-transparent (d'après [Rekimoto 1995])	38
Figure 15 : Jeu de Mah-Jongg augmenté vu par le casque d'un joueur (d'après [Szalavári 1998])	40
Figure 16 : TROC : récupération d'un objet virtuel par un cube réel (d'après [Renevier 2004a])	41
Figure 17 : Vue du système Caméléon montrant les strips, la tablette tactile en dessous et un écran tactile sur la droite (d'après [Mackay 1998a])	42
Figure 18 : Tableau de synthèse de la revue des applications de Réalité Mixte (1 <sup>ère</sup> partie)	44
Figure 19 : Tableau de synthèse de la revue des applications de Réalité Mixte (2 <sup>ème</sup> partie)	45
Figure 20 : Taxonomie de Paul Milgram (d'après [Milgram 1994a])	47
Figure 21 : Taxonomie de Dubois (d'après [Dubois 2001b])	50
Figure 22 : Taxonomie des 4 types d'interactions (d'après [Nigay 2003])	50
Figure 23 : taxonomie des caractéristiques d'une augmentation (d'après [Renevier 2003c])	51
Figure 24 : Taxonomie de Mackinlay des périphériques d'entrée sur laquelle ont été placés les périphériques considérés par Buxton (d'après [Mackinlay 1990])	61
Figure 25 : les sept étapes de l'activité de l'utilisateur (d'après [Norman 1986])	70
Figure 26 : les deux gouffres de l'exécution et de l'évaluation (d'après [Norman 1986])	71
Figure 27 : Modèle de l'interaction instrumentale (adapté de [Beaudouin-Lafon 2000a])	72
Figure 28 : Modèle conceptuel de DPI (d'après [Beaudoux 2001])	74
Figure 29 : Modèle de l'activité médiatisée par l'instrument (d'après [Rabardel 2003b])	75
Figure 30 : Modèle MCRpd (d'après [Ullmer 2000])	76
Figure 31 : Description de CASPER-V2 avec ASUR (d'après [Dubois 2001a])	78
Figure 32 : Représentation de CASPER-v2 avec ASUR 2004 (d'après [Juras 2004])	80
Figure 33 : Modèle Seeheim	81
Figure 34 : Modèle « Arch » (adapté de [UIMS 1992])	82
Figure 35 : Modèle MVC	83
Figure 36 : Modèle PAC	84
Figure 37 : Représentation de la composition d'agents AMF (d'après [Tarpin-Bernard 1997])	85
Figure 38 : Représentation des ports de communication (d'après [Tarpin-Bernard 1997])	85
Figure 39 : Représentation schématique des administrateurs (d'après [Tarpin-Bernard 1997])	85
Figure 40 : Modélisation AMF d'un logiciel de gestion d'agenda (d'après [Tarpin-Bernard 1999])	86
Figure 41 : Modèle Hybride PAC-Amodeus (d'après [Coutaz 2001])	87
Figure 42 : Les couches de l'implémentation de AMF	87
Figure 43 : Adaptation du modèle PAC-Amodeus aux systèmes mixtes (d'après [Dubois 2001b])	88
Figure 44 : Taxonomie espace-temps des collecticiels (d'après [Ellis 1991])	92
Figure 45 : Le trèfle fonctionnel des collecticiels	93
Figure 46 : Modèle « 3C » et les tâches (d'après [David 2001a])	94
Figure 47 : Formalisme de Renevier : (a) représentation des utilisateurs ; (b) représentation des objets ; (c) représentation des interactions et relations spatiales (d'après [Renevier 2004c])	96
Figure 48 : Exemple de représentation de scénario de Navicam+ (d'après [Renevier 2004c])	97
Figure 49 : Modèle de Dewan (d'après [Dewan 1998])	98
Figure 50 : Modèle PAC* (d'après [Calvary 1997])	99
Figure 51 : Modèle Clover (d'après [Laurillau 2002c])	100
Figure 52 : Exemple d'agent AMF fragmenté (d'après [Tarpin-Bernard 1998])	101
Figure 53 : AMF-C répliqué ; (a) gestion centralisée ; (b) gestion répartie (d'après [Tarpin-Bernard 1997])	102

Figure 54 : Exemple d'agents AMF-C répartis (d'après [Delotte 2003]).....	102
Figure 55 : Modèle PAC-Amodeus modifié et les 3 flux de données contextuelles (d'après [Renevier 2003c]).....	103
Figure 56 : Les principales entités et leurs relations.....	108
Figure 57 : Exemples de modélisations du travail collaborative : a: outils et objets partagés; b: outils et objets propres à chaque utilisateur.....	108
Figure 58 : Représentation des frontières.....	109
Figure 59 : Représentation d'un utilisateur.....	111
Figure 60 : Représentations des outils et des objets.....	112
Figure 61 : Représentation d'un objet mixte.....	113
Figure 62 : Représentation du Modèle interne de l'application.....	114
Figure 63 : Représentations des transducteurs.....	115
Figure 64 : Représentation d'un dispositif de pointage simple avec IRVO. (a) : modèle complet ; (b) : représentation simplifiée en fusionnant l'outil réel et le transducteur.....	117
Figure 65 : Relations entre un utilisateur et les objets.....	118
Figure 66 : Représentation de relations interindividuelles.....	118
Figure 67 : Exemple de représentation d'une communication médiatisée avec IRVO.....	119
Figure 68 : Représentation de la perception d'un objet mixte par un utilisateur.....	120
Figure 69 : Représentation des contraintes spatiales.....	121
Figure 70 : Représentation des contraintes géométriques relatives.....	122
Figure 71 : Tableau résumant le formalisme d'IRVO.....	123
Figure 72 : Modélisation de CASPER-v2+ avec IRVO (à droite) Le modèle ASUR du même dispositif est reproduit à gauche (Cf. Figure 31).....	124
Figure 73 : Modélisation de la tâche « lecture des notes » de Audio Notebook avec IRVO.....	126
Figure 74 : Modélisation du DoubleDigitalDesk avec IRVO.....	126
Figure 75 : Modélisation de la tâche « préparer des combinaisons » du jeu de Mah-Jongg avec IRVO.....	127
Figure 76: Modélisation de la tâche « poser une combinaison sur la table de jeu commune » du jeu de Mah-Jongg avec IRVO.....	128
Figure 77 : Modélisation de l'interaction bi-manuelle avec un <i>Toolglass</i> .....	129
Figure 78 : Notre taxonomie des systèmes de réalité mixte (pour un utilisateur).....	131
Figure 79 : Les 44 applications recensées au Chapitre II placées dans notre taxonomie ainsi que les <i>Toolglasses</i> (§ VI.3.1.5) et l'application WordTab du projet Accoréa (présenté au Chapitre VIII).....	134
Figure 80 : Comparaison de la modélisation de CASPER-v2+ avec : à gauche, le formalisme ASUR 2004 (d'après [Juras 2004]), à droite, le formalisme IRVO.....	135
Figure 81 : Comparaison de la modélisation du Navicam+ : à gauche, le formalisme de Renevier (d'après [Renevier 2004c]), à droite, le formalisme IRVO.....	136
Figure 82 : Evaluation du formalisme d'IRVO selon les « principes de Namur ».....	143
Figure 83 : Evaluation d'IRVO selon les critères de Brun.....	145
Figure 84 : liens entre une approche basée sur les modèles et le modèle Arch.....	152
Figure 85 : Exemple de modélisation de tâches collaboratives avec CTT (d'après [Mori 2002]).....	155
Figure 86 : Liens entre l'arbre de tâches et les schémas IRVO.....	156
Figure 87 : Exemple de schéma IRVO correspondant à la Tâche 1.....	156
Figure 88 : Exemple de schéma de synthèse d'une application mono-utilisateur.....	157
Figure 89 : Exemple de schéma de synthèse d'une application collaborative.....	157
Figure 90 : Principaux constituants de l'environnement de conception d'applications collaboratives mobiles et adaptatives (d'après [Delotte 2003]).....	159
Figure 91 : première phase de la méthodologie CoCSys avec le modèle comportemental étendu.....	160
Figure 92 : Exemple d'utilisation d'IRVO pour la représentation de scénarios sur le système Navicam+.....	161
Figure 93 : Relations conceptuelles entre les modèle Arch, AMF-C et IRVO.....	164
Figure 94 : Correspondance entre les agents AMF-C et les objets IRVO.....	164
Figure 95 : Exemple de correspondance agent AMF / artefact IRVO.....	166
Figure 96 : Modélisation de la tâche de rédaction d'article scientifique (d'après [Beldame 2001]).....	176
Figure 97 : Modélisation des tâches principales de WordTab avec CTT.....	179
Figure 98 : Activité de co-édition en local avec le logiciel WordTab représentée avec IRVO.....	180
Figure 99 : Schéma de notre dispositif de Table Interactive.....	181
Figure 100 : Photographie de notre Table Interactive (ancienne installation).....	181
Figure 101 : Interface de WordTab en mode « positionnement des feuilles ».....	183
Figure 102 : Interface de WordTab en mode édition.....	184
Figure 103 : Exemple d'utilisation de WordTab avec 3 utilisateurs.....	187
Figure 104 : Modélisation avec IRVO de la tâche « déplacement d'un camion » dans TableGate II.....	192

# Chapitre I Introduction

*Reality is merely an illusion, albeit a very persistent one.*  
Albert Einstein (1879 – 1955).

Cette thèse s'inscrit dans le domaine de l'Interaction Homme-Machine (IHM) et porte sur la modélisation et la conception des systèmes de Réalité Mixte en particulier dans le contexte d'applications collaboratives.

## Motivations

L'Interface Homme-Machine représente la partie visible d'un système informatisé à tel point que pour l'utilisateur « l'interface c'est l'application ». Depuis les débuts de l'informatique, elle a considérablement évoluée. Aux âges « préhistoriques » (une cinquantaine d'années seulement !), les opérateurs introduisaient les programmes et les données directement en binaire au moyen d'interrupteurs (jeu de clés) ou grâce à des cartes et rubans perforés. Dès les années soixante, des terminaux clavier/imprimantes (les télétypes) puis des terminaux clavier/écran ont permis une saisie plus conviviale mais toutefois limitée au texte.

Au-delà de la vision proposée par Vannevar Bush [Bush 1945] et de son système Memex, la vraie évolution (sinon révolution) est intervenue à la fin des années soixante grâce, en particulier, aux travaux de Doug Engelbart (à l'Université de Standford) puis de Alan Kay (au Xerox PARC) [Engelbart Symposium 1998], avec l'apparition des interfaces graphiques et de la souris comme moyen d'interaction. Ces nouvelles interfaces, et l'effet conjugué d'une baisse drastique des coûts du matériel, ont permis une explosion de l'usage de l'informatique bien au-delà des spécialistes au point d'en faire aujourd'hui un outil indispensable à toute profession.

Depuis l'avènement de l'ordinateur individuel, les interfaces graphiques WIMP (Windows, Icon, Menus and Pointing devices) et la métaphore du bureau se sont imposés comme étant le style d'interaction dominant et la majorité des boîtes à outils sont basées sur ces principes. Cependant, malgré sa puissance et son universalité, la métaphore du bureau telle que nous la connaissons aujourd'hui, montre ses limites dans de nombreuses situations : personnes en situation de travail critique (pilotes d'avions, de machines, ...), personnes handicapées, ...

L'un des principaux reproches tient au fait que le système informatique redéfinit un univers avec ses propres lois déconnectées du monde réel dans lequel les êtres humains évoluent. Même pour des applications bureautiques simples comme un traitement de texte, les utilisateurs gèrent souvent deux versions des documents : l'une électronique conservée dans la mémoire de l'ordinateur et échangeable à volonté avec ses collaborateurs par messagerie électronique, l'autre papier quasiment indispensable à une lecture correcte, permettant des annotations aisées, transportable, échangeable, mais désynchronisée de la version électronique [Wellner 1993b].

Depuis une dizaine d'années, la communauté informatique cherche à réconcilier les 2 mondes : le monde réel, l'environnement physique de l'utilisateur, et le monde virtuel qui regroupe les moyens de traitement, de stockage et de communication proposés par l'informatique. On peut citer 2 grandes tendances :

- La Réalité Augmentée (Augmented Reality), dont l'acte fondateur peut être situé par le dossier spécial de *Communications of the ACM* de juillet 1993, intitulé « Computer-Augmented Environments : Back to the Real World » et proposé par Pierre Wellner, Wendy Mackay et Rich Gold [Wellner 1993a].

- L'informatique ubiquitaire (Ubiquitous Computing), initiée par le célèbre papier de Mark Weiser, « The Computer for the 21st Century » paru en 1991 dans *Scientific American* [Weiser 1991].

Ces deux approches partagent la même philosophie basée sur la primauté du monde réel sur le monde virtuel, mais elles diffèrent par les technologies mises en œuvre : la Réalité Augmentée garde un approche assez centrée sur l'ordinateur qu'il soit fixe ou porté (ordinateur portable, PDA, *wearable computer*) alors que l'informatique ubiquitaire repose plus sur une informatique disséminée, embarquée dans les objets contenant des capacités informatiques et des périphériques d'entrée-sortie et qui sont reliés entre eux [Rekimoto 1995]. Cependant, en faisant une revue de la littérature, nous avons vu que les frontières ne sont pas aussi claires, et que de nombreux systèmes peuvent se réclamer des 2 approches.

### *Objectifs de la thèse et démarche*

Nous nous intéressons dans cette thèse à la première approche, celle de la Réalité Augmentée qui est aujourd'hui nommée, de manière plus générale, Réalité Mixte. La Réalité Mixte s'intéresse aux systèmes interactifs dans lesquels des objets réels et des données informatiques sont mêlées de manière cohérente.

Depuis une dizaine d'année, de nombreuses applications dans ce domaine ont été conçues, ainsi que de nombreuses technologies associées. On constate que la plupart ont été créés de manière indépendante et autonome. Il nous semble maintenant que ce domaine est suffisamment mûr pour dépasser ce stade d'expérimentation et qu'il est nécessaire de dégager des modèles et des méthodologies de conception, condition *sine qua non* pour que le concept de Réalité Mixte puisse essaimer au-delà des laboratoires de recherche.

Ceci constitue l'objectif de notre thèse. Notre but est d'aider à la conception de systèmes de Réalité Mixte en particulier dans le contexte des applications collaboratives.

En effet, on constate aujourd'hui que dans le monde de l'entreprise, le travail collaboratif (TCAO : travail collaboratif assisté par ordinateur) est devenu incontournable, les applications mono-utilisateurs évoluant progressivement de façon à intégrer des fonctionnalités multi-utilisateurs. Actuellement, La grande majorité des applications professionnelles comporte donc une part - plus ou moins importante - de travail collaboratif..

Par conséquent, l'objectif essentiel de cette thèse est d'élaborer un nouveau modèle de l'interaction, que nous avons nommé IRVO (Interacting with Real and Virtual Objects), qui soit adapté aux systèmes de Réalité Mixte et collaboratifs. Un deuxième objectif est de proposer une première version d'une méthodologie de conception permettant de supporter le développement de systèmes de Réalité Mixte et collaboratifs.

Notre démarche pour atteindre ces objectifs a été tout d'abord d'effectuer plusieurs états de l'art sur :

- Des applications existantes de Réalité Mixte,
- Des dispositifs d'interaction pour la Réalité Mixte,
- Des modèles d'interaction « classiques » (sans TCAO), et des modèles d'architecture des logiciels.
- Des modèles d'interaction avec TCAO, et des modèles d'architecture des collecticiels.

A partir de ces analyses, nous avons constaté qu'il n'existait (au début de nos travaux) qu'un seul modèle, ASUR, prenant en compte la Réalité Mixte, mais de façon incomplète. Pour bâtir notre propre modèle, IRVO, nous sommes donc partis de ce modèle existant.

Nous avons validé IRVO, du point de vue de sa capacité descriptive, en modélisant avec IRVO les applications existantes de Réalité Mixte étudiées dans l'état de l'art. Nous avons aussi examiné ses

capacités pour l'analyse prédictive de systèmes en cours de conception. Nous avons incorporé IRVO aux autres modèles d'une méthodologie de conception MBA (Model-Based Approach) en cours de développement dans notre laboratoire, la méthodologie CoCSys, dans le but de supporter le développement de systèmes de Réalité Mixte et collaboratifs.

### *Plan de la thèse*

Le **Chapitre II** vise tout d'abord à définir le concept de Réalité Mixte ce qui n'est pas aussi simple qu'il y paraît au premier abord car la définition a évolué de la Réalité Augmentée vers la Réalité Mixte. De plus, deux communautés de recherche travaillent sur ce thème, la communauté de la synthèse d'images (computer graphic) et la communauté IHM, qui ont chacune leur vision du domaine. Dans la suite du chapitre nous faisons un état de l'art des applications de Réalité Mixte que nous avons recensées dans la littérature. En fin de chapitre nous examinons plusieurs taxonomies qui ont été proposées dans la littérature et qui peuvent nous permettre de mieux cerner ce qu'est ou n'est pas un système de Réalité Mixte. Finalement, pour conclure ce chapitre, nous avançons « notre » définition de la Réalité Mixte.

Le **Chapitre III** examine les différents dispositifs d'entrées/sorties mis en œuvre dans le contexte de la Réalité augmentée et permettant plus particulièrement de réaliser les « augmentations » voulues. Une première section détaille les dispositifs en sortie : affichage des images (sens de la vue), restitution des sons (sens de l'ouïe) et les interfaces haptiques et kinesthésiques (sens du toucher). Une deuxième section aborde les dispositifs en entrée : capture des positions et des mouvements, capture de l'audio et suivi du regard (eye-tracking) et de manière anecdotique les interfaces contrôlés par le cerveau.

Le **Chapitre IV** passe en revue les différents modèles classiques de l'Interaction Homme-Machine conçus pour les interfaces du type WIMP (Windows, Icon, Menus and Pointing devices). Ils sont particulièrement nombreux et nous ne citons que les principaux. Ensuite, nous étudions différents modèles qui prennent en compte en partie les outils et les objets du monde réel comme moyen d'interaction. Nous examinons notamment le modèle ASUR qui est dédié à la Réalité Mixte. Nous étudions également dans ce chapitre les modèles d'architectures logicielles et une extension du modèle PAC-Amodeus pour prendre en compte les dispositifs d'entrée/sortie spécifiques à la Réalité Mixte. Nous concluons ce chapitre par une analyse critique de ces différents modèles en montrant les limitations que nous y voyons ce qui nous conduira à proposer notre propre modèle au Chapitre VI.

Le **Chapitre V** aborde le Travail Collaboratif Assisté par Ordinateur (TCAO). Nous commençons par une définition et un rappel des principaux concepts associés, avant de nous focaliser sur les aspects plus spécifiques de l'interaction homme-machine dans les collecticiels. Nous examinons plus particulièrement le formalisme de Renevier dédié à la modélisation des systèmes mixtes collaboratifs et mobiles. Nous étudions également dans ce chapitre les modèles d'architectures logicielles et une nouvelle extension du modèle PAC-Amodeus pour prendre en compte les aspects liés au contexte. Nous concluons ce chapitre par une analyse critique de ces différents modèles.

Le **Chapitre VI** présente la contribution principale de notre thèse, le modèle IRVO (Interacting with Real and Virtual Objects) comme réponse aux critiques des modèles classiques de l'IHM et du TCAO qui ne sont pas adaptés à la Réalité Mixte, car ils sont conçus dans l'optique d'une interface utilisateur de type WIMP et ignorent tous les aspects liés aux objets et outils réels. Après une présentation générale des objectifs et des fondements du modèle, nous détaillons le formalisme graphique d'IRVO et nous donnons plusieurs exemples montrant le pouvoir d'expression du modèle. Nous faisons une évaluation pragmatique d'IRVO en modélisant les 44 applications extraites de la revue de la littérature au Chapitre II et en les classant dans une taxonomie que nous proposons et qui étend les taxonomies existantes. Nous comparons également les capacités descriptives d'IRVO avec le modèle ASUR, d'une part, et le formalisme de Renevier, d'autre part.

Nous examinons ensuite les capacités d'analyse prédictive d'IRVO en proposant des règles de vérification basées essentiellement sur des propriétés ergonomiques. Nous finissons ce chapitre en proposant une évaluation théorique du modèle et en concluant sur les apports de ce modèle.

Le **Chapitre VII** aborde la conception des systèmes de Réalité Mixte. Nous faisons le choix d'une approche MBA (Model-Based Approach) que nous présentons de manière générale. Nous montrons en particulier qu'IRVO peut harmonieusement s'insérer parmi les autres modèles et nous étudions tout particulièrement les liens entre IRVO et l'arbre de tâches. Nous proposons ensuite une première version d'une méthodologie de conception des systèmes mixtes et collaboratifs comme une extension de la démarche CoCSys dans laquelle nous nous proposons de nous insérer. Nous montrons les apports d'IRVO dans cette démarche pour la conception des applications de Réalité Mixte. Nous étudions également les conséquences sur l'architecture du logiciel et établissons des liens entre le modèle IRVO et le modèle AMF-C qui modélise le contrôleur de dialogue de l'application dans l'architecture préconisée par CoCSys. Nous concluons sur la nécessité de poursuivre ces travaux pour proposer une méthodologie complète et utilisable.

Le **Chapitre VIII** présente le projet ACCORéA (Aide à la Conception Coopérative supportée par la Réalité Augmentée) que nous avons initié au laboratoire ICTT (Interaction Collaborative Téléactivités, Téléformation) en 2001. Il cherche à étudier comment la Réalité Mixte peut aider dans des activités de conception tant dans ses dimensions individuelles que collectives et que les acteurs soient dans un même lieu ou à distance. Ce projet nous sert de support à une première utilisation de la méthodologie présentée au chapitre précédent. Nous présentons les résultats de ce projet qui sont, un dispositif technique de Table Interactive que nous avons mis en place au laboratoire, et un prototype d'éditeur de texte collaboratif augmenté, WordTab.

Le **Chapitre IX** est la conclusion de ce mémoire dans laquelle nous faisons le bilan de nos contributions et où nous dégageons les perspectives de recherche.

# Chapitre II      La Réalité Mixte : un paradigme pour l'Interaction Homme-Machine

*Reality is that which, when you stop believing in it, doesn't go away.*  
Philip K. Dick (1928 – 1982).

Au début de ce chapitre, nous essayons tout d'abord de donner une définition de la Réalité Mixte. Cela n'est pas aussi simple qu'il y paraît au premier abord car chaque auteur a « sa » définition et la notion même de Réalité Mixte a évolué au fil du temps.

Dans la suite du chapitre nous passons en revue une quarantaine d'applications de la littérature et faisant partie du domaine de la Réalité Mixte. Il ne s'agit pas seulement de sacrifier à la tradition de « faire l'état de l'art » mais ce travail servira, d'une part pour examiner les dispositifs d'entrée/sortie utilisés dans ces applications (Chapitre III) et d'autre part comme base de validation de notre modèle (Chapitre VI).

En fin de chapitre nous examinons plusieurs taxonomies qui ont été proposées dans la littérature et qui peuvent nous permettre de mieux cerner ce qu'est ou n'est pas un système de Réalité Mixte.

Finalement, en guise de conclusion, nous énonçons « notre » définition de la Réalité Mixte.

## II.1 Définitions de la Réalité Mixte

Comme nous l'avons dit en introduction, l'acte fondateur de la Réalité Augmentée peut être situé par le dossier spécial de *Communications of the ACM* de juillet 1993, intitulé « Computer-Augmented Environments : Back to the Real World » et proposé par Pierre Wellner, Wendy Mackay et Rich Gold [Wellner 1993a]. Il semblerait donc aisé d'y trouver une définition précise et consensuelle de la Réalité Augmentée, mais malheureusement ce n'est pas vraiment le cas selon la provenance scientifique des chercheurs ou l'angle d'étude de leurs travaux.

### II.1.1 La réalité augmentée – premières définitions

La Réalité Augmentée (RA) est définie par Wendy Mackay comme une manière de réintégrer « l'information électronique dans le monde physique ». Il s'agit de « permettre aux gens de tirer parti de leurs compétences dans l'interaction avec le monde de tous les jours, tout en profitant de la puissance des réseaux informatiques » [Mackay 1996, Mackay 1998a]. Elle distingue 3 types d'augmentations :

- **Augmentation de l'utilisateur** : l'utilisateur porte un dispositif, généralement sur la tête ou les mains, pour obtenir des informations sur les objets physiques,
- **Augmentation des objets physiques** : l'objet physique est modifié en incorporant des périphériques d'entrée, de sortie ou des capacités informatiques sur ou en lui.
- **Augmentation de l'environnement des utilisateurs et des objets** : ni l'utilisateur, ni les objets sont directement affectés. A la place, des périphériques indépendants fournissent et collectent de l'information de l'environnement, en affichant l'information sur les objets et en capturant l'information sur l'interaction de l'utilisateur avec ceux-ci.

La RA est vue selon Robinett [Bowskill 1997] comme un moyen d'augmenter les sens de l'utilisateur, de transformer des événements imperceptibles en phénomènes visibles, audibles ou touchables. C'est vu comme une symbiose entre l'homme et la machine comme jamais auparavant.

De son côté, Milgram définit la RA comme l'ensemble des cas où un environnement réel est – *augmenté* – au moyen d'objets virtuels (graphiques générés par ordinateur)<sup>1</sup> [Milgram 1994a].

Pour Azuma [Azuma 2001] un système de RA complète le monde réel avec des objets virtuels (générés par ordinateur) qui semblent coexister dans le même espace que le monde réel. Un système de RA a les propriétés suivantes :

- il combine des objets réels et virtuels dans un environnement réel,
- il fonctionne de manière interactive, en temps réel,
- il fait coïncider les objets réels avec les objets virtuels.

Cette définition est orientée vers les systèmes de RA qui augmentent la vision mais, d'après l'auteur, la RA, peut potentiellement s'appliquer à tous les sens, incluant l'ouïe, le toucher, et même l'odorat.

Selon Thomas Baudel [Baudel 1995] :

La Réalité Augmentée « vise, selon nous, à "augmenter" les propriétés des objets de notre entourage de capacités de traitement d'information. En sus de leur fonction matérielle (ergotique), ils acquièrent une dimension informatique (sémiotique), par leur capacité de réagir non pas aux seuls phénomènes physiques auxquels ils sont soumis, mais aussi aux informations qu'ils captent sur l'état de leur entourage (personnes, environnement, autres objets "augmentés"...).

Une caractéristique essentielle en est le transfert du lieu de l'interaction: le système interactif n'est plus un lieu mais un environnement. L'interface n'est plus un face à face, mais se dissout dans l'espace et les objets de notre entourage. L'utilisation d'un système de traitement d'information n'est plus exclusivement un acte conscient et intentionnel. »

Cependant, il nous faut bien reconnaître que toutes ces définitions sont assez générales et peuvent s'appliquer tout aussi bien à l'informatique ubiquitaire (Ubiquitous Computing) et ne départagent pas vraiment les deux « concurrents ».

Comme nous le voyons, la Réalité Augmentée fait référence en permanence à des objets réels et virtuels. Avant d'aller plus avant, il s'agit donc d'essayer de définir ces notions, en particulier en reprenant les définitions qui ont été données dans le cadre de la Réalité Virtuelle.

## II.1.2 Objets réels et virtuels

### II.1.2.1 Définitions

Le Dictionnaire de arts médiatiques [GRAM 1996] définit la Réalité virtuelle<sup>2</sup> comme :

Technologie propre aux systèmes informatiques visant à créer des environnements qui simulent le monde naturel ou un monde imaginaire et qui donnent à l'utilisateur l'impression de la réalité

L'affichage peut se faire sur un simple écran complété par des haut-parleurs stéréophoniques. Lorsque l'utilisateur est totalement immergé dans le monde virtuel, c'est-à-dire qu'il est isolé du monde extérieur (utilisation d'un casque de visualisation stéréoscopique, d'un casque audio avec son spatialisé, gants numériques, etc.), on parle alors d'environnement virtuel immersif. Ceci a de nombreux avantages :

<sup>1</sup> "As an operational definition of Augmented Reality, we take the term to refer to any case in which an otherwise real environment is "augmented" by means of virtual (computer graphic) objects."

<sup>2</sup> L'expression Réalité Virtuelle est un oxymoron et de nombreux auteurs préfèrent parler plus justement d'environnements virtuels. Cependant, cette expression a eu un tel impact, voire un tel « succès commercial » qu'il faut désormais vivre avec cet abus de langage.

- L'immersion permet d'éviter la barrière produite par les interfaces et de faciliter la communication personne-machine. Cependant, les dispositifs actuels d'interaction sont encore perfectibles pour qu'ils soient réellement « transparents ».
- Les simulations permettent de réduire les temps et les coûts de développement par rapport à des maquettes physiques (simulations de voitures). De plus, les résultats des calculs peuvent être directement visualisés (écoulement de l'air autour du prototype).
- Il est possible de créer des environnements non encore réalisés (simulation de vols spatiaux) ou difficilement réalisables (simulation d'incendie pour l'entraînement des pompiers).
- Il est possible de recréer un monde virtuel ayant des lois physiques différentes du monde réel ce qui intéresse particulièrement le domaine des jeux.

Une **image virtuelle** est définie par les stimuli visuels, auditifs, tactiles et kinesthésiques qui sont transmis aux organes sensoriels de telle manière qu'ils apparaissent avoir pour origine l'espace en trois dimensions qui entoure les participants de l'environnement virtuel [Furness 1995].

Une **interface virtuelle** est un système de transducteurs, de processeurs de signaux, de périphériques informatiques et de logiciels qui crée un médium interactif par lequel :

- L'information est transmise aux sens sous la forme d'images virtuelles tridimensionnelles, de retour tactile et kinesthésique, de sons spatialisés.
- Le comportement psychomoteur et physiologique de l'utilisateur est détecté et utilisé pour manipuler l'environnement virtuel [Furness 1995].

Paul Milgram définit les objets réels et virtuels ainsi [Milgram 1994a] :

- Les **objets réels** sont tous les objets qui ont une existence concrète et objective. Pour être vu, un objet réel peut être observé directement ou être numérisé puis restitué par un périphérique de sortie.
- Les **objets virtuels** sont les objets qui existent par essence ou par effet mais non formellement ou concrètement. Pour être vu, un objet virtuel doit être simulé puisque par essence il n'existe pas.

De leur côté, Trevisan et al. [Trevisan 2003, Trevisan 2004] distinguent 3 catégories d'objets :

- Les **objets réels ou physiques** sont des objets qui ont une **existence concrète et objective** et qui peuvent être observés directement,
- Les **objets numériques virtuels** sont des objets qui existent par essence ou effet mais non formellement ou concrètement; pour être vus, ils doivent être simulés.
- Les **objets numériques réels** sont des objets réels perçus par des moyens numériques comme la vidéo d'une scène réelle.

Il faut préciser que pour les auteurs qui viennent de la communauté de synthèse d'images, le mot virtuel est implicitement associé à la Réalité Virtuelle, c-à-d que, lorsqu'ils parlent d'augmentation du réel par le virtuel, il s'agit implicitement d'objets virtuels visualisés de la manière la plus réaliste possible et qui fusionnent visuellement avec les objets réels (Cf. définition ci-dessus). Pour notre part **nous prendrons « virtuel » dans un sens le plus général possible, pour toute information provenant de l'ordinateur, mais dont la visualisation peut très bien être symbolique** : par exemple il peut s'agir d'une simple annotation textuelle superposée à l'objet réel comme dans le Musée Augmentée [Rekimoto 1995] (Cf. § II.2.8.1). Certains auteurs préfèrent parler de « numérique » plutôt que de « virtuel » pour éviter cette connotation trop forte avec la Réalité Virtuelle mais « numérique » ne traduit pas le fait que l'objet numérique est perceptible.

### II.1.2.2 La Réalité Virtuelle versus la Réalité Augmentée

La Réalité Augmentée est donc parfois définie comme prenant le contre-pied de cette approche immersive, et au lieu d'enfermer les personnes dans un monde artificiel, se propose d'utiliser l'ordinateur pour simplement augmenter le monde réel des utilisateurs [Wellner 1993a].

Cependant, de son côté, la communauté de synthèse d'images ne voit pas la RA comme une opposition à la RV mais comme une extension : il s'agit de réintroduire des éléments réels dans une scène virtuelle<sup>3</sup> [Azuma 1997]. De nos jours, on parle dans ce cas plutôt de Virtualité Augmentée, comme nous le verrons ci-dessous.

Emmanuel Dubois propose une synthèse des deux visions en proposant un triangle de la réalité augmentée qui complète le triangle de la réalité virtuelle de Burdea [Dubois 2001b] reproduit sur la Figure 1.

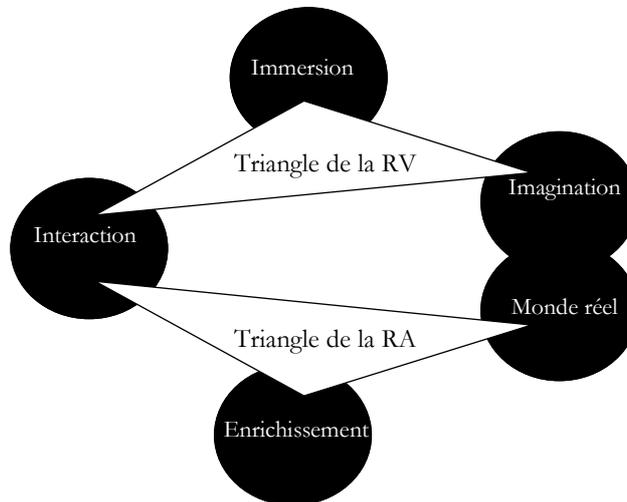


Figure 1 : Triangles de la Réalité Augmentée (RA) et de la Réalité Virtuelle (RV)  
(d'après [Dubois 2001])

### II.1.3 Continuum réel – virtuel

Comme nous le voyons, la définition de la RA n'est pas fixe et dépend des auteurs, de leur communauté scientifique et de leurs intentions. En particulier les chercheurs venant du monde de la synthèse d'images (Computer Graphics) ne voient pas d'opposition entre RA et RV mais plus une ouverture de la RV vers le monde réel.

Ainsi, Paul Milgram a proposé une unification des concepts en considérant un « continuum réel – virtuel » qui va du monde réel à des environnements entièrement virtuels tels que proposés par la RV [Milgram 1994a, Milgram 1994b, Milgram 1995]. Sur ce continuum, il considère les étapes intermédiaires sous l'appellation de Réalité Mixte (Mixed Reality) qui mélange des objets réels et virtuels. La Réalité Mixte (RM) se décline en 2 sous cas : la Réalité Augmentée (RA) et la Virtualité Augmentée (VA) selon la proportion d'objets réels ou virtuels (Cf. Figure 2).



Figure 2 : Le "Continuum réel – virtuel" (d'après [Milgram 1994a])

Azuma précise que dans la virtualité augmentée (dans laquelle les objets réels sont ajoutés aux objets virtuels) et les environnements virtuels (ou réalité virtuelle), l'environnement (*surrounding*

<sup>3</sup> "Augmented Reality (AR) is a variation of Virtual Environment (VE) or Virtual Reality as it is more commonly called" [Azuma 1997]

*environment*) est virtuel, tandis que dans la Réalité Augmentée, l'environnement est réel [Azuma 2001].

Dubois a une approche différente car il distingue la RA et la VA selon 2 continua distincts [Dubois 1999], en considérant **l'objet de la tâche** qui est soit réel (interaction homme-monde réel) ou virtuel (interaction homme-machine) :

- **Le continuum de la Réalité Augmentée** : l'interaction se passe dans le monde réel et elle est augmentée par l'ordinateur,
- **Le continuum de la Virtualité Augmentée** : l'interaction avec l'ordinateur est augmentée par des objets et des actions dans le monde réel.

#### II.1.4 De la Réalité Augmentée à la Réalité Mixte

Ainsi, nous avons pu voir que la Réalité Augmentée telle qu'elle est apparue pour la première fois en 1993 n'a pas (ou n'avait pas) le même sens selon les auteurs et englobait plus ou moins tout système mélangeant objets réels et virtuels. Il semble qu'aujourd'hui les définitions de Milgram et le terme de Réalité Mixte soient partagés par l'ensemble de la communauté et que le terme de Réalité Augmentée a le sens plus restrictif d'environnement réel augmenté par le virtuel. C'est en tout cas, la position que nous adopterons dans ce mémoire ; dans les cas où un auteur utiliserait le terme RA dans le sens plus général de RM, nous le préciserons explicitement pour éviter toute ambiguïté.

Philippe Renevier définit les **systèmes mixtes** comme tout système qui met en œuvre une ou plusieurs augmentations [Renevier 2004c]. Une **augmentation** est l'expression d'un lien entre le monde physique et le monde numérique, c'est-à-dire :

- Soit l'ajout d'une entité d'un des deux mondes dans l'autre monde
- Soit la manipulation d'une entité d'un des deux mondes par des outils de l'autre monde.

Il définit également [Renevier 2004c] :

- Un **système mixte mobile** est un système dont les augmentations dépendent de la localisation de l'utilisateur.
- Un **système mixte collaboratif** est un système dont les augmentations pour un utilisateur dépendent des actions et des connaissances des utilisateurs
- Un **système mixte collaboratif et mobile** est un système dont les augmentations pour un utilisateur dépendent à la fois de sa localisation, de son orientation et des activités des utilisateurs.

On remarquera que l'auteur s'est limité à la prise en compte de la localisation de l'utilisateur pour les systèmes mobiles, mais qu'il s'agirait plus globalement de prendre en compte le contexte physique de l'utilisateur (c'est-à-dire toutes les données physiques mesurées) [Renevier 2004c].

#### II.1.5 L'informatique ubiquitaire ou pervasive

L'informatique ubiquitaire (Ubiquitous Computing) a été proposée par Mark Weiser [Weiser 1991, Weiser 1993] comme une manière de réconcilier les mondes physiques et virtuels.

Son constat est que non seulement l'utilisation de l'informatique est entravée par le problème de l'interface mais que la notion même d'ordinateur « personnel » est dépassée. Il propose donc une vision où les ordinateurs disparaissent en s'intégrant aux objets et à l'environnement au point de ne plus être distinguables en tant que tels. Il s'agit donc d'avoir des ordinateurs petits, à faible consommation et peu coûteux qui incluent des capteurs et des afficheurs pour communiquer avec l'environnement et les utilisateurs et qui sont connectés en réseau.

La Réalité Mixte garde, *a priori*, comme nous le montrera la revue de la littérature à la section II.2, une approche assez centrée sur l'ordinateur qu'il soit fixe ou porté (PDA, *wearable computer*). Cependant, de notre point de vue, ce n'est pas une caractéristique de la Réalité Mixte qui cherche à

la fusion des mondes réels et virtuels en « se moquant » bien de quelle manière on aboutit à cette fusion : augmentation des objets eux-mêmes, augmentation de l'utilisateur (en portant un casque de visualisation par exemple), augmentation de l'environnement (par projection des informations sur les objets) (Cf. Taxonomie de Mackay, § II.3.1).

Là, où l'informatique ubiquitaire dépasse réellement la Réalité Mixte dans ses ambitions est sur son caractère omniprésent (elle est partout) et sur sa transparence : on n'a même plus conscience de l'existence d'ordinateurs et de programmes informatiques [Saha 2003].

### II.1.6 Les interfaces tangibles : une forme de Réalité Mixte

La notion d'interface tangible (TUI = Tangible User Interface) a été introduite par Ishii et Ullmer en 1997 [Ishii 1997]. Elle se base sur les travaux de Fitzmaurice, Ishii et Buxton sur les interfaces « empoignables » (*Graspable User Interface*). Il s'agit d'aller au-delà des interfaces graphiques traditionnelles (appelées WIMP=Windows, Icons, Menus and Pointing Devices<sup>4</sup>) en proposant d'utiliser des objets physiques pour interagir avec l'ordinateur :

Les interfaces tangibles vont augmenter le monde physique réel en couplant de l'information numérique aux objets et environnements physiques de tous les jours [Ishii 1997].

Les auteurs voient donc une évolution de l'informatisation de la métaphore du bureau :

- D'une part, vers notre peau et nos corps,
- D'autre part, dans l'intégration dans les environnements physiques dans lesquels nous vivons.

Cet objectif est très proche de celui de la Réalité Mixte mais en se focalisant clairement sur l'interaction alors que, comme nous l'avons vu, pour de nombreux auteurs, la Réalité Mixte se focalise plus sur la fusion des mondes réels et virtuels.

Dans la perspective des interfaces tangibles, les objets de la tâche sont toujours considérés être numériques, seuls les outils pour interagir avec ceux-ci sont augmentés par l'utilisation d'objets réels [Dubois 2003a]. Ainsi, les interfaces tangibles sont à considérer comme un cas particulier de la Virtualité Augmentée et donc comme faisant partie de la Réalité Mixte au sens large.

### II.1.7 Synthèse

Jun Rekimoto [Rekimoto 1995] a proposé une taxonomie pour distinguer quatre styles d'IHM (Figure 3) qui nous semble particulièrement pertinente car elle résume bien les différentes approches de manière synthétiques :

- avec une **interface graphique classique** : l'interaction entre l'utilisateur et l'ordinateur est totalement séparée de l'interaction entre l'utilisateur et le monde réel : il y a un gouffre.
- dans la **Réalité Virtuelle** : l'ordinateur entoure complètement l'utilisateur et l'interaction avec le monde réel disparaît.
- dans **l'informatique ubiquitaire** : l'utilisateur interagit avec le monde réel mais aussi avec des ordinateurs incorporés aux objets réels.
- dans la **Réalité Augmentée** : l'interaction de l'utilisateur avec le monde réel est augmenté par des informations venant de l'ordinateur.

Cependant nous ne pensons pas que les deux derniers cas soient vraiment disjoints et de nombreux systèmes, comme nous le verrons par la suite, peuvent être considérés comme mixtes et ubiquitaires.

---

<sup>4</sup> Ou encore Windows, Icons, Mice and Pulldown Menus.

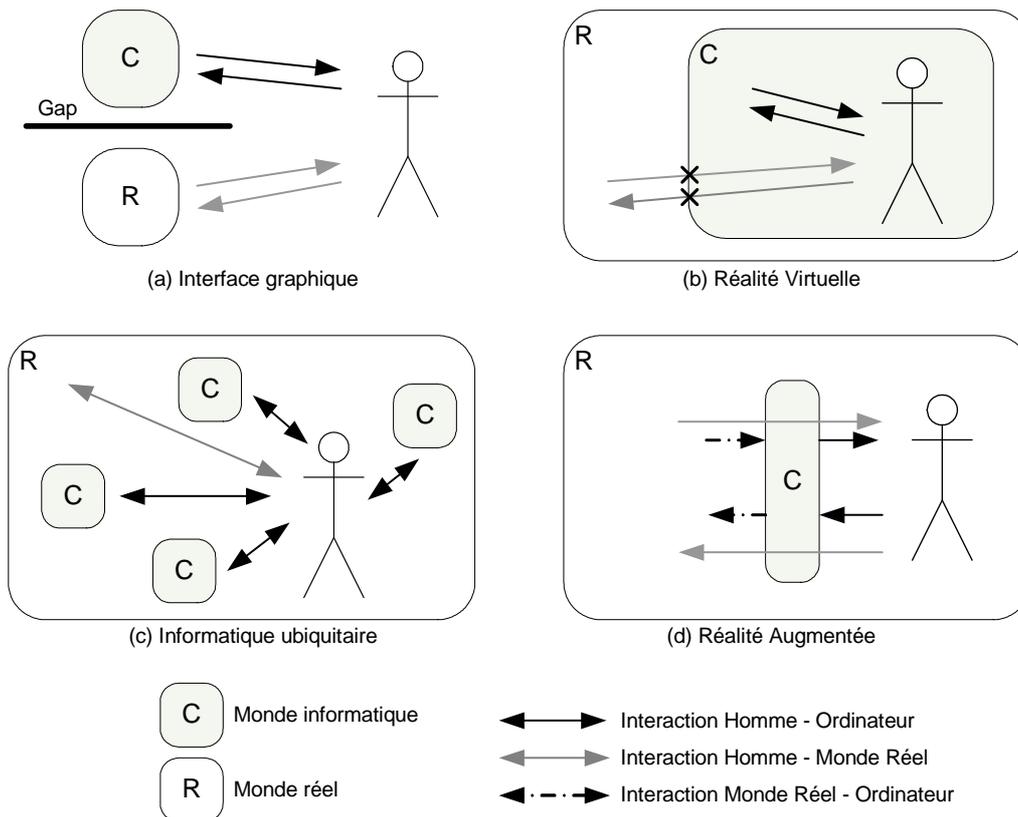


Figure 3 : Comparaison des styles d'interaction (d'après [Rekimoto 1995])

Il est donc difficile de trouver une définition précise et consensuelle de la Réalité Mixte, chaque auteur en fonction de sa communauté scientifique d'origine, de ses intérêts et de ses travaux, propose sa définition. Pour mieux cerner ce concept nous nous proposons donc de faire une revue des systèmes existants pour dégager des traits communs.

## II.2 Exemples de systèmes de Réalité Mixte

Dans cette section, nous proposons une revue de la littérature en examinant les différentes applications qui ont été proposées dans la précédente décennie et qui se réclament du concept de Réalité Mixte ou Augmentée. Nous ne prétendons pas à l'exhaustivité même si nous pensons avoir couvert une bonne partie des applications et en tout cas une palette représentative des domaines d'application et des technologies mises en œuvre.

Pour présenter ces applications, nous avons choisi de les classer par domaines d'application. Ce choix n'est pas forcément le plus pertinent, surtout que nos catégories ne sont pas totalement disjointes. Ce classement nous a cependant paru plus judicieux qu'un classement chronologique.

En conclusion, un tableau de synthèse permettra, en fonction de différents critères, de proposer d'autres classifications.

### II.2.1 Applications bureautiques

Nous avons regroupé dans cette catégorie des applications qui sont principalement dérivées d'applications classiques de bureautique tels les traitements de texte, les tableurs, les logiciels de présentation, les calculatrices, etc... sans être liées à un domaine d'application spécifique.

Les applications qui cherchent à augmenter les documents papiers sont nombreuses. En effet, dans notre société actuelle, l'écrit a une place prédominante dans tous les domaines d'activités. Les

applications bureautiques en général et le traitement de texte en particulier sont, de loin, les applications les plus utilisées par toute personne disposant d'un ordinateur.

### II.2.1.1 Charade [Baudel 1993]

Charade [Baudel 1993] permet le contrôle d'une présentation en amphithéâtre de diapositives ou d'une vidéo par des gestes de l'utilisateur. Pour cela, l'utilisateur porte un gant numérique qui permet de suivre la position de la main dans l'espace et de reconnaître les gestes. Un langage assez complet a été développé permettant en particulier de bien distinguer les gestes qui sont des commandes, des gestes naturels de l'utilisateur et donc d'éviter « l'effet Midas »<sup>5</sup>.

Cette application, présentée dans le numéro spécial des ACM [Wellner 1993a], est aussi citée par MacKay comme un exemple d'augmentation de l'utilisateur [MacKay 1998]. Cependant, l'objet de la tâche étant une projection de diapositives il est entièrement dans le monde virtuel et il n'y a pas de perception augmentée dans cette application. Mais le gant peut être vu comme un outil réel manipulé par l'utilisateur ce qui nous permet de rapprocher cette application des interfaces tangibles et donc de la classer dans le domaine de la Virtualité Augmentée.

### II.2.1.2 DigitalDesk [Wellner 1993b]

Le DigitalDesk [Wellner 1993b] est certainement l'un des systèmes de RA les plus connus. A la différence des systèmes d'interfaces graphiques comme Windows ou X/Motif qui sont basés sur la métaphore du bureau en cherchant à reproduire des actions (couper/coller) ou des objets (poubelle, etc..) sur l'écran de l'ordinateur, ce système est *littéralement* le bureau de l'utilisateur.

Il utilise donc le bureau réel comme le support principal de l'activité en l'augmentant par la projection d'informations provenant du système informatique.

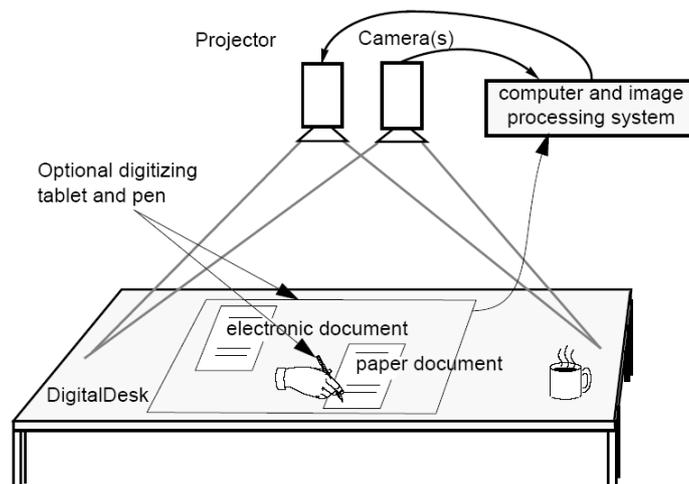


Figure 4 : Schéma du DigitalDesk (d'après [Wellner 1993b])

Le dispositif technologique est représenté sur la Figure 4 :

- Un vidéo-projecteur affiche des documents virtuels sur le bureau et peut aussi compléter des documents réels posés sur le bureau.
- Une caméra permet de suivre les actions de l'utilisateur (avec un stylo ou directement avec son doigt, d'où le terme « digital ») mais aussi de numériser les documents réels ou des annotations graphiques et textuelles.

- Un micro détecte les tapotements de l'utilisateur sur le bureau qui sont le substitut au clic de la souris.

Deux applications ont été expérimentées avec ce système :

- **Calculator** : Une calculatrice virtuelle est projetée sur le bureau. Il est bien sûr possible d'entrer des nombres de manière traditionnelle avec les touches, mais il est aussi possible de désigner (avec un stylo ou le doigt) des nombres se trouvant sur des feuilles de papier, le système se chargeant de les numériser « au vol » pour alimenter la calculatrice.
- **PaperPaint** : Cette application permet de réaliser des opérations de copier/coller avec des dessins réels. L'utilisateur peut sélectionner une zone rectangulaire sur une feuille ordinaire à l'aide d'un stylet. Lorsqu'il lève le stylet, une capture numérique est réalisée et projetée (après traitement numérique) sur la feuille. Cette copie peut ensuite être librement déplacée et même copiée de nouveau à d'autres endroits.

Il existe de très nombreux projets sur les tables et les tableaux augmentés. Nous les abordons plus en détail dans le § III.3.1.4.

### II.2.1.3 DoubleDigitalDesk [Wellner 1993b]

Le DoubleDigitalDesk permet le partage de documents réels entre plusieurs utilisateurs. Il permet à des personnes situées dans deux lieux séparés de « partager » leur bureau physique, chacun pouvant voir, éditer et écrire sur les documents papiers de l'autre.

Dans cette application, chaque DigitalDesk capture en continu l'image du bureau local et projette une image seuillée et mise à l'échelle de la vidéo du bureau distant. Le résultat est que chaque utilisateur voit ce qu'il y a sur chaque bureau. Quant un document est posé sur le premier bureau, son image est projetée sur le deuxième bureau et vice versa. Les projections sont numériquement mises à l'échelle et positionnées pour fournir la même vue aux deux utilisateurs, et le contraste est ajusté pour que l'effet de boucle vidéo (c'est-à-dire l'image de son propre bureau retransmise par la vidéo du bureau distant) soit minimisé. Chaque utilisateur peut écrire ou dessiner avec un stylo réel sur des documents papier ou électroniques et l'autre utilisateur verra ces tracés apparaître à l'endroit correspondant. Les mouvements de la main sont aussi transmis ce qui permet aux utilisateurs de pointer mutuellement des zones sur des documents. Comme la main de l'utilisateur distant masque ce qui est dessous comme sur un bureau ordinaire, ceci doit être réglé par un protocole social et le discours. Il existe pour cela un lien audio à travers une liaison téléphonique ou des combinés casque-micro.

### II.2.1.4 Chameleon [Fitzmaurice 1993]

Ce système est basé sur l'utilisation d'un PDA dont l'affichage dépend de sa position physique (à l'aide d'un localisateur 3D). On peut donc le voir comme une « loupe informationnelle » qui permet d'avoir des informations sur l'objet prêt duquel on le place. De ce point de vue il est très proche de la deuxième version du Navicam (Cf. § II.2.1.5) sauf que l'image du monde réel n'est pas reprise sur l'affichage du PDA ce qui réduit l'effet d'augmentation de la réalité. Plusieurs applications sont proposées :

- **Une carte active** : l'utilisateur approche le PDA d'une carte murale de grande dimension et peut obtenir des informations sur la ville que se trouve à proximité.
- **Une bibliothèque augmentée** : Le PDA permet d'avoir des informations sur les sections pendant le parcours dans la bibliothèque. Il permet aussi d'afficher par exemple le sommaire d'un livre en se plaçant à proximité sans avoir à l'ouvrir.

---

<sup>5</sup> En référence à ce roi de Phrygie (VIII<sup>e</sup> siècle av. J.-C.) à qui Dionysos avait donné le pouvoir de transformer en or tout ce qu'il touchait, y compris ses aliments, et qui failli donc mourir de faim. Cette référence est prêtée à Steve Freeman d'après Pierre Wellner [Wellner 1993b].

- **Un substitut de bureau portable** : une image panoramique en 360° du bureau de l'utilisateur est prise et stockée sur le PDA avant son départ. Pendant son déplacement, l'utilisateur peut consulter ce qui se trouve dans son bureau sans être physiquement présent. Une version améliorée pourrait bien sûr utiliser une vidéo en direct du bureau si un réseau haut débit est disponible.

Une utilisation particulière du dispositif est la navigation dans un espace informationnel de grande dimension (où les mouvements gauche/droite et haut/bas permettent une navigation 2D et les mouvements d'avant en arrière permettent de régler le niveau de zoom).

### II.2.1.5 Applications basées sur Navicam [Rekimoto 1995]

Le dispositif Navicam [Rekimoto 1995] consiste en un casque de visualisation semi-transparent (permettant l'affichage d'informations sur la scène réelle) porté par l'utilisateur et muni d'une caméra qui reconnaît les situations du monde réel en détectant des étiquettes portant un code-barres en couleur.

Une deuxième version du dispositif utilise un ordinateur de poche (PDA) à la place du casque. Dans ce cas l'image filmée par la caméra est affichée sur le PDA et mélangée aux informations virtuelles pour reproduire une vue augmentée similaire à celle du casque. Selon les auteurs, cette configuration est décrite par la métaphore de la loupe : une loupe optique classique agrandit l'image du monde réel alors que cette loupe l'élargit en terme d'informations.

Plusieurs applications ont été développées avec ce dispositif ; nous citons ici les deux liées à la bureautique, les autres seront présentées dans les paragraphes II.2.8.1 (Augmented Museum) et II.2.10.1 (Augmented Library) :

- **Active Paper Calendar** : Dans cette application l'utilisateur regarde un calendrier classique (muni d'un code-barres) accroché au mur ; à cette vue directe est superposée la vue virtuelle grâce au Navicam de son agenda personnel pour la journée en cours.
- **Active Door** : Des portes de bureaux<sup>6</sup> sont munies d'un code-barres et lorsqu'un visiteur (muni de Navicam) se présente devant cette porte, son casque peut afficher des informations sur l'occupant du bureau en cas d'absence : par exemple le visiteur peut savoir où se trouve la personne en ce moment et si elle va revenir bientôt. Il est aussi possible à l'occupant du bureau de laisser un message vidéo. Bien que l'application porte le nom de « porte active » en réalité la porte est parfaitement passive (seule une étiquette est ajoutée) et l'augmentation par le casque la fait se comporter comme une porte active.

### II.2.1.6 Paper-based audio notebook [Stifelman 1996]

L'Audio Notebook proposé par Lisa Stifelman [Stifelman 1996] permet à un utilisateur de capturer et d'accéder à un enregistrement audio d'une conférence ou d'un cours de manière synchronisée avec des annotations écrites sur du papier.

Ainsi, lorsque l'utilisateur relit ses notes et tourne les pages de son carnet, la lecture audio en cours s'interrompt et saute jusqu'au point correspondant au début de la page. La synchronisation de l'audio est également assurée par le pointage direct dans la page de notes. Enfin une barre de défilement latérale permet de naviguer directement dans l'enregistrement audio ce qui est pratique lorsque les notes sont éparpillées et ne permettent que des sauts à « gros grains » dans le fichier audio. Les pages servent ainsi d'index implicite au fichier sonore : le carnet papier se trouve ainsi augmenté par l'audio.

Le dispositif expérimental est présenté Figure 5. Un détecteur optique placé en haut à droite permet de savoir quelle est la page « active ». Le carnet est posé sur une tablette numérique qui permet de

<sup>6</sup> Il s'agit ici de bureaux au sens de pièce (*office*) et non au sens de meuble (*desk*).

détecter l'endroit où l'utilisateur prend des notes pour les synchroniser avec l'enregistrement audio. Des boutons permettent de contrôler l'enregistrement (démarrage, pause, arrêt).

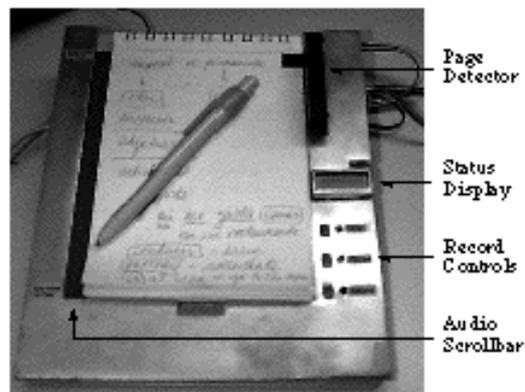


Figure 5 : Prototype de l'Audio Notebook (d'après [Stifelman 1996]).

### II.2.1.7 LivePaper [Robinson 2001]

Les dispositifs basés sur des tables interactives similaires au DigitalDesk pour la manipulation de documents réels sont nombreux et seront abordés au paragraphe III.3.1.4. Plus nombreuses encore sont les applications basées sur ces dispositifs et dans presque tous les domaines (Cf. paragraphes suivants).

Nous citerons une application collaborative dans le domaine de la bureautique assez proche fonctionnellement du DoubleDigitalDesk basé sur le système **LivePaper** [Robinson 2001]. Dans ce dispositif, seul un utilisateur dispose d'une table interactive, les autres utilisateurs distants disposent d'un ordinateur classique qui leur permet de visualiser l'image des objets réels posés sur la table et de dessiner des annotations virtuelles avec la souris ; ces annotations sont bien sûr projetées sur la table interactive avec un vidéo-projecteur pour que l'utilisateur local puisse les voir.

Une autre application mono-utilisateur du système LivePaper permet de contrôler des fichiers musicaux associés à des documents réels en fournissant des boutons de commandes virtuels (lecture, pause, arrêt, rembobinage, avance rapide) projetés à côté du document.

Une troisième application liée à l'architecture, permet la visualisation des murs en 3D à partir d'un plan papier 2D posé sur le bureau.

## II.2.2 Applications pour la conception

Nous avons regroupé dans ce paragraphe des applications dont le but principal est l'aide à la conception de produits dans le domaine industriel.

### II.2.2.1 Video Mosaic [MacKay 1993, Mackay 1998a]

L'application Video Mosaic est destinée aux producteurs de films et de vidéos pour gérer des story-boards papier. Un story-board est constitué de plans, chacun contenant une image représentative de la prise de vue, le dialogue des acteurs (ou une voix off) et des notes sur le plan ou la scène.

Video Mosaic fait le lien entre les story-boards papier et un système de montage vidéo informatisé.

Le système est composé : d'une caméra à fort zoom (ou d'un scanner à main) pour capturer les détails, d'une caméra grand angle pour détecter la position des éléments du story-board sur la table et d'un vidéo-projecteur pour afficher des informations pour l'utilisateur.

### II.2.2.2 Build-IT [Fjeld 1999, Fjeld 2002a]

Build-IT est un système permettant à un groupe de coopérer en vue de la conception et l'aménagement d'objets du monde réel comme des pièces, des écoles ou des usines. Le dispositif est similaire au DigitalDesk sur lequel un plan 2D de l'espace en cours d'aménagement est projeté. Les utilisateurs interagissent avec le système à l'aide de briques physiques leur permettant de sélectionner et manipuler les objets virtuels qui sont projetés sur la table. Plusieurs briques peuvent être utilisées simultanément ce qui permet de travailler à plusieurs en même temps mais aussi d'avoir des interactions à 2 mains permettant des actions complexes. Enfin, un deuxième vidéo projecteur (ou un écran) affiche une vue en 3D de l'environnement en cours de conception.

La Figure 6 montre des utilisateurs travaillant à plusieurs autour d'une table. La Figure 7 montre en détail une interaction où deux objets sont rapprochés l'un de l'autre en utilisant deux briques physiques.



Figure 6 : Vue générale du dispositif Build-IT (d'après [Fjeld 2002]).

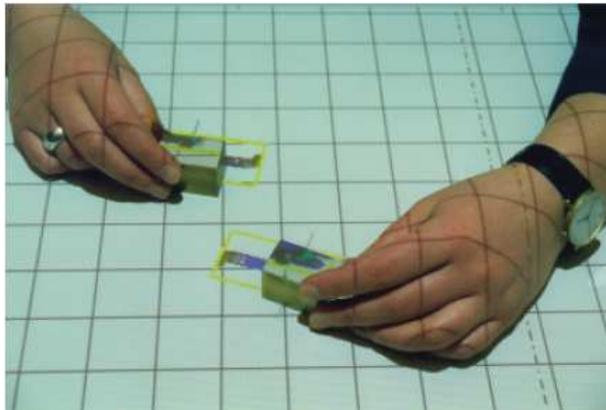


Figure 7 : Vue détaillée d'une interaction avec deux briques (d'après [Fjeld 2002]).

Le système Build-IT est à ranger du côté des interfaces tangibles tel que les auteurs le présentent. Selon la distinction abordée ci-avant, ce système est typiquement à classer dans la Virtualité Augmentée dont les interfaces tangibles sont une sous partie.

### II.2.2.3 SeamlessDesign [Kiyokawa 1999]

SeamlessDesign propose un environnement de travail collaboratif pour le prototypage rapide d'objets 3D ayant des contraintes géométriques. Les deux utilisateurs portent des casques de visualisation semi transparent muni de localisateurs (pour le suivi des mouvements de la tête) ainsi que des gants numériques pour interagir (également localisés par le système). Cette application est de type Réalité Virtuelle car l'ensemble des objets manipulés est virtuel. Cependant les auteurs présentent leur dispositif comme étant un *environnement augmenté partagé* (SAE – Shared Augmented

Environnement) car les deux utilisateurs peuvent se voir en train d'agir et aussi voir la direction de leur regard ce qui est important dans le contexte du travail collaboratif. On peut donc finalement la voir comme une forme dégradée de Virtualité Augmentée selon la définition du § II.1.3, car il s'agit plutôt d'une juxtaposition des objets virtuels et des utilisateurs réels que d'une fusion des deux mondes.

#### **II.2.2.4 Réalité Mixte pour les premières étapes de la conception de produit [Santos 2003]**

Cette application mono-utilisateur cherche à proposer un système générique pour les premières étapes de la conception de produits industriels et notamment pour la conception de voitures. Le dispositif technique utilise un casque de visualisation avec incrustation vidéo, un stylo et une tablette pour l'interaction et un localisateur basé sur deux caméras pour le suivi du casque, du stylo et de la tablette.

Les auteurs ont particulièrement étudié l'interaction à deux mains et plus spécifiquement un mode d'interaction avec un « ruban adhésif ». Dans la conception des automobiles, les concepteurs utilisent du ruban adhésif pour concevoir les lignes caractéristiques sur des prototypes : la main gauche maintient la bande adhésive pendant que la main droite détermine l'orientation de la courbe tangente au point d'adhésion. Cette technique a été transposée avec un ruban virtuel au dispositif des auteurs qui utilise une maquette réelle à échelle réduite d'une voiture en cours de conception.

### **II.2.3 Applications pour l'architecture et l'urbanisme**

#### **II.2.3.1 Ariel [MacKay 1998a]**

Cette application est destinée à assister des responsables de chantier qui manipulent des plans d'architecture sur une table à dessin. Une tablette tactile au format A0 permet de capturer les annotations et les commandes de l'utilisateur. Un vidéo-projecteur affiche des images informatiques et des menus interactifs sur la table.

Les plans sont identifiés par des code-barres lus par un lecteur manipulé par l'utilisateur. Le positionnement précis du plan sur la table est réalisé par l'utilisateur par pointage des 4 coins.

L'utilisateur peut faire des annotations sous forme d'hyperliens indiqués comme des « hotspots ». Il peut s'agir de notes textuelles, mais aussi audio ou vidéo.

Pour éviter que les informations projetées ne se mêlent au dessin, ce qui peut-être perturbant, une fenêtre réelle (une simple feuille de papier blanc) est déplaçable par l'utilisateur et est suivie par le système à l'aide d'une diode électroluminescente placée dans le coin.

Il est également possible de travailler à plusieurs et de partager des annotations grâce à un « media space ».

#### **II.2.3.2 Collaborative Interior Design [Ahlers 1995, Klinker 1997]**

Interior Design est une application collaborative développée pour l'aménagement intérieur de pièces. Les utilisateurs peuvent déplacer des meubles virtuels superposés à une vue réelle d'une pièce (éventuellement partiellement meublée par des meubles réels). Seule la vue est augmentée, l'interaction se faisant classiquement avec la souris dans une application sous X-Window.

D'un point de vue technique, la pièce est filmée par une caméra vidéo et le signal est mélangé à l'image des objets virtuels générée par le système informatique avant d'être affichée sur un moniteur vidéo. Ceci a l'avantage de pouvoir « distribuer » la réalité à tous les utilisateurs et de leur fournir la même vision que l'utilisateur soit localement dans la pièce ou à distance. Cependant, seul l'utilisateur local peut bouger la caméra (qui est suivie par un localisateur) pour changer l'angle de vue.

### II.2.3.3 Architectural Anatomy [Webster 1996]

Cette application utilise la Réalité Augmentée pour montrer à l'utilisateur des parties de l'immeuble qui sont cachées derrière l'architecture ou des structures de finition (comme des faux plafonds). L'utilisateur peut aussi obtenir des informations supplémentaires sur ces objets cachés.

Le dispositif utilise un casque de visualisation semi-transparent qui est suivi par un localisateur pour pouvoir générer une image virtuelle alignée sur la vue réelle. La visualisation se fait sous la forme de « fils de fer » et l'utilisateur peut utiliser une souris pour pointer ces éléments et obtenir des informations additionnelles.

### II.2.3.4 Augmented Urban Planning Workbench [Ishii 2002]

Cette application aborde la conception dans le domaine de l'urbanisme. Il s'agit de réconcilier les trois représentations que les concepteurs en urbanisme utilisent :

- Des plans 2D sur des feuilles de papier, des schémas dessinés à la main, des cartes, des photographies par satellite
- Des maquettes en trois dimensions. Ce sont des modèles physiques à échelle réduite des immeubles et des paysages
- Des modèles numériques en deux ou trois dimensions stockés dans l'ordinateur. Ce sont des modèles d'immeubles et des paysages, des modèles de simulation de l'éclairage, du vent et des flots de trafic.

Le dispositif technique est baptisé Luminous Table. Il s'agit d'un vaste plan de travail sur lequel des plans papiers en 2D ou des photos par satellite peuvent être posées. Sur ces documents, des maquettes en 3D des immeubles peuvent être posées. Enfin, deux vidéoprojecteurs suspendus au-dessus de la table projettent des simulations numériques dynamiques sur l'ensemble. Le dispositif est complété par deux caméras fixes qui captent l'activité sur la table et permettent d'aligner les projections dynamiques avec les dessins et les maquettes à l'aide d'étiquettes optiques.

Le logiciel propose plusieurs fonctionnalités. En particulier il permet de simuler l'éclairage par le soleil et les ombres portées par les immeubles et la simulation du trafic automobile.

L'interaction se fait en déplaçant des blocs transparents en plexiglas qui représentent les immeubles. Ce dispositif est donc proche des interfaces tangibles. Plusieurs personnes peuvent prendre place autour de la table pour des séances de conception collaborative.

### II.2.3.5 Collaborative City-Planning System [Kato 2003]

Cette application collaborative est très similaire à l'Augmented Urban Planning Workbench (Cf. ci-dessus) dans ses principes et son mode d'interaction. Par contre, d'un point de vue technique, chaque utilisateur porte un casque de visualisation muni de deux caméras en face de chaque œil : sur l'image filmée par chaque caméra sont incrustées les objets virtuels, avant d'être affichée dans le casque. L'image des caméras est aussi analysée par le système pour reconnaître des étiquettes permettant de calculer la position de l'utilisateur par rapport à la table de travail et de connaître la position des différents objets manipulés par les utilisateurs.

L'interaction, du type interface tangible, est basée sur une métaphore de *tasse*. Un cube réel en plexiglas est manipulé par l'utilisateur :

- Pour prendre un objet virtuel il suffit de le recouvrir par la *tasse* puis de la soulever de la table. L'objet est alors lié à la *tasse*.
- Pour poser un objet il suffit de poser la *tasse* à l'endroit voulu et de la soulever pour que l'objet reste en place.
- Pour déplacer un objet, il faut poser la *tasse* sur l'objet virtuel, déplacer la *tasse* sur la table pour emmener l'objet à l'endroit voulu, puis retirer la *tasse*

- Pour détruire un objet virtuel contenu dans la *tasse*, il suffit de la secouer et l'objet virtuel disparaît.

### II.2.3.6 Bird's-eye View System [Kim 1999]

Ce système présente l'originalité d'utiliser un hélicoptère miniature qui embarque une caméra et un système de localisation par GPS pour obtenir des vues de l'environnement réel prises d'en haut comme si l'utilisateur volait. L'image, retransmise au sol par une transmission hertzienne, est mélangée avec des images virtuelles générées par un système de Réalité Virtuelle pour produire une vue augmentée qui est affichée sur un écran classique. Une application expérimentale liée à l'architecture et l'urbanisme a été développée par l'équipe qui permet de visualiser un immeuble non encore construit au milieu des autres bâtiments et infrastructures existants.

## II.2.4 Applications pour la maintenance

### II.2.4.1 Karma [Feiner 1993]

Karma (acronyme de « Knowledge-based Augmented Reality for Maintenance Assistance ») permet l'accès à une base de données de gestes techniques pour aider des techniciens qui font la maintenance de dispositifs. Plutôt que de présenter l'information sur un écran d'ordinateur de manière textuelle et graphique, elle est directement présentée sous forme d'animations virtuelles superposées à la scène réelle.

Le prototype montre le cas de la maintenance d'une imprimante laser, en particulier deux tâches, le remplissage du bac à papier et le changement de la cartouche de toner. Le technicien porte un casque de visualisation semi-transparent qui permet de superposer les graphiques générés par l'ordinateur à la scène réelle (Figure 8).

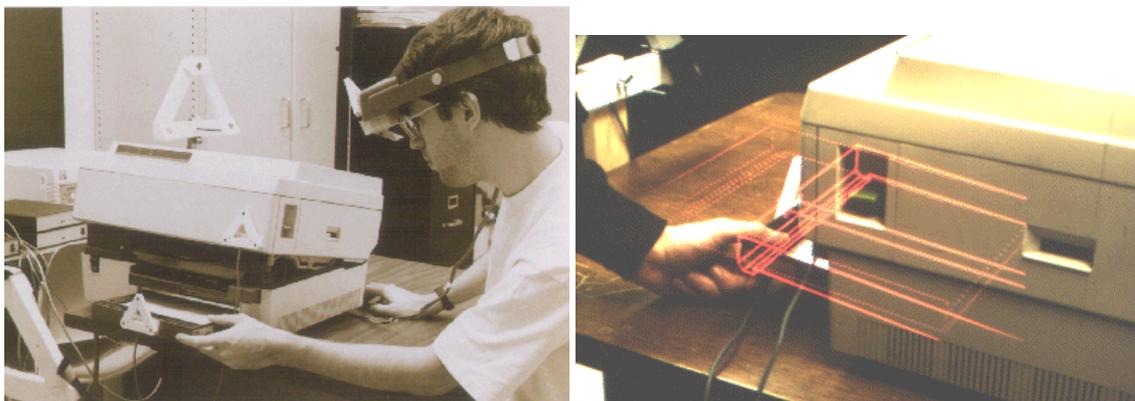


Figure 8 : Karma (d'après [Feiner 1993])

Un dispositif de suivi (localisateur ultrasonique) permet d'obtenir en temps-réel la position et l'orientation en 3D du casque ainsi que des principaux constituants de l'imprimante ( tiroir du papier, couvercle). Le suivi permet de réaliser deux fonctions :

- Calculer l'image 3D virtuelle de manière à la superposer (aux précisions de mesure près) à la scène réelle en permanence.
- Connaître la position des constituants de l'imprimante, ce qui permet de savoir si les actions demandées ont été accomplies par l'utilisateur.

En pratique, le dispositif présente une action à effectuer, comme par exemple : tirer le bac à papier, en montrant la position initiale (en lignes pleines) et finale du bac (en lignes pointillées), ainsi qu'une flèche montrant le mouvement à effectuer. Comme la position du bac à papier est suivie par le système, dès que l'utilisateur a fait le geste demandé, le système est averti de la fin de la tâche, ce qui permet de passer à la tâche suivante.

#### II.2.4.2 SEAR [Goose 2003a, Goose 2003b]

Le système SEAR (Speech-Enabled Augmented Reality), également appelé PARIS [Goose 2003b] est un système générique de RA mobile utilisant un système de localisation basé sur la vidéo et offrant une interface multimodale graphique et vocale.

L'application proposée permet à des techniciens de maintenance dans un environnement industriel de visualiser sur un ordinateur portable léger des données techniques sur les composants du système technique qu'il doit maintenir (Figure 9).



Figure 9 : Système SEAR (d'après [Goose 2003a])

Les composants sont identifiés et localisés par une caméra portée par l'utilisateur (et à l'aide d'étiquettes spéciales collées sur les composants) ce qui permet d'afficher sur l'écran de l'ordinateur une vue 3D virtuelle contextualisée du dispositif technique. L'interaction se fait principalement par un système de reconnaissance vocale (dont le vocabulaire dépend lui aussi du contexte) dans le but de libérer les mains de l'utilisateur qui doit en général effectuer des gestes techniques sur le dispositif industriel qu'il maintient. Enfin, en sortie, un synthétiseur de parole permet de générer les réponses du système.

#### II.2.4.3 Collaborative Mechanical Repair [Klinker 1997]

Dans cette application, un mécanicien est assisté par un système de RA pendant qu'il examine et répare un moteur complexe. Le système présente au mécanicien un ensemble d'informations superposées au moteur réel : il peut s'agir du nom des pièces, de la description de leur fonction ou encore l'historique de la fabrication et de la maintenance. Le mécanicien peut aussi bénéficier de l'assistance d'un expert distant qui peut contrôler quelle information est affichée sur le système de RA du mécanicien.

Techniquement, le système utilise le même dispositif que Collaborative Interior Design (§ II.2.3.2). L'augmentation est donc réalisée par superposition des informations virtuelles à une image vidéo du moteur, le tout étant affiché sur un moniteur vidéo.

### II.2.5 Applications pour l'enseignement, la formation et l'apprentissage

#### II.2.5.1 SharedView et GestureCam [Kuzuoka 1994]

Le système SharedView permet à un opérateur de piloter une machine complexe à commande numérique en recevant l'aide d'un formateur distant au moyen d'informations vocales et de données affichées dans un casque de visualisation semi-transparent. Une caméra est attachée au casque de visualisation qui filme ainsi à peu près le champ de vision du technicien. Cette image est affichée sur un écran et l'expert pointe directement son doigt sur l'écran pour désigner des objets.

Une deuxième caméra fixe filme le tout (écran + doigt) et envoie cette image dans le casque de visualisation du technicien qui voit ainsi l'image du doigt de l'expert en superposition avec l'image réel du système sur lequel il travaille.

Ce système a montré ses limites car il se peut que l'opérateur bouge la tête au moment où le formateur veut montrer quelque chose sur l'écran. De plus, si le formateur veut montrer une partie qui est hors du champ de vision, il doit demander à l'opérateur de tourner sa tête en transmettant des instructions orales.

Le système GestureCam est une version améliorée où la caméra de l'opérateur, au lieu d'être mobile avec lui, est posée à côté du dispositif et est pilotée à distance par le formateur. Sur cette image est superposé un pointeur piloté par le formateur et l'image composite est affichée devant le formateur et l'opérateur. De plus, l'orientation de la caméra montre directement à l'opérateur quelle partie le formateur examine, et optionnellement un pointeur laser solidaire de la caméra permet de pointer encore plus précisément un élément.

### II.2.5.2 Training Tool for Assembly Tasks [Boud 1999]

Cette application vise à fournir un support pour l'apprentissage de tâches d'assemblage de dispositifs mécaniques. Elle utilise un casque d'affichage semi-transparent (noir et blanc et monoculaire) pour afficher le plan de montage en superposition à la tâche réelle. Des expérimentations ont été conduites avec 20 sujets sur l'assemblage d'une pompe à 8 composants en comparant diverses techniques d'apprentissage de la tâche : avec un schéma sur papier, avec un système de Réalité Virtuelle (3 interfaces : moniteur et souris 2D, moniteur avec vision stéréoscopique et souris 2D et interface immersive et souris 3D) et avec le système de Réalité Augmentée décrit ci-dessus. Les résultats montrent que les meilleurs temps de montage de la pompe après apprentissage sont faits par les personnes qui ont été entraînées avec le système de RA, les plus mauvais temps étant ceux qui n'ont eu que le schéma papier.

### II.2.5.3 MagicBook [Billingshurst 2001a, Billingshurst 2001b]

Le projet MagicBook se propose d'augmenter des livres ordinaires qui peuvent être feuilletés et lus comme tout livre (y compris les images) sans technologie additionnelle. Cependant, les utilisateurs munis d'un casque de visualisation semi-transparent peuvent voir apparaître des modèles virtuels 3D posés sur les pages. Les contenus virtuels peuvent être de taille quelconque et animés. S'ils le souhaitent les utilisateurs peuvent complètement s'immerger dans la scène virtuelle comme dans un système de Réalité Virtuelle (Figure 10).



Figure 10 : les 3 modes d'utilisation du MagicBook : (a) réalité ; (b) Réalité Augmentée ; (c) immersion en Réalité Virtuelle (d'après [Billingshurst 2001b])

Techniquement, le dispositif de visualisation est composé d'un casque de vision opaque monté au bout d'un manche. Une caméra placée au centre du casque filme la scène réelle et l'affiche dans le casque augmenté par les objets virtuels incrustés (technique de *video see-through HMD*, Cf. § III.2.1.1). Des marqueurs spécifiques sont imprimés sur les pages et sont reconnus par la caméra, ce

qui permet d'assurer le positionnement correct des objets virtuels sur la scène réelle (utilisation de l'ARToolkit des auteurs, Cf. § III.3.1.1).

Sur le manche se trouvent également un localisateur de direction inertiel, un interrupteur et un capteur de pression. Lorsque l'utilisateur veut passer en mode de Réalité Virtuelle immersive, il appuie sur l'interrupteur : l'affichage bascule alors en mode RV pure (il n'y a plus transmission de l'image réelle) ; par une pression plus ou moins forte sur le capteur de pression, l'utilisateur peut choisir sa vitesse de vol dans le monde virtuel ; pour changer de direction dans le monde virtuel, il lui suffit de physiquement changer sa direction ce qui est capté par le localisateur de direction.

Ce dispositif a été utilisé pour de nombreuses applications en particulier en formation pour des enfants japonais ou encore pour apprendre à construire pas à pas la célèbre *Red and Blue Chair* (1918) de Gerrit Rietveld (1888-1965). Des applications pour la visualisation de données scientifiques (données sismiques) ont également été explorées.

Un mode collaboratif est également possible, chaque utilisateur tenant un dispositif d'affichage dans la main. Chaque utilisateur voit la scène virtuelle de son point de vue et peut désigner des objets virtuels avec son doigt à l'autre utilisateur. Si les utilisateurs passent en mode RV immersif, alors ils sont représentés sous forme d'avatar dans la scène virtuelle ce qui leur permet de se voir mutuellement. Un cas intéressant est lorsqu'un des utilisateurs est en mode RV immersif alors que l'autre est en mode RA : dans ce cas, le premier voit les lunettes de l'autre utilisateur flotter « dans le ciel » ce qui lui permet d'avoir un retour sur sa position même si le deuxième utilisateur n'est pas immergé dans la scène virtuelle.

#### II.2.5.4 Augmented Chemistry [Fjeld 2002b]

Cette application permet d'apprendre la chimie en utilisant une paillasse augmentée. La paillasse est constituée d'une table et d'un écran à projection arrière (Figure 11). Sous l'écran est posée une caméra qui permet d'enregistrer les interactions de l'utilisateur qui ont lieu sur et au-dessus de la table. L'écran affiche ce que la caméra filme de telle manière que l'utilisateur a l'impression de se trouver devant un miroir. Ce miroir est augmenté par les objets virtuels qui sont superposés à des cubes manipulés par l'utilisateur. Les cubes sont localisés par le système à l'aide de la caméra et de marqueurs spéciaux collés (utilisant l'ARToolkit).

Pour interagir, l'utilisateur dispose d'un calepin dans lequel se trouve les éléments chimiques (un par page). Il peut les prendre avec un outil (appelé *gripper*) et les amener sur la paillasse. Pour connecter les éléments chimiques entre eux et construire des molécules, il dispose d'un cube qu'il fait tourner avec sa deuxième main pour dire s'il veut une liaison chimique simple, double ou triple entre les éléments.



Figure 11 : La paillasse augmentée de Augmented Chemistry (d'après [Fjeld 2002])

Une « augmentation » audio est également fournie :

- Lorsqu'une molécule est saturée (c-à-d qu'il n'est plus possible d'y lier d'autres atomes), un signal acoustique l'indique à l'utilisateur qui peut commencer à construire la molécule suivante.
- Lorsqu'une molécule fait partie d'une liste prédéfinie, il est offert à l'utilisateur des informations parlées additionnelles sur celle-ci.

### **II.2.5.5 Caretta [Hosoi 2003]**

Le système Caretta est destiné à l'apprentissage collaboratif en face-à-face en intégrant des espaces privés et partagés. Chaque utilisateur dispose d'un PDA qui constitue son espace privé pour des activités individuelles. L'espace partagé est constitué par une table tactile multi-utilisateur sur laquelle des objets peuvent être posés et détectés (nature et position) à l'aide d'étiquettes RFID. Un vidéo-projecteur placé au-dessus de la table permet de projeter des informations et ainsi augmenter les objets posés.

L'application développée sur ce système permet l'apprentissage de l'urbanisme et des questions environnementales. Sur ce point il rejoint des systèmes tels que Augmented Urban Planning Workbench (§ II.2.3.4) ou Collaborative City-Planning System (§ II.2.3.5). Il comporte un module de support intelligent qui permet de donner des suggestions à l'utilisateur en se basant sur ses manipulations précédentes. Également en mode partagé, Caretta infère à partir des intentions de chaque utilisateur et génère différentes alternatives acceptables par le plus d'utilisateurs possible.

## **II.2.6 Applications dans le domaine médical**

### **II.2.6.1 CASPER v2 [Dubois 2001a, Dubois 2001b, Dubois 2003]**

Le système CASPER (Computer Assisted PERicardial puncture) est conçu pour aider un chirurgien à réaliser la ponction d'un liquide pathologique situé à proximité immédiate du cœur sans avoir à faire une intervention lourde : pour cela le chirurgien introduit une longue aiguille de ponction à travers la peau et les muscles jusqu'à l'épanchement qu'il retire ensuite par aspiration avec une seringue. On se doute que le risque est très élevé d'endommager d'autres organes (cœur, foie, poumon) avec des conséquences létales. Le système CASPER réalise donc une assistance pour cette tâche :

- Dans un premier temps, une série d'images échographiques du patient permet de planifier la trajectoire idéale de l'aiguille. La difficulté est que le cœur est mobile (par ses battements et la respiration pulmonaire) et que les tissus voisins sont déformables sous l'action d'introduction de l'aiguille.
- Dans un deuxième temps, pendant l'opération, la position de l'aiguille est localisée en temps réel pour aider le chirurgien à l'introduire en suivant au plus près la trajectoire idéale.

Une première version utilisait un affichage de croix de couleurs sur un écran ordinaire ce qui pose des problèmes d'interprétation de la direction de l'aiguille et surtout du sens de la correction à apporter pour rectifier la trajectoire. De plus, l'affichage étant sur un écran séparé, le regard du chirurgien fait continuellement des allers-retours entre le patient et l'écran.

La deuxième version, CASPER v2, utilise un affichage de réalité augmentée sur un casque de vision semi-transparent, ce qui permet une superposition directe de la trajectoire, représentée sous forme d'un cône de sécurité, sur le corps du patient. Le référentiel d'affichage étant celui du chirurgien, les erreurs d'interprétation de référentiel qui pouvaient exister avec l'affichage sur écran sont donc supprimées. La Figure 12 montre la vision à travers le casque en superposition sur un modèle physique représentant la cible.

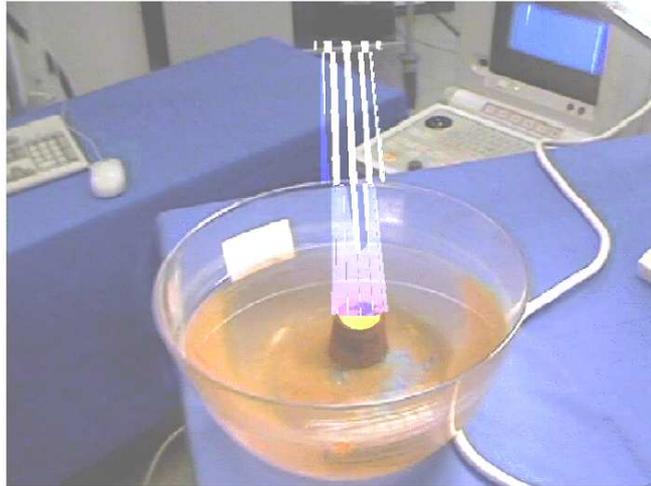


Figure 12 : Prototype du système CASPER v2 (d'après [Dubois 2001b]).

### II.2.6.2 Virtual Mirror Box [O'Neill 2003]

Cette application cherche à traiter le phénomène de membre fantôme dont souffrent de nombreux amputés parfois de manière sévère. Le principe de traitement est d'utiliser une boîte dans laquelle les patients mettent le bras intact et qui comporte un miroir placé sur la ligne médiane de telle sorte que la réflexion de leur bras donne l'impression aux patients qu'ils ont deux bras intacts.

Le dispositif Virtual Mirror Box reproduit ce principe mais en utilisant un écran plat à la place du miroir qui affiche l'image 3D d'un bras virtuel. Le bras intact est équipé d'un gant numérique qui pilote le modèle du bras virtuel ce qui permet de reproduire l'effet miroir du dispositif classique. L'avantage est que le bras virtuel peut être déformé de manière à représenter les pathologies ressenties par le patient : bras plus court ou plus long, variation en épaisseur, être continu ou avoir des lacunes, par rapport au bras originel. On peut aussi configurer le dispositif pour que les mouvements du bras virtuel soient différents de ce que permet un miroir comme suivre la même direction que le bras intact. Les auteurs espèrent ainsi mieux traiter cette pathologie avec les patients pour lesquels la boîte miroir classique ne convient pas.

### II.2.6.3 Autres applications

Emmanuel Dubois a fait une très bonne revue des applications dans le domaine des Gestes Médicaux Assistés par Ordinateur (GMCAO) [Dubois 2001b, chap. 1]. Beaucoup de ces applications utilisent essentiellement des retours d'informations visuelles dans le but de « montrer l'invisible ». Rare sont ceux qui exploitent la Réalité Augmentée comme le système de **craniotomie augmentée** (Grimson 1996, cité par [Dubois 2001b]). Quelques dispositifs offrent un retour haptique et un seul dispositif propose un retour sonore. Quelques dispositifs peuvent même aider le chirurgien dans son geste, voire réaliser complètement le travail (sous la supervision du chirurgien tout de même !) comme **Robodoc** (Mittelstadt 1993, cité par [Dubois 2001b]) ou **Minerva** (Galuser 1995, cité par [Dubois 2001b]).

Une autre application liée à la gestion des dossiers médicaux, **NISChart**, sera examinée au paragraphe II.2.10.3.

## II.2.7 Applications pour l'archéologie, la paléontologie

### II.2.7.1 MAGIC [Renevier 2001, Nigay 2002, Renevier 2004a]

MAGIC est une plateforme et une application de Réalité Augmentée mobile et collaborative destinée aux archéologues dans leur activité sur le terrain. Chaque archéologue est muni d'un

ordinateur mobile du type Tablet PC et d'un casque de visualisation semi-transparent sur lequel est fixée une caméra pour la prise de vues et la reconnaissance d'objets réels.

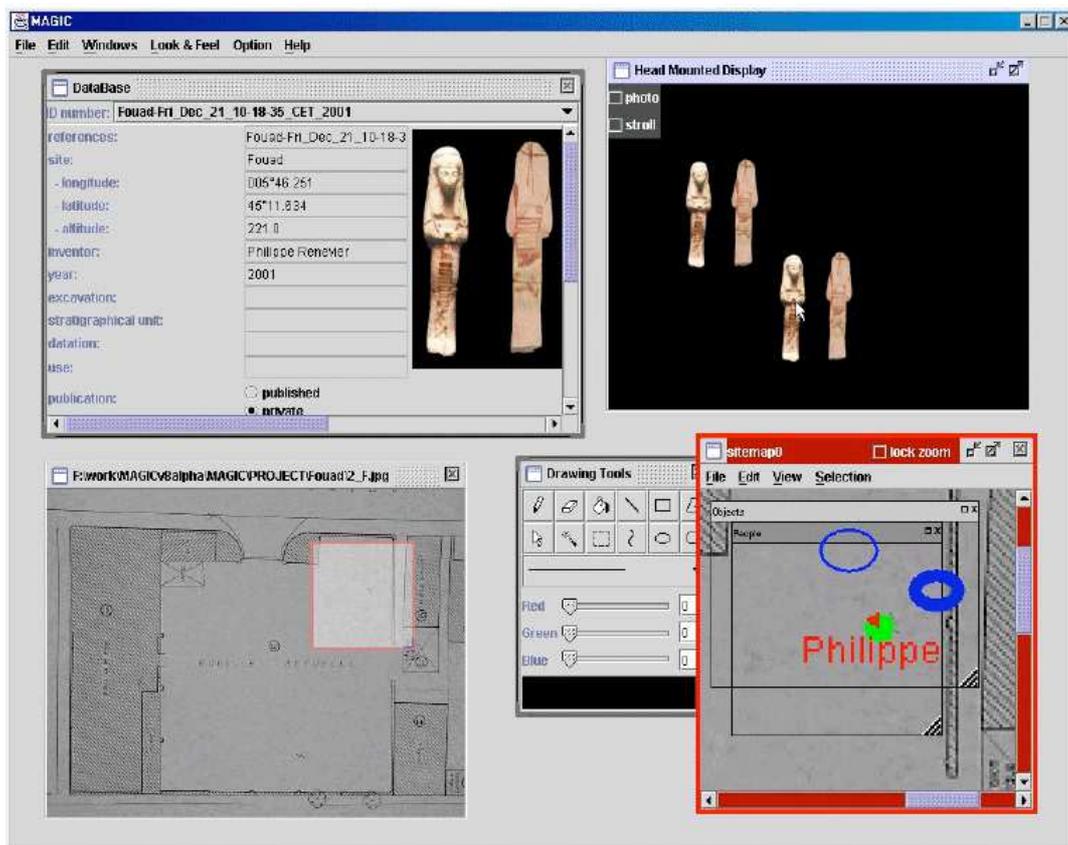


Figure 13 : Interface de MAGIC sur le Tablet PC ; la fenêtre en haut à droite est la vue qui est affichée sur le casque de visualisation semi-transparent (d'après [Renevier 2004a])

Le champ de fouilles est représenté sous forme d'une carte dans l'interface du logiciel sur laquelle est superposée la position des différents utilisateurs (Figure 13). Lorsqu'un archéologue fait une découverte, il peut directement saisir les informations sur la base de données archéologiques et prendre une photo de l'objet grâce à la caméra en cliquant l'objet avec le pointeur de la souris visualisé sur le casque (procédé appelé par les auteurs « réalité clickable »). Il peut également parcourir la base pour voir si d'autres objets similaires ont été découverts au même endroit ou sur d'autres sites ce qui permet de mieux analyser la nouvelle découverte. Il peut aussi entamer une discussion avec les autres archéologues présents sur le terrain ou avec un expert distant. Pendant qu'il parcourt le champ de fouilles, tous les objets qui ont été découverts et qui ont été retirés (car ils étaient sur des strates supérieures), sont visualisés *in situ* en réalité augmentée grâce au casque semi-transparent comme s'ils étaient toujours en place.

### II.2.7.2 Augmented Paleontology [Bimber 2002]

Le terme de Paléontologie Augmentée est proposé par les auteurs pour couvrir l'usage des techniques de Réalité Augmentée dans le domaine de la paléontologie pour aider les scientifiques de cette discipline dans leur recherche. Par exemple il est possible de visualiser en superposition sur le fossile réel d'un crâne de dinosaure les images d'os manquants, des muscles, des yeux et de la peau générées par un modèle de Réalité Virtuelle. Un autre exemple est la superposition du mouvement du pied (virtuel) d'un dinosaure sur le fossile réel de l'empreinte.

Le dispositif technique utilisé est le *Virtual Showcase* qui est une vitrine d'exposition composée d'un miroir semi-transparent et d'un vidéo projecteur qui est utilisé comme une source de lumière contrôlable au pixel près pour illuminer l'objet réel placé dans la vitrine. L'affichage des images

virtuelles est stéréographique, la localisation des utilisateurs étant réalisée par un localisateur infrarouge pour générer une vue augmentée selon la bonne perspective.

L'usage de ces dispositifs dans des musées pour communiquer les résultats au grand public sous une forme attrayante et efficace a également été étudiée par les auteurs et rejoint dans ce cas les applications examinées au paragraphe suivant.

Olivier Bimber [Bimber 2004] a également étudié un prolongement en utilisant, à la place des fossiles réels, des hologrammes optiques de ces fossiles, ce qui présente l'avantage, en particulier dans le cas des musées, de ne pas nécessiter d'avoir l'objet réel présent. Les hologrammes ont une résolution bien supérieure aux affichages 3D existants, puisque les meilleurs hologrammes sont équivalents à un affichage de 1 terapixel. Par contre, ils sont statiques ce qui justifie de les augmenter pour les mêmes raisons que les objets réels sont augmentés.

On peut cependant se demander si l'on peut encore parler ici de « Réalité » Augmentée. Si l'hologramme est bien enregistré sur un support physique (la plaque photographique), l'image vue par l'utilisateur n'a pas de réalité tangible comme un objet physique mais se rapproche plus d'une image virtuelle au sens de la Réalité Virtuelle. Serait-ce donc un cas de virtualité augmentée par le virtuel ?

## **II.2.8 Applications pour les musées, les expositions et les visites de sites**

Nous regroupons dans cette section les applications développées pour les musées, les expositions ou la visite de sites car elles partagent à peu près les mêmes objectifs. Margarita Martinez et al. [Martinez 2004] ont plus particulièrement étudié la conception de systèmes de Réalité Augmentée pour fournir des visites guidées dans le contexte des sites historiques et du patrimoine.

### **II.2.8.1 Musée Augmenté [Rekimoto 1995]**

Cette application utilise le dispositif Navicam [Rekimoto 1995] présenté au paragraphe II.2.1.4. Les visiteurs d'un musée se déplacent devant des œuvres près desquelles ont été disposés des codes-barres en couleur. Lorsqu'un visiteur (muni d'un Navicam) se présente devant une œuvre, la caméra du Navicam reconnaît le code-barres ce qui permet d'afficher dans le casque du visiteur des informations sur l'œuvre qu'il regarde (Figure 14).



Figure 14 : Musée augmenté : vue prise à travers le casque semi-transparent (d'après [Rekimoto 1995])

Les informations peuvent être personnalisées en fonction de l'âge de l'utilisateur, de son niveau de connaissance ou de sa langue préférée. En effet, à l'heure actuelle, les contenus des panneaux d'explication dans les musées sont en général trop basiques pour des experts et trop difficiles à comprendre pour des enfants ou des visiteurs étrangers.

### II.2.8.2 LISTEN [Terrenghi 2004]

Cette application propose un environnement augmenté par du son pour des visiteurs d'expositions d'art. Les utilisateurs portent des casques audio sans-fil qui peuvent fournir du son spatialisé en 3D et écoutent des sources sonores virtuelles disposées dans l'environnement. Les visiteurs en se déplaçant dans l'espace réel naviguent automatiquement dans un espace d'information acoustique conçu comme un complément ou une augmentation de l'espace réel. Ces informations audio sont personnalisées en fonction du contexte des utilisateurs. Par rapport à un audio guide classique le système mémorise les répétitions des actions et réagit en proposant de nouvelles sources audio.

### II.2.8.3 Touring Machine et MARS [Feiner 1997, Höllerer 1999a, Höllerer 1999b]

Steven Feiner et al. [Feiner 1997] proposent un prototype pour l'exploration d'un campus permettant d'obtenir des informations sur les bâtiments. Le dispositif technique, appelé MARS (Mobile Augmented Reality System), est basé sur un ordinateur porté (*wearable computer*) complété par un récepteur GPS pour la localisation 3D de l'utilisateur, un magnétomètre/inclinomètre pour connaître l'orientation de l'utilisateur, un casque de visualisation semi-transparent et un PDA. L'utilisateur sélectionne avec son doigt dans un menu affiché sur le casque le département ou le bâtiment qu'il cherche sur le campus à l'aide d'une tablette tactile (*touchpad*) placée au dos du PDA. Lorsque l'utilisateur a sélectionné le bâtiment voulu, une flèche apparaît dans le casque pointant sur le bâtiment ; dans le cas où le bâtiment n'est pas dans le champ de vision, la flèche est rouge et indique la direction dans laquelle l'utilisateur doit se tourner pour le voir.

Dans une version améliorée du dispositif [Höllerer 1999a, Höllerer 1999b] des points d'intérêt sont signalés sous forme de drapeaux superposés à l'image réelle. L'utilisateur peut sélectionner un de ces points d'intérêt en tournant sa tête de telle manière que le drapeau choisi se trouve le plus proche du centre de l'affichage du casque. Dans ce cas le drapeau passe du vert au jaune en guise de retour visuel. En cas de fixation pendant plus d'une demi-seconde le drapeau est activé ce qui déclenche l'apparition d'un menu que l'utilisateur peut manipuler avec la tablette tactile décrite ci-dessus.

## II.2.9 Jeux

Le domaine des jeux est souvent le plus en pointe pour la diffusion vers le grand public de techniques de visualisation graphique ou de moyens d'interactions innovants. Cependant, aucun jeu commercial vendu à l'heure actuelle ne fait explicitement usage de techniques de Réalité Augmentée car il manque les outils et les plateformes nécessaires pour un développement « industriel » de ces jeux comme le montre [Stapleton 2002]. Les applications que nous présentons ici sont donc des prototypes permettant de tester des dispositifs de recherche.

### II.2.9.1 Mah-Jongg game [Szalavári 1998]

Cette application utilise la Réalité Augmentée pour mettre en œuvre le très ancien jeu traditionnel chinois de Mah-Jongg. Le jeu se joue à 4 joueurs placés avec 144 pièces appelées *tuiles* comportant des figures différentes. Au début du jeu les tuiles sont placées figures cachées pour constituer une *muraille* carrée sur une double épaisseur, l'espace à l'intérieur étant appelé *Empire Céleste*. Sans rentrer dans les détails de la règle, les manipulations que peuvent faire les joueurs sont :

- Prendre des tuiles dans la muraille ou sur la table pour les placer dans sa *main*,
- Manipuler les tuiles dans sa main pour constituer des *combinaisons*,
- Poser une tuile ou une combinaison sur la table.

Dans l'application proposée, la table de jeu et la *main* des joueurs (représentée par une tablette) sont réelles et l'ensemble des pièces est virtuel. Les joueurs portent des casques de visualisation semi-transparent permettant de voir la vue augmentée tant de la table (muraille et tuiles posées) que de leur propre tablette (tuiles de leur *main*) comme le montre la Figure 15. Les joueurs disposent

également chacun d'un stylo avec un bouton pour sélectionner et manipuler les tuiles. L'ensemble des casques, des tablettes et des stylos est localisé en 3D par le système.

L'interaction est donc directe, les joueurs saisissant directement les tuiles virtuelles avec le stylo (qui est augmenté par une pelle virtuelle). Le système vérifie bien sûr que les règles du jeu sont respectées et interdit toute action cherchant à les violer. Pour respecter l'intimité des joueurs, un joueur ne peut voir à travers son casque que les tuiles de la table et les tuiles de sa tablette privée ; ainsi, les tuiles des autres joueurs vues à travers son casque apparaissent blanches ce qui interdit toute velléité de tricher. Quand un utilisateur prend une combinaison sur sa tablette, les figures sont donc invisibles aux autres joueurs jusqu'à ce qu'il la dépose sur la table : à ce moment précis, la combinaison devient visible à tous.



Figure 15 : Jeu de Mah-Jongg augmenté vu par le casque d'un joueur (d'après [Szalavári 1998])

Le système fournit également une aide locale aux joueurs en leur signalant des combinaisons.

Cette application est donc un exemple de travail collaboratif dont la partie coproduction est entièrement supportée par le système. Pour les aspects coordination, une partie est assurée par le système en particulier pour le respect des règles du jeu. Quant à la communication elle est orale et directe entre les joueurs sans intervention du système.

### **II.2.9.2 Rasende Roboter [Demeure 2002]**

A l'origine il s'agit d'un jeu de plateau allemand sur lequel des robots se déplacent. Le but est d'essayer de trouver le plus court chemin entre un robot et un objectif en respectant des règles de déplacement sur le plateau qui comporte différents obstacles. Dans la version augmentée, le plateau de jeu est virtuel, ce qui permet de la paramétrer bien plus que dans la version réelle, mais les robots restent réels pour pouvoir être manipulables directement par les utilisateurs (interface tangible). Une possibilité de calcul du meilleur coup par l'ordinateur est proposée ce qui est intéressant en phase d'apprentissage.

Techniquement, un vidéoprojecteur est placé au-dessus de la table pour afficher le plateau de jeu et les parcours des robots. Une caméra filme la scène et permet de détecter la position des robots.

### **II.2.9.3 TROC [Bouchet 2002, Renevier 2004a]**

TROC est un jeu collaboratif et mobile qui utilise la même plateforme que l'application MAGIC (Cf. § II.2.7.1). Chaque joueur doit collecter une liste d'objets numériques qui sont placés sur le

terrain de jeu au début de la partie. Grâce au dispositif de réalité augmentée complété par une carte du terrain, les joueurs en se déplaçant découvrent les objets. TROC inclue aussi des sons 3D qui aident le joueur à trouver les objets. Les joueurs peuvent aussi utiliser des « outils magiques » pour les aider à localiser les objets ou les autres joueurs sur la carte. Les objets numériques récoltés sont stockés dans 4 cubes physiques transportés par l'utilisateur, le contenu de ces cubes étant affiché sur l'ordinateur ainsi que sur le dessus des cubes en Réalité Augmentée s'ils sont placés devant le casque de visualisation. Pour récupérer les objets, l'utilisateur peut soit utiliser la technique de la « réalité cliquable » (Cf. MAGIC, § II.2.7.1), soit placer le cube devant lui et utiliser une commande vocale « prendre » (Figure 16). Pour gagner et obtenir tous les objets de sa liste, les joueurs doivent collaborer et échanger des objets récupérés. Bien sûr (et c'est même conseillé pour gagner) il est possible de mentir en prétendant avoir un objet et en donner un autre !

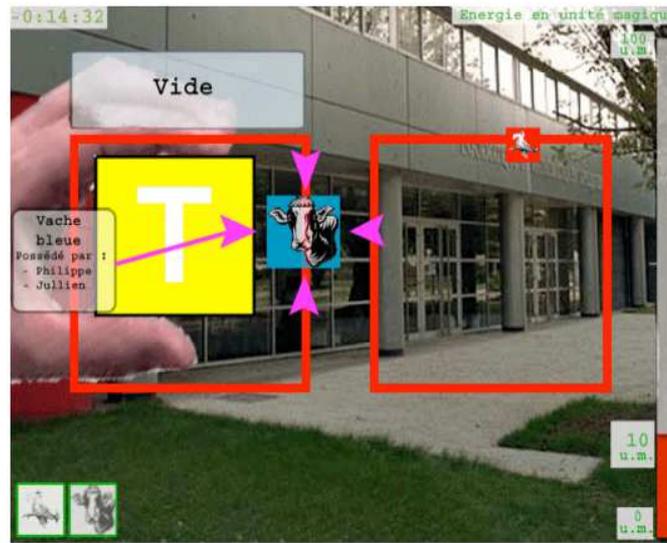


Figure 16 : TROC : récupération d'un objet virtuel par un cube réel (d'après [Renevier 2004a])

## II.2.10 Autres domaines d'application

### II.2.10.1 Augmented library [Rekimoto 1995]

Ce dispositif utilise le dispositif Ubiquitous Talker [Rekimoto 1995] qui est une version étendue de Navicam (Cf. § II.2.1.4) : en plus du casque de visualisation et de la caméra de Navicam, l'utilisateur porte un casque audio et un micro qui lui permettent de dialoguer avec le système par la voix. Dans cette application l'utilisateur peut rechercher des livres dans une bibliothèque en interrogeant le catalogue ou en demandant des informations sur le livre qu'il a en face de lui. Les informations peuvent être retournées vocalement ou sur l'affichage du casque.

### II.2.10.2 Caméléon [Mackay 1997, Mackay 1998a, MacKay 1998b]

Les contrôleurs aériens utilisent des radars et des bandes de papiers (« strips ») sur lesquelles se trouve le plan de vol de l'avion qu'un contrôleur a en charge. Ces strips sont annotés manuellement pour noter des problèmes, pour se souvenir d'actions à prendre et pour communiquer le statut de l'avion aux autres contrôleurs. Le remplacement de ces strips papiers par un affichage sur écran s'est avéré être un échec. L'idée de Caméléon est donc d'augmenter les strips papier en incorporant une résistance dans le support (ce qui permet de les identifier et de capturer leur position) et en les posant sur une tablette tactile (pour capturer les annotations du contrôleur). Un écran tactile placé à côté affiche des informations en regard de chaque strip (Figure 17). Par exemple, si un autre contrôleur suggère un nouveau niveau de vol, cela apparaît sur l'écran à côté du strip concerné. Le contrôleur peut souligner le nouveau niveau s'il est d'accord, en écrire un nouveau ou cliquer sur l'icône « téléphone » pour en discuter. Le contrôleur dispose aussi de commande comme taper une

fois sur un strip pour voir le vol mis en évidence sur le radar ou 2 fois pour voir la route et d'autres informations.



Figure 17 : Vue du système Caméléon montrant les strips, la tablette tactile en dessous et un écran tactile sur la droite (d'après [Mackay 1998a])

### II.2.10.3 Rasa, NISMap et NISChart [Cohen 2004]

Les deux premières applications sont des applications militaires collaboratives basées sur des cartes géographiques en papier.

Avec **Rasa**, les officiers peuvent suivre une bataille en utilisant des Post-it sur lesquels ils écrivent des symboles décrivant les unités, qu'ils collent sur la carte, puis ultérieurement qu'ils déplacent en fonction des informations qui viennent du terrain. La carte murale est fixée sur un grand tableau tactile et les Post-it sont initialement posés sur une petite tablette tactile. Un vidéoprojecteur peut ajouter des informations virtuelles sur la carte. Enfin, un système de reconnaissance de la parole complète le dispositif.

Lorsque l'utilisateur veut placer une nouvelle unité sur la carte il dessine un symbole qui est capturé par la petite tablette tactile et en même temps il la nomme oralement ce qui permet au système de l'identifier (c'est donc un cas d'interaction multimodale). Ensuite il pose le Post-it sur la carte ce qui provoque l'enregistrement de l'unité dans la base de données avec sa position (détectée par le grand tableau tactile). Le système projette un symbole significatif de l'unité sur la carte et distribue le résultat aux autres systèmes. Un des avantages cité par les auteurs et la fiabilité du dispositif : en effet, si l'ordinateur « plante », les utilisateurs peuvent continuer d'ajouter et de déplacer les Post-it ; lorsque le système démarre, il va afficher les anciennes positions, et il suffira juste de réconcilier manuellement la version informatique et la version papier.

**NISMap** est globalement basée sur les mêmes principes mais utilise, à la place du tableau tactile qui est difficilement transportable, le stylo **Anoto**<sup>7</sup> et son concept « d'encre électronique » : Le stylo Anoto peut écrire sur du papier normal, mais s'il est utilisé sur du papier spécial (une fine grille de repérage est pré-imprimée) alors une caméra miniature intégrée au stylo enregistre le tracé et peut le transmettre à un ordinateur par une liaison sans-fil (Bluetooth). Cette écriture peut être simplement archivée telle quelle ou bien être analysée par un système de reconnaissance d'écriture. NISMap utilise donc des cartes géographiques imprimées sur le papier spécial Anoto, ce qui permet de

<sup>7</sup> <http://www.anoto.com/>

capturer l'annotation tracée et sa position sur la carte. L'intérêt, selon les auteurs, est que la carte papier peut être transportée voire être pliée, déchirée, trouée et « fonctionner » encore...

Une application basée sur la même technologie, baptisée **NISChart**, a été développée pour traiter les dossiers des patients dans un hôpital : des formulaires sont imprimés sur le papier Anoto que le médecin remplit de manière habituelle mais grâce à la transmission sans-fil, les informations peuvent directement être stockées dans une base de données sous forme brute ou après reconnaissance de l'écriture. De plus le médecin peut simultanément enregistrer des informations vocales avec le dossier.

### II.2.11 Synthèse et discussion

Après avoir présenté individuellement ces applications regroupées par domaine d'application nous allons présenter une synthèse sous forme d'un tableau faisant apparaître les caractéristiques les plus importantes de ces applications et des dispositifs utilisés (Figure 18 et Figure 19) :

- La première colonne donne le nom de l'application
- La deuxième colonne indique le domaine couvert par cette application.
- La troisième colonne résume les dispositifs d'entrées/sorties utilisés par l'application. de manière très brève. Nous faisons au Chapitre III une revue détaillée de ces dispositifs en les classant et en les analysant en vue de leur représentation dans le modèle IRVO.
- La quatrième colonne précise le type de l'augmentation : RA (Réalité Augmentée) ou VA (Virtualité Augmentée) selon la classification de Dubois en fonction de la nature de l'objet principal de la tâche.
- La cinquième colonne précise la nature de l'augmentation du point de vue de la perception par l'utilisateur. Nous distinguons principalement les augmentations visuelles, sonores et haptiques.
- La sixième colonne précise si l'application exploite la multimodalité (utilisation de plus d'une modalité d'interaction). C'est un critère important des interfaces et nous prendrons les modalités en compte dans la modélisation des relations entre utilisateur et objets.
- La septième colonne précise si l'application permet le travail collaboratif. Cette caractéristique est importante car nos travaux se situent dans le cadre des applications de Réalité Mixte supportant le travail collaboratif. Nous distinguons, selon la taxinomie de Ellis [Ellis 1991] : CLA (collaboratif local asynchrone), CDA (collaboratif distant asynchrone), CLS (collaboratif local synchrone), CDS (collaboratif distant synchrone). Nous pouvons rencontrer aussi CA (collaboratif asynchrone) et CS (collaboratif synchrone) lorsque l'application permet la collaboration tant en local qu'à distance.
- Enfin la huitième colonne précise si l'application supporte la mobilité. Cette caractéristique est moins importante dans le cadre de nos travaux car nous avons choisi de ne pas la traiter spécifiquement contrairement à [Renevier 2003c]. Cependant, comme il s'agit de faire une synthèse, nous avons choisis de la faire figurer. Nous distinguons : F (l'application est fixe par construction), M (l'application est mobile), I (l'application est mobile mais utilisable à l'intérieur de bâtiments seulement), E (l'application est mobile mais utilisable à l'extérieur de bâtiments seulement).

Nom de l'application	Domaine d'application	Dispositifs	Augmentation		multimodalité	Travail collaboratif	mobilité
			Type	Nature			
Charade	Bureautique	Gant numérique	VA	-	-	-	F
DigitalDesk	Bureautique	Table, caméra,	RA	Visuelle	-	-	F
DoubleDigitalDesk	Bureautique	vidéoproj.	RA	Visuelle	-	CDS	F
Chameleon	Bureautique	PDA, localisateur	RA	Visuelle	-	-	I
Navicam	Active Paper Calendar	HMD ou PDA, caméra	RA	visuelle	-	-	I
	Active Door				Bureautique	-	CLA
Audio notebook	Bureautique	Tablette tactile, micro, HP	RA	sonore	oui	-	M
LivePaper	Appli 1	Bureautique	RA	Visuelle	-	CDS	F
	Appli 2	Musique	RA	Visuelle, sonore	-	-	F
	Appli 3	Architecture	RA	Visuelle	-	-	F
Video Mosaic	Montage vidéo	Table, caméras, vidéoproj., HP	RA	Visuelle, sonore	-	-	F
Build-IT	Conception	Table, briques, vidéoprojecteurs	VA	Visuelle	-	CLS	F
SeamlessDesign	Conception	HMD, gants localisateur	VA ?	Visuelle	-	CLS	F
RM et premières étapes de conception	Conception	HMD, caméras, localisateur	RA	Visuelle	-	-	F
Ariel	Architecture	Table, localisateur	RA	Visuelle, sonore	-	(CDS)	F
Interior Design	Architecture	Caméra, écran	RA	Visuelle	-	CDS	F
Architectural Anatomy	Architecture	HMD, localisateur	RA	Visuelle	-	-	I
Augmented Urban planning workbench	Urbanisme	Table, caméras, vidéoprojecteurs	VA	Visuelle	-	CLS	F
Collaborative City Planning	Urbanisme	HMD, caméras	VA	Visuelle	-	CLS	F
Bird's-eye View System	Urbanisme	Caméra, hélicoptère, GPS	RA	Visuelle	-	-	E
Karma	Maintenance	HMD, localisateur	RA	Visuelle	-	-	F
SEAR	Maintenance	PDA, caméra, micro/casque audio	RA	Visuelle sonore	oui	-	M
Collaborative Mechanical Repair	Maintenance	Caméra, écran	RA	Visuelle	-	CDS	F
SharedView	Formation	Caméras, HMD, écran	RA	Visuelle	-	CDS	F
GestureCam	Formation	Caméra pilotée, écrans	RA	Visuelle	-	CDS	F
Training Tool for Assembly Tasks	Enseignement	HMD	RA	Visuelle	-	-	F
MagicBook	Enseignement	Casque, caméra, localisateur	RA	Visuelle	-	CLS	M
Augmented Chemistry	Enseignement	Table, étiquettes, caméra, vidéoproj.	VA	Visuelle	-	-	F
Caretta	Enseignement	Table tactile, PDA vidéoproj., tags RFID	VA	Visuelle	-	CLS	F
CASPER v2	Chirurgie	HMD, localisateur	RA	Visuelle	-	-	F
Virtual Mirror Box	Médecine	Écran, gant	RA	Visuelle	-	-	F
NISChart	Médecine	Stylo Anoto	RA	Visuelle, sonore	oui	CA	M

Figure 18 : Tableau de synthèse de la revue des applications de Réalité Mixte (1<sup>ère</sup> partie)

Nom de l'application	Domaine d'application	Dispositifs	Augmentation		multimodalité	Travail collaboratif	mobilité
			Type	Nature			
MAGIC	Archéologie	HMD, PDA, caméra, GPS	RA	Visuelle	-	CDS	E
Augmented Paleontology	Paleontologie	Écran, miroir semi-transparent, localisateur	RA	visuelle	-	-	F
Musée Augmenté	Musée	HMD ou PDA, caméra	RA	Visuelle	-	-	I
LISTEN	Expositions	Casque audio sans-fil, localisateur	RA	sonore	-	-	I
Touring Machine	Visite de sites	HMD, PDA, GPS, magnétomètre	RA	Visuelle	-	-	E
Mah-Jongg game	Jeux	HMD, localisateur	VA	Visuelle	-	CLS	F
Rasende Roboter	Jeux	Table, caméra, vidéoprojecteur	VA	Visuelle	-	CLS	F
TROC	Jeux	HMD, PDA, caméra, GPS	VA	Visuelle	-	CDS	E
Augmented library	Bibliothèque	HMD ou PDA, caméra, micro/casque audio	RA	Visuelle, sonore	oui	-	I
Caméléon	Contrôle Aérien	Réglettes, tablette tactile, écran tactile	RA	Visuelle	-	CLS	F
Rasa	Militaire	Tablette et tableau tactiles, vidéoproj.	RA	Visuelle, sonore	oui	CS	F
NISMap	Militaire	Stylo Anoto, vidéoproj.	RA	Visuelle, sonore	oui	CS	F

Figure 19 : Tableau de synthèse de la revue des applications de Réalité Mixte (2<sup>ème</sup> partie)

Sur les 44 applications recensées, seules 10 sont à ranger du côté de la Virtualité Augmentée et, en général, dans la catégorie des interfaces tangibles, les 34 autres étant des applications de Réalité Augmentée. Par contre sur la nature de l'augmentation on constate une majorité écrasante d'utilisation d'augmentation visuelle éventuellement couplée à une augmentation sonore. Seules 2 applications sont uniquement du type augmentation sonore. Ceci s'explique aisément par la prédominance du sens de la vision chez l'être humain devant tous les autres sens.

Nous n'avons trouvé, par ailleurs, aucune application qui utilise une augmentation haptique : les seuls papiers trouvés dans ce domaine [Vallino 1999], [Noma 2003] ne proposent que des dispositifs expérimentaux sans véritable application concrète. On peut expliquer cette absence par le fait que, la Réalité Augmentée étant basée sur des objets réels, il y a implicitement un retour tactile et kinesthésique par la simple manipulation des artefacts sans qu'il y ait besoin de rajouter une information supplémentaire provenant du monde virtuel. A l'inverse, la Réalité Virtuelle se préoccupe des interfaces haptiques pour suppléer justement au manque de « réalité » dans ces environnements.

Six applications utilisent le concept de multimodalité : c'est peu et en même temps beaucoup si l'on considère que les travaux sur la multimodalité avaient eu un grand retentissement dans le milieu des années 1990 [Bernsen 1994a, Bernsen 1994b, Bernsen 1995, Vigouroux 1996].

Sur les 44 applications recensées, la moitié supporte le travail collaboratif : nous ne nous attendions pas à un tel nombre. Elles se partagent à nouveau pour moitié entre collaboratif local et collaboratif à distance, certaines offrant les deux possibilités. La collaboration locale n'est parfois pas totalement explicite par les concepteurs : le dispositif est utilisable par plusieurs utilisateurs simultanément ou alternativement et donc, permet une certaine forme de production collaborative, mais souvent sans

supporter les aspects de coordination explicitement. Cependant, les utilisateurs étant en face-à-face, ils peuvent communiquer librement et la coordination est ainsi assurée par un protocole social.

Du point de vue de la mobilité, seules 15 applications supportent la mobilité des utilisateurs et, pour la plupart, celle-ci est limitée à une zone de travail restreinte (une pièce ou bâtiment) en général à cause des technologies de localisation utilisées. Cependant, 4 d'entre-elles sont utilisables en extérieur (grâce au GPS) et 4 autres peuvent être utilisées n'importe où si l'environnement est préparé au préalable (par des étiquettes *ad hoc*).

Nous ne ferons pas de commentaires ici sur les différentes techniques d'augmentation et d'interaction qui sont mises en œuvre car faisant l'objet du chapitre suivant où nous examinons les dispositifs d'entrée et de sorties utilisés dans les systèmes de Réalité Mixte et leur mise en œuvre dans ce contexte.

Enfin nous concluons cette synthèse par le constat qu'il n'y a aucune application commerciale parmi tous les systèmes recensés en dehors du stylo Anoto. Il faut remarquer que le système BUILD-IT était proposé à la vente à la fin des années 1990 par une start-up (TellWare GmbH) qui semble avoir disparue depuis. Cela ne veut pas dire non plus que toutes les applications recensées soient de simples prototypes de laboratoire ! Par exemple, le dispositif SEAR est développé par Siemens dans un but explicite de maintenance industrielle mais rien ne nous permet cependant d'affirmer qu'il est « vraiment » utilisé.

## II.3 Taxonomies

Dans cette section nous présentons plusieurs taxonomies des systèmes de Réalité Mixte. L'intérêt de telles taxonomies est qu'elles donnent des critères précis permettant de classer les différents systèmes les uns par rapport aux autres mais également de les classer vis-à-vis de systèmes qui ne sont pas considérés comme faisant partie de la Réalité Mixte. Ces taxonomies sont aussi utiles en phase de conception pour aider le concepteur dans son choix de dispositifs en lui présentant les différents critères et paramètres qui interviennent et sur lesquels il peut jouer.

Nous avons déjà présenté deux taxonomies : le continuum de Virtualité de Paul Milgram [Milgram 1994] au paragraphe II.1.3 et les styles d'interaction de Jun Rekimoto [Rékimoto 1995] au paragraphe II.1.7.

### II.3.1 Taxonomie des types d'augmentation de Mackay

Wendy Mackay propose de classer les différentes formes de Réalité Augmentée selon le type de l'augmentation [Mackay1996, 1998a] :

- Augmentation de l'utilisateur : l'utilisateur est « augmenté » par le port d'équipements généralement sur la tête comme un casque de visualisation (Cf. Karma, § II.2.4.1) ou sur les mains comme des gants numériques (Cf. Charade, § II.2.1.1).
- Augmentation de l'objet physique : l'objet réel est changé en lui ajoutant des dispositifs d'entrée, de sortie ou un processeur sur ou dans l'objet ; elle cite l'exemple de Paper et sa tortue ou encore des PARCTabs.
- Augmentation de l'environnement autour de l'utilisateur et de l'objet réel : ni l'utilisateur, ni l'objet ne sont affectés directement. Des périphériques indépendants fournissent et capturent de l'information de l'environnement, affichent des informations sur les objets et capturent de l'information sur l'interaction de l'utilisateur avec ces objets. Elle cite l'exemple du DigitalDesk (Cf. § II.2.1.2).

Nous trouvons cette taxonomie trop technocentrée et peu satisfaisante car elle classe, par exemple, « Augmented Urban Planning Workbench » (§ II.2.3.4) et « Collaborative City Planning » (§ II.2.3.5), qui sont deux applications fonctionnellement très proches, dans deux catégories différentes uniquement parce que l'une choisit un vidéoprojecteur et l'autre des casques de

visualisation pour fournir l'augmentation aux utilisateurs. Cette limitation avait été aussi relevée par [Dubois 2001b].

D'autre part, l'augmentation des objets physiques fait aussi partie de l'approche de l'informatique ubiquitaire et cette taxonomie n'aide donc pas à séparer la Réalité Augmentée de l'informatique ubiquitaire.

### II.3.2 Taxonomie de Paul Milgram

Paul Milgram [Milgram 1994a, Milgram 1994b] propose une taxonomie en 3 dimensions des systèmes de Réalité Mixte en considérant 3 axes (Figure 20) :

- **Etendue de la connaissance du monde** (EWK = Extent of World Knowledge) : cet axe détermine dans quelle mesure le système informatique a une connaissance des objets réels ou pas. Il s'agit de mesurer le degré de modélisation du monde réel dans le système informatique :
  - Pas de modèle du tout,
  - Modélisation partielle : où, quoi, où et quoi,
  - Modélisation complète.
- La **fidélité de reproduction** (RF = Reproduction Fidelity) : cet axe détermine la qualité de l'image tant pour les objets réels que virtuels :
  - vidéo monoscopique, modèle fil-de-fer,
  - Vidéo couleur, rendu des surfaces visibles,
  - Vidéo stéréoscopique, gestion des ombres, des textures, de la transparence,
  - Vidéo haute définition, utilisation de techniques de ray-tracing,
  - TVHD 3D avec animations temps-réel et de haute fidélité.
- **Etendue de la métaphore de présence** (EPM = Extent of Presence Metaphor) : détermine comment l'utilisateur se trouve « présent » dans la scène ou immergé. Cet axe s'applique aussi bien aux objets réels que virtuels. Il est à remarquer qu'il n'est pas totalement orthogonal au précédent :
  - Imagerie monoscopique basée sur un moniteur,
  - Imagerie stéréoscopique
  - Ecrans larges, images panoramiques
  - Casques de visualisation, images panoramiques
  - Imagerie en temps réel.

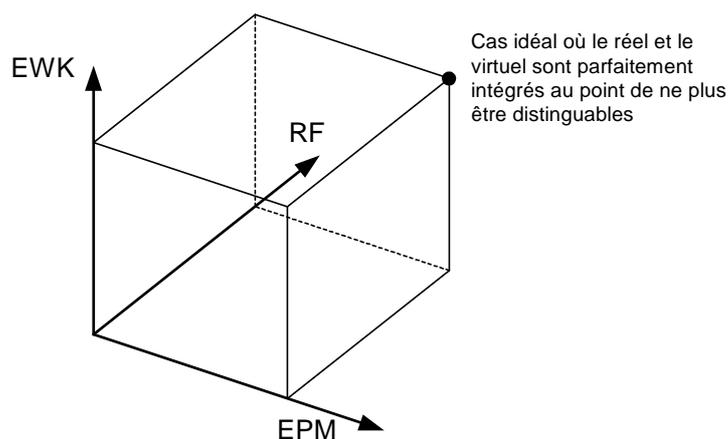


Figure 20 : Taxonomie de Paul Milgram (d'après [Milgram 1994a])

Il faut remarquer que Paul Milgram considère implicitement la Réalité Augmentée visuelle comme une extension de la Réalité Virtuelle ce qui limite ainsi la portée de cette taxonomie. De même, les augmentations du type sonore, tactile ou haptique ne sont pas prises en compte par cette taxonomie.

D'autre part cette taxonomie est basée sur les capacités du logiciel (pour EWK) ou des possibilités techniques des dispositifs d'affichages (pour RF et EPM) sans que cela ait un impact direct sur l'utilisabilité du système par l'utilisateur [Dubois 2001b].

### II.3.3 Taxonomie de Bowskill

Bowskill [Bowskill 1997] considère d'abord la Réalité Augmentée comme la possibilité pour un utilisateur de percevoir des phénomènes physiques, grâce à des capteurs et un système informatique, qui seraient autrement imperceptibles. On voit donc que sa définition n'est pas aussi générale que celle des autres auteurs. Malgré cette remarque, la taxonomie qu'il propose est suffisamment générale pour que nous l'examinions.

Il considère tout d'abord que les systèmes peuvent se classer en deux catégories principales selon le type d'annotation que procure le système :

- **Annotation dépendante** : l'annotation est dépendante de l'intervention de l'utilisateur. Par exemple il peut s'agir de reconnaissance de la parole permettant d'avancer dans un manuel de maintenance en ligne sous forme de pages Web. Il fait une distinction entre dépendance locale et à distance selon que l'opérateur est local ou à distance.
- **Annotation indépendante** : l'annotation ne demande pas d'intervention de l'utilisateur. Dans le cas le plus générique le système de réalité augmentée est une boîte noire qui reçoit des stimuli du monde réel en entrée et génère en sortie une représentation « augmentée » que l'utilisateur perçoit en symbiose avec le monde réel. Il distingue deux sous-catégories :
  - Les systèmes non-conscients de l'environnement (*unaware*), par exemple un casque de visualisation peut fournir un affichage sans tenir compte du contexte spatial.
  - Les systèmes conscients de l'environnement (*aware*), par exemple pour l'alignement des images générées par l'ordinateur avec le patient dans les dispositifs de chirurgie augmentée.

Cette distinction dans le type d'annotation est intéressante car elle est basée sur l'action de l'utilisateur et non sur une caractéristique technique comme la plupart des autres classifications. Elle correspond à la notion de modalité active ou passive de Renevier [Renevier 2003c].

Il est à remarquer que la notion d'awareness de Bowskill est très proche de la notion de Etendue de la connaissance du monde (EWK) de la taxonomie de Paul Milgram (Cf. § II.3.2).

Il identifie également 4 mécanismes communs pour les annotations :

- **Spectral** : dérivé de l'intensité ou de la longueur d'onde de la lumière (ou d'un autre rayonnement électromagnétique) dans le champ visuel de l'utilisateur.
- **Spatial** : annotations contextuelles déterminées par le contenu ou les relations spatiales entre les objets de la scène en cours de visualisation ou plus largement dans l'environnement spatial.
- **Temporel** : il y a beaucoup d'exemples où le mouvement des objets est suivi en utilisant le traitement d'image. Les annotations temporelles peuvent englober ces techniques autorisant l'utilisateur de voir les changements au cours du temps.
- **Emotionnel** : Détecter les états émotionnels d'un utilisateur est la clé à l'intelligence affective et est, par conséquent, un mécanisme potentiel par lequel une réalité augmentée est maintenue.

Cette classification nous semble peu pertinente car elle concerne seulement les augmentations visuelles sauf pour la dernière catégorie qui est d'une toute autre nature et qui nous semble difficile à placer sur le même plan.

### II.3.4 Taxonomies de Robinson & Robertson

Robinson et Robertson nous proposent deux taxonomies [Robinson 2001] qui s'appliquent tant à l'informatique ubiquitaire, qu'aux systèmes de Réalité augmentée par de l'informatique portée<sup>8</sup> ou par videoprojection.

Ils considèrent que les augmentations peuvent tout d'abord être catégorisées par le lieu de l'augmentation. Ils identifient 3 cas :

- **Orienté objet** (*Objet-oriented*) : la fonctionnalité est fortement liée à un objet particulier même s'il y a plusieurs objets avec la même fonctionnalité. Par exemple la possibilité de partager une feuille de papier en transmettant son contenu à distance et en ayant des annotations distantes projetées sur elle, est orienté objet.
- **Multi-objet** (*Multi-object*) : une fonctionnalité qui implique l'interaction entre plusieurs objets est multi-objet. Un cas particulier est quand seulement un type d'augmentation peut avoir lieu à la fois. Par exemple, l'augmentation audio se fait par un seul haut-parleur, il n'est donc pas souhaitable d'augmenter plus d'une feuille de papier avec de l'audio en même temps. Il y a contention lorsque plusieurs objets tentent une augmentation audio simultanée et une communication implicite entre eux est requise pour résoudre lequel "parle".
- **Environnemental** (*Environmental*) : une fonctionnalité qui s'applique à l'ensemble de la zone augmentée est appelée environnementale. Dans le cas de l'informatique ubiquitaire, l'environnement indique une fonctionnalité intégrative qui s'applique à la pièce entière, tandis que pour l'informatique portée, chaque opération qui est sensible au contexte, mais non associée à un objet particulier, est environnementale.

Dans le cas de l'informatique portée, les informations projetées n'augmentent pas toujours des objets réels ou l'environnement réel et sont pourtant encore appelées réalité augmentée. Pour cela, ils considèrent 2 catégories supplémentaires:

- **Non-lié** (*Unbound*): une fonctionnalité qui n'a pas d'ancrage physique.
- **Centrée utilisateur** (*Subject-centric*): une fonctionnalité privée au porteur de l'ordinateur.

D'autre part, pour les fonctionnalités orientées objet, ils distinguent 2 classes:

- Les **propriétés littérales** (*literal features*), qui exploitent la métaphore de la manipulation directe ; dans le cas de la Réalité Mixte, les objets physiques servent d'ancre à ces propriétés.
- Les **propriétés magiques** (*magic features*), comme les menus pop-up, sont des moyens d'interactions puissants mais éloignés de la manière dont les objets réels fonctionnent. Ces fonctionnalités sont donc plus difficiles à apprendre pour l'utilisateur mais offrent plus de puissance à l'utilisateur. Ils proposent que ces propriétés soient associées aux objets physiques en les plaçant à proximité immédiate mais pas directement au-dessus.

Cette distinction est intéressante car elle fait intervenir une notion importante de la Réalité Augmentée : si on cherche à réintroduire le réel, c'est entre autres choses pour profiter de la capacité qu'à un objet physique d'être manipulable par l'utilisateur et donc de permettre une véritable « manipulation directe » de l'interface, l'objet physique servant d'ancre réelle pour manipuler l'augmentation virtuelle. Les auteurs proposent donc de clairement séparer les augmentations virtuelles en adéquation avec l'objet physique des autres fonctionnalités qui sont purement du domaine de l'informatique.

---

<sup>8</sup> Nous avons choisi d'utiliser ce terme proposé par Jean-Marc Robert (Ecole Polytechnique de Montréal) pour traduire le terme anglais de *wearable computer*. La traduction classique en « ordinateur vestimentaire » est en effet trop réductrice et le terme d'« ordinateur corporel » proposé par Jean-Yves Tigli (I3S/ESSI) ne nous semble guère plus satisfaisant.

### II.3.5 Taxonomie de Dubois

Nous avons déjà présenté au paragraphe II.1.3 le **premier axe** de la taxonomie que propose Emmanuel Dubois où il distingue la RA et la VA selon 2 continua distincts [Dubois 1999], en considérant **l'objet de la tâche** qui est soit réel (interaction homme-monde réel) soit virtuel (interaction homme-machine) :

- **Le continuum de la Réalité Augmentée** : l'interaction se passe dans le monde réel et elle est augmentée par l'ordinateur,
- **Le continuum de la Virtualité Augmentée** : l'interaction avec l'ordinateur est augmentée par des objets et des actions dans le monde réel.

Le **deuxième axe** de la taxonomie considère le type d'augmentation :

- **L'exécution augmentée** : un système mixte augmente l'exécution dans le cas où il ajoute des actions nouvelles ou améliore la qualité de la réalisation de la tâche.
- **L'évaluation augmentée** : un système mixte augmente l'évaluation dans le cas où il ajoute des données qui ne sont pas perceptibles lors de la réalisation de la tâche sans le système.

Les deux derniers cas peuvent bien sûr être combinés, ce qui conduit à une taxonomie comportant 6 cas (Figure 21).

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

Figure 21 : Taxonomie de Dubois (d'après [Dubois 2001b])

### II.3.6 Taxonomie des systèmes interactifs mixtes de Nigay et al.

Nigay *et al.* [Nigay 2002, Nigay 2003, Renevier 2004a] proposent une taxonomie pour classer les techniques d'interaction dans les systèmes de Réalité Mixte (Figure 22). Le premier axe considère si l'objet de la tâche est dans le monde réel (Réalité Augmentée) ou virtuel (Virtualité Augmentée). Le deuxième axe caractérise la nature de l'augmentation : l'interaction a lieu en utilisant des objets physiques ou virtuels. Seuls deux cas sur les quatre sont donc des interactions mixtes, les deux autres étant soit une interaction physique soit une interaction avec une interface informatique.

	<i>Interaction utilisant des objets physiques</i>	<i>Interaction utilisant des objets numériques</i>
<i>Tâche dans le monde physique</i>	Interaction physique	<b>Interaction Mixte en Réalité Augmentée</b>
<i>Tâche dans le monde numérique</i>	<b>Interaction Mixte en Virtualité Augmentée</b>	Interface graphique

Figure 22 : Taxonomie des 4 types d'interactions (d'après [Nigay 2003])

### II.3.7 Taxonomies de Renevier

Philippe Renevier propose plusieurs taxinomies dans [Renevier 2003c]. Nous présentons ici les deux plus importantes.

### II.3.7.1 Caractéristiques des augmentations

La première taxinomie propose de caractériser les augmentations selon 5 axes complémentaires [Renevier 2003c] :

- La cible de l'augmentation : *utilisateurs augmentés, objets augmentés, environnement augmenté*, en reprenant la taxinomie de Mackay (Cf. § II.3.1),
- Le type d'augmentation : *exécution augmentée* ou *évaluation augmentée* selon la distinction faite par Dubois (Cf. § II.3.5)
- La temporalité de l'augmentation qui définit la durée de vie des augmentations : *éphémères* ou *persistantes* d'une session à l'autre
- Le mode d'interaction pour la création de l'augmentation qui définit comment l'utilisateur interagit avec une entité augmentée. Ce mode est *passif* si l'augmentation est déterminée automatiquement sans l'intervention de l'utilisateur. Il est *actif* si l'augmentation est spécifiée explicitement (intégralement ou partiellement) par l'utilisateur.
- Le mode d'interaction pour la modification d'une augmentation qui peut être *passif* ou *actif* comme pour la création.

La Figure 23 présente cette taxinomie complétée par des exemples de systèmes.

Cible de l'augmentation	Utilisateur	Objets	Environnement
	<i>Navicam</i>	<i>Phicons</i>	<i>Digital Desk</i>
Type de l'augmentation	Evaluation		Exécution
	<i>Navicam</i>		<i>Phicons, Digital Desk</i>
Temporalité de l'augmentation	Persistant		Ephémère
	<i>Navicam, Phicons</i>		<i>Phicons, Digital Desk</i>
Mode d'interaction lors de la création de l'augmentation	Actif		Passif
	<i>Phicons, Digital Desk</i>		<i>Navicam</i>
Mode d'interaction lors de la modification de l'augmentation	Actif		Passif
	<i>NaviCam (configuration par l'utilisateur), Phicons, Digital Desk</i>		

Figure 23 : taxinomie des caractéristiques d'une augmentation (d'après [Renevier 2003c])

### II.3.7.2 Taxinomie des systèmes mixtes collaboratifs et mobiles

La deuxième taxinomie propose de caractériser les systèmes mixtes collaboratifs et mobiles (tels qu'ils sont définis au § II.1.4) selon 2 axes [Renevier 2003c] :

- La localisation des liens (ou augmentations) par rapport aux utilisateurs, qui comporte trois cas :
  - Les liens sont co-localisés, les utilisateurs se déplaçant autour. C'est par exemple le cas du Virtual Showcase (Cf. § II.2.7.2)
  - Les utilisateurs sont co-localisés, les liens sont liés à cette position et ils n'existent qu'en la présence des utilisateurs.
  - Les liens sont répartis dans tout l'espace d'utilisation. Les utilisateurs se déplacent alors de lien en lien.
- Les types d'interaction qui comportent aussi 3 cas :
  - Les utilisateurs interagissent avec des objets à travers des liens ; les utilisateurs sont à l'initiative de l'interaction.

- Les objets interagissent avec les utilisateurs à travers des liens ; il s'agit du cas inverse du précédent et cela correspond au mode d'interaction passif (Cf. ci-dessus).
- Les utilisateurs interagissent entre eux ; il s'agit d'une communication homme-homme médiatisée.

### II.3.8 Synthèse des taxonomies

La quasi-totalité de ces taxonomies cherche à caractériser le domaine de la Réalité Mixte selon un ou plusieurs points de vue pour distinguer plusieurs sous-catégories mais pas dans le but de séparer les systèmes mixtes des systèmes non-mixte. L'exception notable est la taxonomie de Nigay *et al.* qui permet de distinguer les interactions mixtes des interactions graphiques des environnements WIMP et des interactions dans le monde réel.

Nous proposerons notre propre taxonomie au § VI.3.2 où nous affinerons la classification RA/VA de Dubois et proposerons un moyen de distinguer les systèmes mixtes des systèmes non-mixtes.

Chaque taxonomie se base sur un ou plusieurs critères (jusqu'à 5) pour classer les systèmes. Certains de ces critères peuvent être des sous critères permettant d'affiner un critère de premier niveau. Ces critères peuvent se classer dans 4 grandes catégories :

- **Les critères en rapport avec la technologie :**
  - *Cible de l'augmentation* : apparaît dans la taxonomie de Mackay et comme la 1<sup>ère</sup> caractéristique de la 1<sup>ère</sup> taxonomie de Renevier.
  - *Niveau de modélisation du monde réel par le logiciel* : c'est le critère EWF de la taxonomie de Milgram et awareness dans la taxonomie de Bowskill.
  - *Réalisme de l'affichage* : c'est le critère RF de la taxonomie de Milgram.
  - *Niveau d'immersion* : c'est le critère EPM de la taxonomie de Milgram.
  - *Mécanisme d'augmentation* : ce critère est présent dans la taxonomie de Bowskill (avec la valeur 'émotionnel' qui paraît déplacée dans cette catégorie !)
- **Les critères en rapport avec les objets :**
  - *Nature de l'objet de la tâche* (réel/virtuel) : ce critère est présent dans les taxonomies de Dubois et de Nigay,
  - *Lieu d'augmentation* : 1<sup>ère</sup> critère de la taxonomie de Robinson & Robertson
  - *Propriétés littérales et magiques* : 3<sup>ème</sup> critère de la taxonomie de Robinson & Robertson
- **Les critères en rapport avec les augmentations :**
  - *Liaison des augmentations avec les utilisateurs* : nous classons dans cette catégorie le 2<sup>ème</sup> critère (non-lié/lié) de la taxonomie de Robinson & Robertson et le critère de localisation des augmentations de la 2<sup>ème</sup> taxonomie de Renevier.
  - *Temporalité de l'augmentation* : critère présent dans la 1<sup>ère</sup> taxonomie de Renevier.
- **Les critères en rapport avec l'interaction de l'utilisateur:**
  - *Augmentation de l'exécution ou de l'évaluation* : ce critère est présent dans la taxonomie de Dubois et dans la 1<sup>ère</sup> taxonomie de Renevier.
  - *Utilisateur actif ou passif vis-à-vis de l'augmentation* : nous classons dans cette catégorie les critères 'annotation' de la taxonomie de Bowskill, 'modalité' (4<sup>ème</sup> et 5<sup>ème</sup> critères) de la 1<sup>ère</sup> taxonomie de Renevier, et 'types d'interaction' de la 2<sup>ème</sup> taxonomie de Renevier (qui intègre aussi l'interaction personne-personne médiatisée).
  - *Nature de l'objet supportant l'interaction* : 2<sup>ème</sup> critère de la taxonomie de Nigay

Même si quelques taxonomies partagent des critères similaires de classement (voire réutilisent le même critère) elles divergent sensiblement sur le choix de ces critères ce qui s'explique par le point de vue que chaque auteur adopte et surtout parce que les différents critères ne sont pas orthogonaux. Cette divergence de point de vue montre en fait la difficulté de s'entendre sur une définition commune pour le terme de Réalité Mixte.

On pourrait être tenté de reprendre les 13 catégories de critères recensées pour construire une « super taxonomie » qui croiserait ces 13 catégories. Mais, étant donné le constat que nous venons de faire sur l'hétérogénéité des approches, cela nous semble voué à l'échec.

### *Critique de ces taxonomies*

Il faut remarquer que ces critères sont tous centrés sur la caractérisation du système et aucun n'est centré sur l'humain et sa perception du monde. En effet, même les critères liés à l'interaction ne sont qu'une vue externe de l'humain. Pourtant, les Sciences Humaines pourraient apporter un point de vue intéressant sur les systèmes de Réalité Mixte. En particulier, les travaux sur la perception, et plus particulièrement sur la perception de la vision, constitueraient un apport fondamental. Les philosophes du courant de la phénoménologie, et notamment Merleau-Ponty (1945), ont étudié la perception des objets.

Plus récemment, dans le domaine de l'IHM, les travaux de Paul Dourish [Dourish 1999] avancent l'idée d'une « *embodied interaction* » (interaction incorporée ou incarnée), c'est à dire de la création, de la manipulation et du partage de la signification à travers l'interaction engagée avec les artefacts. Pour cela, P. Dourish se réfère aux philosophes de la phénoménologie (Heidegger, Schutz, Wittgenstein, Merleau-Ponty).

## **II.4 Bilan**

Nous avons essayé de mieux cerner dans ce chapitre le concept de Réalité Mixte qui est au centre de ce mémoire. Nous avons vu que ce domaine de recherche est particulièrement jeune (une dizaine d'année) mais qu'il rassemble des chercheurs provenant principalement de deux communautés :

- Des chercheurs de la communauté de synthèse d'images (*Computer Graphic*)
- Des chercheurs de la communauté d'Interaction Homme-Machine

Les premiers voient la Réalité Mixte comme un prolongement de la Réalité Virtuelle ; la Réalité Virtuelle cherche à reproduire un monde entièrement synthétique et coupé de la réalité. La Réalité Mixte est donc vue comme une réintroduction d'objets réels dans les mondes virtuels. Ainsi leur approche est principalement centrée sur les problèmes de synthèse d'image et d'alignement de ces images virtuelles avec les objets réels (ou les images des objets réels). Le « graal » de ces chercheurs est d'arriver à un environnement dans lequel le réel et le virtuel ne soient plus discernables.

Pour la communauté IHM, l'importance est mise du côté de l'utilisateur et de l'interaction avec le système. Il s'agit de s'attaquer aux gouffres de l'interaction (tels que définis par [Norman 1986]) que provoque l'usage des interfaces informatiques par rapport au monde réel. Le but est donc de profiter du meilleur des deux mondes en les intégrant de la manière la plus harmonieuse possible dans un objectif de cohérence et de continuité. Le but ultime n'est pas ici la fusion des mondes mais bien la transparence dans l'interaction : l'interaction avec les objets virtuels doit être aussi naturelle qu'avec des objets réels.

Ces deux approches ne sont pas contradictoires mais bien complémentaires, pouvant s'enrichir l'une de l'autre.

Il semble cependant regrettable que l'on n'arrive pas à une définition qui puisse être partagée par tous les chercheurs. La tentative de Thomas Baudel [Baudel 1995] est malheureusement restée sans suite et le workshop MIXER [Dubois 2004a] n'a pas mieux réussi.

Nous n'aurons pas la prétention d'apporter la solution dans ce mémoire et nous retiendrons donc les définitions suivantes assez générales mais cependant clairement centrées sur l'interaction de l'utilisateur avec le système :

La **Réalité Mixte (RM)** qualifie un dispositif interactif dans lequel des objets réels et des données informatiques sont mêlées. L'ensemble est perçu de manière cohérente par l'utilisateur par un ou plusieurs de ses sens ce qui lui permet d'interagir avec les différents objets réels ou virtuels de manière cohérente.

La **Réalité Augmentée (RA)** décrit plus spécifiquement un système de RM dans lequel l'interaction avec le monde réel est augmentée par l'ordinateur.

La **Virtualité Augmentée (VA)** décrit plus spécifiquement un système de RM dans lequel l'interaction avec l'ordinateur est augmentée par des objets et des actions du monde réel.<sup>9</sup>

---

<sup>9</sup> Ces deux dernières définitions sont celles de [Dubois 1999].

# Chapitre III Dispositifs d'entrée/sortie pour la Réalité Mixte

*It's called a pen. It's like a printer, hooked straight to my brain.*  
Dale Dauten

Nous avons vu au chapitre précédent, qu'en dehors de dispositifs très spécifiques comme la réglette des strips augmentés de Caméléon, la plupart des applications utilisent le même ensemble de dispositifs d'entrée/sortie permettant de réaliser les augmentations voulues et que nous étudions dans ce chapitre. Il ne s'agit pas de détailler les différentes technologies mises en œuvre dans tous ces dispositifs mais d'avoir un aperçu de leurs caractéristiques principales tout en faisant ressortir leurs usages possibles dans le contexte de la Réalité Mixte.

## III.1 Introduction

Les dispositifs d'entrée/sortie peuvent être classés en deux grandes catégories :

- Les dispositifs qui permettent à l'utilisateur d'interagir avec le système informatique,
- Les dispositifs qui permettent au système informatique de capturer des paramètres de l'environnement physique ou d'agir sur cet environnement physique indépendamment de l'utilisateur.

Comme notre champ d'étude est l'interaction homme-machine, nous nous limitons donc à l'exploration de la première catégorie de dispositifs. Cependant, certains dispositifs entrent dans les deux catégories comme, par exemple, les localisateurs (*trackers*) selon la nature de l'objet qu'ils localisent ; en effet, on peut localiser un cube manipulé par l'utilisateur en remplacement de la souris, ou localiser la position d'objets dans l'environnement.

Notre revue se partagera en deux grandes parties : dans une première section les dispositifs en sortie du système informatique et dans la deuxième section, les dispositifs en entrée.

Pour chaque section, nous classons les dispositifs selon le sens perceptif ou actionnel qu'ils considèrent.

Ainsi, pour les dispositifs en sortie, nous considérons :

- Les dispositifs s'adressant au sens de la vue : c'est la vaste catégorie des dispositifs d'affichage que ce soit du texte ou des images. Cet affichage peut être simple ou sophistiqué allant jusqu'à des casques de visualisation 3D immersifs utilisés en Réalité Virtuelle. Cette catégorie est particulièrement foisonnante car le sens de la vue est le principal sens utilisé par l'être humain.
- Les dispositifs s'adressant au sens de l'ouïe : les dispositifs sont ici assez variés et vont de la simple restitution monaurale de sons, de musique ou de voix, jusqu'aux dispositifs de son spatialisés.
- Les dispositifs s'adressant au sens du toucher : il existe de nombreuses interfaces haptiques et kinesthésiques selon le nombre de degrés de liberté qu'ils prennent en compte.

- Nous ne présentons pas les dispositifs s'adressant aux sens de l'odorat et du goût car ils sont encore à un stade assez expérimental<sup>10</sup> et aucune application recensée à la section II.2 ne les utilise.

Pour les dispositifs en entrée, nous considérons :

- Les dispositifs capturant les mouvements et les actions physiques de l'utilisateur. Cette catégorie est très foisonnante et nous n'en ferons qu'un parcours général en insistant plus particulièrement sur les dispositifs très utilisés en Réalité Mixte.
- Les dispositifs capturant la parole (ou plus généralement les sons) émis par une personne.
- Les dispositifs capturant la direction du regard de l'utilisateur (*eye-tracking*). Bien que ce soient des dispositifs qui capturent un mouvement, ce mouvement est totalement couplé à ce que la personne regarde ce qui le fait considérer à part.
- Les dispositifs liés à la détection d'odeurs ne sont pas abordés car, même si le corps humain émet en permanence des odeurs qu'il est tout à fait possible de détecter par des capteurs biochimiques, il n'est pas possible de le faire de manière consciente et volontaire.
- Enfin, nous abordons de manière anecdotique les interfaces cerveau-ordinateur car ils sont à un niveau encore très expérimental mais présentent des perspectives passionnantes.

Pour chaque dispositif nous examinons dans ses grandes lignes les technologies utilisées surtout pour en faire ressortir l'intérêt et les limitations du point de vue de l'interaction homme-machine. Cette distinction dispositif/technologie est souvent difficile à faire car une même technologie comme la vision artificielle peut servir pour des dispositifs de localisation d'objets ou de manière plus vaste à la reconnaissance directe des objets et de leurs caractéristiques physiques.

## III.2 Dispositifs en sortie

### III.2.1 Dispositifs d'affichage

Nous examinons dans cette section les différentes technologies qui sont utilisées pour afficher des images visibles par les utilisateurs et permettant le mélange d'images réelles et virtuelles. Nous ne parlons pas des classiques écrans d'ordinateurs ou moniteurs vidéo bien qu'ils soient utilisés dans certaines applications de RM (comme Interior Design — § II.2.3.2 — ou Collaborative Mechanical Repair — § II.2.4.3) mais seulement des périphériques « non-standards ».

#### III.2.1.1 Casques de visualisation ou HMD (Head-Mounted Display)

Ces casques ont principalement été développés pour la Réalité Virtuelle. Kocian a examiné de manière détaillée les technologies des casques de visualisation [Kocian 1995]. Même si les produits ont évolué depuis cette époque, les principes généraux restent vrais.

Tout d'abord, l'image générée par l'ordinateur est affichée sur un écran vidéo qui peut être de différentes technologies. Les affichages à tubes cathodiques sont aujourd'hui dépassés (car trop lourd et trop gourmand en énergie) et ce sont surtout les afficheurs à cristaux liquides (LCD) qui sont utilisés. Cette image est ensuite visualisée par l'intermédiaire d'un système optique plus ou moins sophistiqué qui permet en général de construire une image virtuelle (au sens de l'optique !) qui est placée à l'infini pour l'utilisateur. Depuis peu de temps, une nouvelle technologie a fait son apparition qui consiste en l'affichage direct de l'image sur la rétine de l'utilisateur en utilisant un laser de faible puissance [Lewis 2004], ce qui permet d'avoir une très bonne luminosité en utilisant très peu d'énergie tout en évitant des optiques complexes et lourdes à porter.

---

<sup>10</sup> Pour la génération d'odeurs, il existe un produit commercial proposé par la société Olfacom (<http://www.olfacom.com>) mais le site Web n'est plus actualisé depuis le 3 mars 2004 (consulté le 28/11/2004) ce qui ne laisse pas présager de bonnes nouvelles quant à la pérennité de cette société...

Il existe 2 grandes catégories de casques de visualisation :

- Les casques opaques (*non-see-through HMD*) qui ne permettent que de voir l'image virtuelle en coupant l'utilisateur de la vision du monde extérieur,
- Les casques semi-transparents (*see-through HMD*) qui permettent de visualiser l'image du monde virtuel en surimpression de la vision directe du monde réel.

C'est cette deuxième catégorie de casques qui est principalement utilisée pour la Réalité Mixte. Cependant, les casques opaques peuvent aussi être utilisés en filmant le monde réel et en mélangeant cette image avec celle du monde virtuel par un dispositif électronique *ad hoc* ou par logiciel [Azuma 1997], [Barfield 1995] ; la caméra peut être unique, placée entre les deux yeux, ce qui ne fournit qu'une image monoscopique ou il peut y en avoir une pour chaque œil, ce qui permet une vision stéréoscopique.

En pratique, pour la Réalité Mixte sont donc utilisés deux types de casques :

- Des casques semi transparents (*direct see-through HMD*), permettant une vision directe du monde réel, comme dans le dispositif Navicam (§ II.2.1.5), ou les applications Architectural Anatomy (§ II.2.3.3), CASPER v2 (§ II.2.6.1), ou encore MAGIC (§ II.2.7.1).
- Des casques de visualisation à transmission vidéo (*video see-through HMD* ou *HMD with video overlay*), permettant une vision indirecte du monde réel, comme dans MagicBook (§ II.2.5.3) qui utilise un casque à vision monoscopique, ou Collaborative City-Planning System (§ II.2.3.5) qui utilise un casque à vision stéréoscopique.

Les principales difficultés liées à l'utilisation de casques pour la réalité mixte sont [Barfield 1995b] :

- L'alignement des images (*image registration*) : il s'agit de faire coïncider la scène réelle et la scène virtuelle en alignant les objets réels et virtuels pour qu'ils soient correctement perçus par l'utilisateur. Cet alignement est sensible à :
  - la qualité des localisateurs utilisés pour connaître la position du casque (précision, justesse, dérive dans le temps),
  - au délai apporté par les temps de calcul de la scène virtuelle qui font que, en cas de mouvement rapide de l'utilisateur, la scène virtuelle « traîne » derrière la scène réelle.
- L'incrustation vidéo (*video keying*) pour mixer la vidéo de la scène réelle et de la scène virtuelle, dans le cas de l'utilisation de casques opaques. Le mélange peut être réalisé par des techniques classiques utilisées par la télévision ou par des techniques modernes de traitement d'image numérique beaucoup plus souples d'usage et performantes.
- La différence de luminosité entre les scènes réelles et virtuelles qui font que l'une des deux scènes est mal perçue par l'utilisateur car trop sombre. Ceci est particulièrement sensible avec les casques semi-transparents.

Le principal avantage du casque est une meilleure immersion pour des affichages de scènes 3D. Par contre, son usage est strictement individuel ce qui nécessite donc plusieurs casques (et autant de calculs de scène) pour des dispositifs collaboratifs. De plus, la tête de l'utilisateur doit être localisée et la scène recalculée à chaque mouvement pour présenter une image alignée avec le monde réel. Enfin, le poids et le coût de ces casques représentent aussi un frein à leur usage.

### III.2.1.2 Vidéoprojecteurs

Les vidéoprojecteurs sont utilisés en Réalité Mixte de deux manières différentes :

- Pour afficher une image vidéo sur grand écran : l'augmentation est réalisée par incrustation vidéo des images virtuelles avec la vidéo de la scène réelle comme pour les casques de visualisation à transmission vidéo. Le principal intérêt par rapport à la version avec casque est que l'affichage est partageable par plusieurs utilisateurs.
- Par projection sur un miroir semi transparent selon le même principe que les casques semi transparents. Ce principe est utilisé dans le Virtual Showcase (Cf. § II.2.7.2).

- Pour afficher des informations sur des objets réels par projection directe. Cependant les objets doivent être disposés selon un plan perpendiculaire à l'axe de projection pour éviter des déformations optiques de l'image (déformations en trapèze, flou dû à une profondeur de champ limitée) et éviter les ombres portées.

C'est donc principalement pour des applications de type « table interactive » que sont utilisés les vidéoprojecteurs, la table servant de plan de projection sur laquelle on peut poser des objets ayant un faible relief (feuilles de papiers, petites briques). Cependant Ishii et al. [Ishii 2002], dans l'application Augmented Urban Planning Workbench (§ II.2.3.4), ont réussi à utiliser des maquettes d'immeubles en plexiglas qui sont suffisamment transparentes pour limiter les ombres portées tout en étant suffisamment réalistes pour les utilisateurs.

Les principaux avantages sont :

- Dispositif non-intrusif : l'utilisateur n'a pas de casque à porter.
- Image non dépendante de la position de l'utilisateur : il n'y a pas à recalculer la scène quand l'utilisateur bouge.
- L'affichage est partageable par plusieurs utilisateurs ce qui est bien indiqué pour des applications collaboratives en local.
- Le coût : les vidéoprojecteurs sont désormais peu onéreux grâce au marché grand public alors que les casques de visualisation restent chers.

Par contre, les dispositifs utilisant des vidéoprojecteurs sont par construction fixes et donc inappropriés aux applications mobiles.

### III.2.1.3 Mini-écrans

Ces écrans sont souvent utilisés dans des dispositifs mobiles car légers et peu coûteux. Il s'agit souvent de l'affichage d'un PDA ou d'un Tablet PC, comme par exemple dans le Musée Augmenté (§ II.2.8.1) ou SEAR (§ II.2.4.2).

Un dispositif, TUISTER [Butz2004], du domaine des interfaces tangibles, utilise des mini-écrans (64 x 16 pixels) pour créer un périphérique innovant pour la navigation dans des structures hiérarchiques.

### III.2.1.4 Classes des systèmes de Réalité Mixte de Paul Milgram

Au-delà du continuum Réel – Virtuel exposé au paragraphe II.1.3, Paul Milgram spécifie plus concrètement 6 classes de systèmes de Réalité Mixte [Milgram 1994a, Milgram 1994b]. Il veut les présenter comme génériques mais en pratique ces classes ne considèrent que l'augmentation de la vision par mélange d'images réelles et virtuelles :

- Classe 1 : Réalité Augmentée basée sur un moniteur vidéo (type WoW = Window-on-the-world, non-immersif) : les images générées par le système informatique sont simplement superposées sur les images du monde réel (avec des techniques de mixage vidéo). Cette affichage peut être monoscopique ou stéréoscopique en utilisant des lunettes *ad hoc*.
- Classe 2 : même principe que la classe 1 mais en utilisant un casque de visualisation opaque.
- Classe 3 : système de RA basés sur des casques de visualisation semi-transparentes (*HMD with optical see-through*) permettant de voir directement le monde réel et sur lequel se superpose des images informatiques.
- Classe 4 : même principe que la classe 3 mais utilisant une source vidéo plutôt qu'une vision directe (*video see-through HMD*). La différence avec la classe 2 est que la vidéo du monde extérieur doit correspondre de manière orthogonale au monde extérieur réel simulé ce qui se fait avec la classe 3.

- Classe 5 : Virtualité Augmentée basée sur un moniteur vidéo : les images générées par le système informatique sont affichées sur le moniteur auquel sont superposées l'image du monde réel par des techniques de mixage vidéo.
- Classe 6 : Virtualité Augmentée partiellement ou totalement immersive (par exemple, affichage sur grand écran) ; le mixage réel – virtuel est obtenu comme avec la classe 5.
- Classe 7 : Virtualité Augmentée partiellement immersive qui permet des interactions avec des objets réels comme 'atteindre' et 'saisir' avec la main réelle de l'utilisateur.

Cette taxonomie pourrait faire une bonne synthèse pour les dispositifs d'affichage de Réalité Augmentée. Malheureusement elle est incomplète puisqu'elle ne considère pas le cas de la projection d'images sur des objets réels ce qui est pourtant une technique massivement utilisée en Réalité Augmentée, par exemple avec les tables interactives (Cf. § III.3.1.4), ainsi que le cas des mini-écrans.

### III.2.2 Restitution du son

Michael Cohen et al. ont étudié de manière détaillée l'usage du son dans les interfaces informatiques et en particulier pour la Réalité Virtuelle [Cohen 1995].

Pour la restitution sonore, on peut utiliser :

- Un ou plusieurs haut-parleurs (HP) disposés à côté des dispositifs ou dans la pièce ; cette disposition permet un partage aisé par plusieurs personnes mais est, par construction, fixe.
- Un casque audio porté par l'utilisateur ; comme principaux avantages on peut noter que : le dispositif est mobile, les signaux sonores peuvent être individualisés pour chaque utilisateur, en cas d'utilisation simultanément de micros pour la prise de son, les effets Larsen sont facilement annulés.

Les sons restitués peuvent être :

- Monauraux : un seul haut-parleur est utilisé. Ceci est suffisant, pour restituer des sons simples comme des alarmes.
- Stéréophoniques : il faut deux haut-parleurs (ou un casque). La restitution du son est partiellement spatialisée (droite-gauche) ce qui permet d'améliorer le réalisme du rendu sonore.
- Spatialisés : il faut deux haut-parleurs ou plus. Basé sur la perception psycho-acoustique des sons par les êtres humains et grâce à un traitement numérique du son *ad hoc* il est possible de placer une source sonore virtuelle n'importe où dans l'espace. Ce dispositif fonctionne bien avec un casque audio, mais si l'utilisateur bouge la tête, il faut en temps réel modifier la position de la source ce qui oblige à utiliser un localisateur. On évite cet écueil en utilisant des haut-parleurs disposés dans la pièce mais alors l'utilisateur ne peut se déplacer que dans un espace restreint car l'effet de spatialisation est limité à une zone donnée. C'est également un problème si plusieurs utilisateurs se trouvent dans la pièce car il perturbent la propagation du son et annihilent l'effet de spatialisation.

Si la Réalité Virtuelle, qui vise à une immersion maximale de l'utilisateur dans l'environnement virtuel, utilise surtout les sons spatialisés, la Réalité Mixte se contente souvent des technologies les plus simples : l'Audio Notebook (§ II.2.1.6), DoubleDigitalDesk (§ II.2.1.3) ou SEAR (§ II.2.4.2) utilisent du son monaural, Video Mosaic (§ II.2.2.1) ou LivePaper (§ II.2.1.7) utilisent du son stéréophonique et seul LISTEN (§ II.2.8.2) utilise du son spatialisé. Il convient de noter que l'utilisation du son pour la Réalité Mixte est encore peu abordée par les chercheurs comme l'explique [Sundareswaran 2003].

Une façon toute simple d'obtenir un son spatialisé pour augmenter des objets réels est de placer directement des petits haut-parleurs sur ou dans les objets eux-mêmes. Cette technique a été mise en œuvre avec succès dans le projet Telepresence [Buxton 1992] pour un système de

visioconférence avec les dispositifs Hydra qui sont des « substituts vidéo » (*video surrogate*) des personnes distantes : chaque dispositif contient une caméra, un petit écran vidéo, un micro et un haut-parleur et est disposé de manière *ad hoc* sur la table pour simuler la présence de la personne distante.

### III.2.3 Interfaces haptiques et kinesthésiques

Ces interfaces s'adressent globalement au sens du toucher qui se scinde en [Hannaford 1995] :

- Perception tactile ou haptique : perception des formes et des structures (mais aussi de la température) grâce aux terminaisons nerveuses de notre peau.
- Perception kinesthésique ou proprioceptive : perception des forces et des mouvements par l'intermédiaire de nos muscles et articulations.

Les interfaces tactiles ont été étudiées extensivement par [Kaczmarek 1995]. Les affichages tactiles les plus connus sont certainement les afficheurs Braille utilisés par les non-voyants pour lire le contenu d'un écran (et généralement associé à un synthétiseur de parole). Des afficheurs électromécaniques basés sur le même principe pour « lire » tactilement des images ont aussi été développés : cependant le temps de parcours de l'image peut être très élevé. Nous pouvons aussi citer les dispositifs vibro-tactiles montés dans des souris qui permettent de simuler des textures lorsque l'utilisateur déplace la souris sur la table. Enfin des simulations électriques directes par des électrodes permettent de simuler une grande gamme de sensations tactiles et elles peuvent être placées sur différentes parties du corps.

Les interfaces kinesthésiques ont été étudiées en détail par [Hannaford 1995]. Il existe deux types de dispositifs : les dispositifs à retour d'effort (force-feedback) et ceux à retour de mouvement (displacement-feedback). Dans tous les cas il s'agit de simuler le plus possible les efforts qui s'exercent sur la main de l'utilisateur lorsqu'il la déplace ou essaye d'aller au-delà d'un point donné. Les systèmes peuvent aller des plus simples (unidimensionnels) aux plus sophistiqués (bras de télémanipulateurs à 6 degrés de liberté).

En pratique les interfaces haptiques et kinesthésiques intéressent peu la Réalité Mixte. A notre connaissance, seuls [Vallino 1999] et [Noma 2003] se sont intéressés explicitement à la question. En effet, soit on est dans le cas d'objets réels directement manipulés par l'utilisateur et il n'y a donc pas besoin de simuler des perceptions tactiles ou kinesthésiques puisqu'elles sont directes, soit les objets sont virtuels et dans ce cas, soit le retour tactile n'est pas considéré comme important, soit les techniques des interfaces tangibles sont exploitées et l'utilisateur a donc à nouveau une perception physique directe de ces objets qu'il manipule.

## III.3 Dispositifs en entrée

Nous explorons dans le premier paragraphe les différents périphériques en entrée qui sont principalement liés aux gestes de l'utilisateur, généralement sa main dominante mais il peut aussi s'agir de ses pieds (pédales), des mouvements de sa tête ou de tout son corps. Nous parlerons dans les deux paragraphes suivants des périphériques d'entrée pour l'audio et des eye-trackers qui sont liés à la vision. Nous finirons par les interfaces cerveau-ordinateur qui sont encore à un stade très expérimental.

### III.3.1 Dispositifs liés à la position, au mouvement, à la force

De nombreux auteurs ont étudié les périphériques d'entrée comme [Buxton 1983], [Mackinlay 1990] ou [Mackenzie 1995]. Buxton avait proposé une taxonomie de ces périphériques [Buxton 1983] en fonction du nombre de dimensions spatiales supportées (1, 2 ou 3) et du type de propriété physique mesurée (position, mouvement ou pression). Foley en 1984 (cité par [Mackinlay 1990])

avait aussi proposé une taxinomie arborescente. Mackinlay [Mackinlay 1990] propose une taxonomie regroupant les deux précédentes (Figure 24) :

- En abscisse il considère les 6 degrés de liberté possibles : en position (X, Y, Z) et en rotation (rX, rY, rZ). Chacune de ces grandeurs peut être mesurée avec une précision plus ou moins grande.
- En ordonnée il considère 4 types de mesures qu'effectuent les périphériques :
  - Pour la position, il s'agit de : la position (P), le mouvement (différence de position=dP), la force (F), la différence de force (dF).
  - Pour la rotation, il s'agit de : l'angle (R), la différence d'angle (dR), le couple (T), la différence de couple (dT).

	Linear			Rotary		
	X	Y	Z	rX	rY	rZ
P	Tablet Light Pen Touch Panel Touch Screen		Sliding Pot	Absolute Joystick 3D Joystick		Rotary Pot
dP	Mouse Tasa X-Y Pad		Treadmill Tasa Ferinstat	Trackball 3D Trackball		Continuous Pot Thumbwheel
F			Pressure Pad	Isometric Joystick		Torque Sensing
dF						T dT
	1 10 100 Inf	1 10 100 Inf	1 10 100 Inf	1 10 100 Inf	1 10 100 Inf	1 10 100 Inf

Figure 24 : Taxonomie de Mackinlay des périphériques d'entrée sur laquelle ont été placés les périphériques considérés par Buxton (d'après [Mackinlay 1990]).

Nous ne détaillons pas tous les périphériques d'entrée par la suite mais nous nous focalisons sur ceux qui présentent un réel intérêt pour la Réalité Mixte.

### III.3.1.1 Capteurs de positions – localisateurs (trackers)

Ces dispositifs sont en général composés d'un système central immobile et de dispositifs posés sur les objets mobiles dont on veut suivre la position. On peut distinguer 2 catégories :

- Les dispositifs mobiles émettent des signaux caractéristiques qui sont captés par le système central et utilisés pour déterminer leur position.
- Le système central émet un signal (champ électrique, champ magnétique, lumière visible ou infrarouge, etc.) que les dispositifs mobiles captent et utilisent pour calculer leur position.

La liaison avec les dispositifs mobiles peut utiliser le même signal que pour la mesure ou une autre technique de transmission filaire ou sans-fil.

Ces dispositifs ont été étudiés en détail par [Kocian 1995, Azuma 1993, Azuma 1997]. A titre d'exemple, nous citerons le système Flock of Birds de Ascension Technology Corporation<sup>11</sup> qui utilise des émetteurs mobiles qui transmettent un signal sous forme de champ magnétique qui est détecté par un récepteur central. Chaque émetteur (un petit cube d'environ 1 cm<sup>3</sup>) doit être collé sur l'objet pour pouvoir le suivre en 6 degrés de liberté. La précision est millimétrique mais le volume de travail est limité à environ 1 m<sup>3</sup>. De plus, la présence de pièces métalliques dans la zone de travail peut perturber le système de mesure ce qui conduit les chercheurs à utiliser au maximum le bois pour la construction de leurs prototypes !

<sup>11</sup> <http://www.ascension-tech.com/>

Une autre technique de plus en plus utilisée, car simple à mettre en œuvre et moins coûteuse, est par l'utilisation des techniques de vision artificielle. La scène filmée par une caméra vidéo (ou deux pour avoir une localisation en 3D) est analysée pour retrouver les objets et les suivre dans leur mouvement. Une étape de calibration préliminaire permet d'obtenir les coordonnées directement dans le référentiel de la pièce. Pour assurer la reconnaissance des objets et améliorer la robustesse du suivi, il est fait usage de marqueurs spécifiques (étiquettes en papier avec un dessin en noir et blanc) collés sur les objets. Ainsi Navicam (§ II.2.1.5) ou CyberCode [Rekimoto 2000] utilisent ce principe de localisation d'objets. Nous citerons plus particulièrement l'AR toolkit [Kato 2000, Billingham 2001a, Billingham 2001b, Poupyrev 2002] qui est aussi basée sur ce principe, et qui présente l'avantage d'être proposée sous forme de logiciel libre téléchargeable et prêt à l'emploi. Ainsi, plusieurs applications de Réalité Mixte comme SEAR (§ II.2.4.2), MagicBook (§ II.2.5.3), Augmented Chemistry (§ II.2.5.4) l'utilisent et nous l'avons également utilisé pour nos propres prototypes (Cf. Chapitre VIII).

Pour être complet, il ne faut pas oublier le GPS (Global Positioning System) basé sur une constellation de satellites tournant autour de la terre et qui permet de se repérer avec une précision de l'ordre de quelques mètres n'importe où dans le monde en utilisant un récepteur peu coûteux. Cette précision peut être suffisante pour des applications de Réalité Mixte comme Bird's-eye View System (§ II.2.3.6), MAGIC (§ II.2.7.1) ou Touring Machine (§ II.2.8.3). Le GPS ne donne qu'une position en 3D (latitude, longitude, altitude) et doit être complété par un inclinomètre/magnétomètre pour mesurer les 3 autres degrés de liberté en rotation. Le plus gros inconvénient du GPS est que son utilisation n'est possible qu'à l'extérieur des bâtiments car il faut être en vue directe des satellites. C'est pourquoi, à l'intérieur des bâtiments, de nombreuses technologies sont utilisées ; elles sont recensées dans [Hightower 2001].

Toutes ces techniques de localisation sont massivement utilisées en Réalité Augmentée car la plupart des applications ont besoin de connaître la position des objets réels utilisés dans le système comme, par exemple :

- Quand la position d'un outil ou d'un objet est suivie en temps réel pour servir d'interface avec le système (cas typique des interfaces tangibles). Par exemple Morten Fjeld *et al.* [Fjeld 1999, Fjeld 2000] ont étudié l'utilisation de briques comme interface pour interagir avec la table interactive BUILD-IT (§ II.2.2.2). L'objet peut même être directement la main ou le doigt de l'utilisateur comme dans [Crowley 1995], [Roussel 1999], [Berard 2000].
- Si la vision de l'objet réel est augmentée, il faut pouvoir connaître la position de cet objet pour pouvoir aligner les données virtuelles avec l'objet réel.
- Si l'augmentation visuelle est fournie par un casque de visualisation, et que l'on souhaite un alignement précis des données virtuelles avec le monde réel, le casque doit être localisé en temps réel pour recalculer l'image virtuelle en fonction de la position et de l'angle de vue de l'utilisateur.

Bien sûr, il existe de nombreux cas où la configuration du dispositif permet de s'affranchir de localiser les objets : par exemple si l'objet de la tâche est fixe et qu'on utilise un vidéoprojecteur pour fournir l'augmentation comme dans le DoubleDigitalDesk (§ II.2.1.3) ou encore dans des systèmes « auto-alignés » par construction comme SharedView (§ II.2.5.1). De même si on travaille en 2D sur une table, alors les tables tactiles sont utilisables (Cf. ci-après).

### III.3.1.2 Vision artificielle

Les techniques de vision artificielle et de reconnaissance de formes ont été développées initialement surtout pour des usages en robotique. Grâce à l'augmentation de puissance de calcul et à la diminution des coûts ces techniques sont désormais accessibles sur de simples ordinateurs [Crowley 1995].

Les applications de ces techniques sont nombreuses. Comme nous l'avons vu au § III.3.1.1 avec l'ARToolkit il est possible de localiser et suivre des objets réels, en leur ajoutant des marqueurs spécifiques. Il est également possible de suivre un doigt qui peut servir de pointeur d'interaction directe soit dans un dispositif fixe [Crowley 1995] soit avec un ordinateur porté (*wearable computer*) [Starner 1997]. Thevenin *et al.* ont appliqué ces techniques pour capturer des inscriptions sur un tableau blanc et ainsi pouvoir en faire des copies virtuelles [Thevenin 1999], qu'ils ont mis en œuvre dans le Tableau Magique [Bérard 2000].

Ces techniques sont également mises en œuvre dans certains eye-trackers (Cf. § III.3.3) comme par exemple le dispositif de [Choi 1999] qui permet simultanément de détecter les mouvements de la tête et la direction du regard d'un utilisateur.

### III.3.1.3 Gants numériques

Les gants numériques sont principalement utilisés dans les applications de réalité virtuelle [Mackenzie 1995]. Les plus simples permettent de connaître la flexion des 5 doigts de manière binaire (fléchi / non fléchi). Les plus sophistiqués mesurent précisément l'angle de chaque phalange. Ils sont souvent couplés à un localisateur pour connaître la position absolue de la main dans l'espace (position et rotation). Certains gants sont aussi couplés à des dispositifs de retour d'efforts.

Ces périphériques sont peu utilisés dans les applications de Réalité Mixte : seuls Charade (§ II.2.1.1), SeamlessDesign (§ II.2.2.3) et Virtual Mirror Box (§ II.2.6.2), parmi les applications que nous avons examinées dans la section II.2 en utilisent un ou plusieurs. Et sur ces trois applications, les deux premières se trouvent à la frange de la Virtualité Augmentée et sont très proches de la Réalité Virtuelle.

### III.3.1.4 Tablettes et tableaux tactiles ou numériques

Les tablettes tactiles (*touchpad*) sont souvent utilisées par les infographistes à la place de la souris car elles travaillent en coordonnées absolues et non relatives. La Réalité Augmentée s'en sert principalement pour la capture de dessins ou de textes réalisés à la main sur une feuille de papier placée sur la tableau comme dans Audio Notebook (§ II.2.1.6), Caméléon (§ II.2.10.2) ou Rasa (§ II.2.10.3).

Les tablettes tactiles couplées à des écrans forment des écrans tactiles comme ceux des PDA et des Tablet PC. Leur principal intérêt est, par rapport à la souris, de permettre une véritable manipulation directe des objets virtuels affichés sur l'écran. SEAR (§ II.2.4.2), Caméléon (§ II.2.10.2) ou MAGIC (§ II.2.7.1) sont trois exemples d'applications utilisant ce genre de dispositif.

Il existe des tableaux numériques de grande dimension (jusqu'à 1 m<sup>2</sup>) qui permettent à plusieurs personnes de travailler ensemble. Ils peuvent être utilisés comme des tableaux blancs classiques ou bien, en ajoutant un vidéoprojecteur, comme un écran tactile. Plusieurs produits commerciaux existent comme TeamBoard ou SmartBoard<sup>12</sup> pour citer les plus connus. Ces dispositifs sont utilisables en Réalité Augmentée par exemple en collant des objets fins comme la carte et les Post-it de Rasa (§ II.2.10.3).

En plaçant de tels tableaux horizontalement, on obtient des tables interactives sur lesquelles on peut poser des objets légers et qui sont donc utilisables en Réalité Mixte.

Le domaine de recherche autour de ces tables et tableaux interactifs est particulièrement actif, avec pour preuves le « workshop on Collaboration with Interactive Walls and Tables » qui s'est tenu à la conférence Ubiquitous Computing 2002 [UbiComp-WS 2002] et les nombreuses applications qui

---

<sup>12</sup> <http://www.teamboard.com/>, <http://www.smarttech.com/>

exploitent ce principe comme le DigitalDesk (§ II.2.1.2), LivePaper (§ II.2.1.7), BUILD-IT (§ II.2.2.2), Augmented Urban Planning Workbench (§ II.2.3.4), Caretta (§ II.2.5.5), DiamondTouch [Shen 2002, Leigh 2002], EnhancedDesk et EnhancedWall [Nakanishi 2002], Collaboration Table [Umemuro 2003].

Dans ce domaine, et lié aux interfaces haptiques, il faut citer le Proactive Desk [Noma 2003] qui propose une table interactive qui utilise des moteurs linéaires pour fournir un retour d'effort en translation sur les objets (ou sur la main de l'utilisateur) sans qu'il y ait une quelconque liaison mécanique ou électrique avec les objets.

### III.3.2 Périphériques d'entrée audio

Le seul périphérique de capture du son qui existe est le microphone dont le signal est, en général, immédiatement numérisé dans le système informatique. Ce microphone peut être placé de manière fixe dans l'environnement de travail, soit le plus souvent intégré à un casque audio. Dans tous les cas, la prise de son (niveau sonore, bruit ambiant) est une opération délicate qui limite son usage dans les applications surtout à l'extérieur.

Le son est utilisé de trois manières dans les dispositifs de Réalité Mixte :

- Il est enregistré sur un support numérique en vue d'une restitution ultérieure comme dans l'Audio notebook (§ II.2.1.6),
- Il est directement transmis à distance pour fournir un canal de communication audio dans les applications collaboratives à distance comme dans le DoubleDigitalDesk (§ II.2.1.3),
- Il est utilisé comme commande vocale pour piloter l'interface du logiciel ou saisir des données. Dans ce cas, il faut un module de reconnaissance de la parole qui, même si les progrès ont été conséquents ces dernières années, nécessite des phases d'apprentissage pour pouvoir espérer un résultat correct. Les applications SEAR (§ II.2.4.2) ou Rasa (§ II.2.10.3) en sont deux exemples.

### III.3.3 Eye-trackers ou oculomètres

Nous avons exploré l'usage des oculomètres (ou eye-trackers) dans [Chalon 2001]. Ce sont des dispositifs qui permettent de suivre la direction du regard d'un utilisateur et éventuellement des paramètres complémentaires comme la taille de la pupille, la rotation axiale de l'œil, ou le clignement des paupières.

Ils ont surtout été utilisés pour des études sur l'analyse du regard et plus particulièrement de la lecture [Rayner 1998]. Cependant, la direction du regard peut être utilisée pour interagir avec l'ordinateur [Jacob 1993, Hutchinson 1993, Jacob 1995, Charbonnier 1995, Colombo 1997], ce que nous appelons de manière un peu ironique : « interaction à l'œil » [David 2001b].

Il existe 4 principales familles d'oculomètres [Charbonnier 1995] :

- Les dispositifs utilisant des lentilles de contact magnétiques (posées sur l'œil) qui sont placées dans un champ magnétique créé autour de la tête de l'utilisateur. Les variations du champ magnétique sont détectées ce qui permet d'en déduire la direction du regard.
- Les dispositifs utilisant l'électro-oculographie (EOG). Il existe une différence de potentiel entre la cornée et la rétine qui produit un champ électrique qui peut être détecté par des électrodes placées sur la peau autour de l'œil. Les variations de ce champ électrique sont ainsi détectées ce qui permet d'en déduire la direction du regard.
- Les dispositifs par photodétecteurs. Ils utilisent la différence de pouvoir réfléchissant entre l'iris et la sclérotique blanche sous éclairage infrarouge. Cette méthode présente de nombreux inconvénients et n'est pratiquement pas utilisée.
- Les dispositifs par caméra et traitement d'image :

- Les dispositifs par réflexion cornéenne. L'œil est éclairé par des sources de lumière infrarouge et il est filmé par une caméra infrarouge. Les reflets mobiles des sources infrarouges sur le cristallin et la cornée sont analysés par un logiciel qui permet de calculer la direction du regard.
- En lumière naturelle. Une caméra filme le visage de l'utilisateur et un logiciel spécialisé dans la reconnaissance visuelle se charge d'identifier l'œil et la direction du regard.

Les dispositifs utilisent le plus souvent la méthode de réflexion cornéenne car elle a l'avantage d'être assez précise tout en étant moins intrusive que les deux premières méthodes. Cependant, avec l'amélioration des algorithmes de traitement d'images et la puissance des ordinateurs, la méthode par traitement d'images en lumière naturelle est de plus en plus utilisée car moins coûteuse et plus simple à mettre en œuvre.

Les dispositifs peuvent être :

- Fixes, ce qui impose que l'utilisateur soit immobile, voire immobilisé par une têtère pour de bons résultats de mesure. Cependant certains dispositifs utilisent une caméra réglable par ordinateur (orientation et zoom) ce qui permet de suivre les mouvements de la tête dans une plage d'une trentaine de centimètres.
- Mobiles : dans ce cas ils sont fixés sur un casque porté par l'utilisateur, ce casque pouvant être aussi un casque de visualisation. La direction du regard est alors calculée dans le référentiel de l'utilisateur et non plus de la pièce, ce qui peut être suffisant pour certaines applications. Cependant, dans la plupart des cas, un localisateur est couplé au eye-tracker pour pouvoir faire le changement de référentiel par calcul.

Malgré de nombreux progrès techniques les eye-trackers sont loin d'être parfaits et demandent des phases de calibration longues pour chaque utilisateur. De plus, malgré une certaine concurrence entre les fabricants, les prix restent élevés ce qui fait que ces dispositifs sont encore assez marginaux comme périphériques d'interaction homme-machine sauf dans le contexte des interfaces pour personnes handicapées (tétraplégiques) où ils permettent une avancée significative en terme de confort d'utilisation.

Dans les applications que nous avons recensées, aucune n'utilise la direction du regard. Il serait pourtant agréable dans la Touring Machine (§ II.2.8.3) [Höllner 1999a, Höllner 1999b] de pouvoir sélectionner un drapeau en le regardant plutôt que d'avoir à tourner la tête pour mettre le drapeau au centre du champ de vision.

### III.3.4 Interfaces cerveau-ordinateur

Le principe de ces interfaces appelées BCI (*Brain-Computer Interfaces* ou parfois *Brain-Controlled Interaction*) est de capter les signaux EEG (électro-encéphalogrammes) des utilisateurs à partir d'électrodes posées sur leur scalp afin de détecter l'activité du cerveau et s'en servir pour commander des robots ou des claviers virtuels par exemple [Millán 2004].

Il est possible de détecter plusieurs types d'activité cérébrale [Millán 2002] :

- Les potentiels évoqués qui sont générés en réponse à des stimuli externes,
- Les potentiels corticaux lents,
- Les variations dans les rythmes des signaux EEG,
- L'activité directe des neurones dans le cortex.

Ces dispositifs sont encore très expérimentaux et nécessitent une préparation lourde pour poser les électrodes sur la tête de l'utilisateur ce qui limite leur diffusion. De plus, ces systèmes nécessitent un apprentissage important (soit de l'utilisateur, soit du système, soit des deux) avant d'arriver à s'en servir et sont souvent sujets à des erreurs d'interprétation des signaux. Cependant ils peuvent avoir un grand intérêt pour des personnes handicapées.

## III.4 Synthèse

### III.4.1 Influence des dispositifs sur l'interaction

L'examen des différents dispositifs d'entrée/sortie nous montre que tout dispositif est composé :

- D'une partie physique présente dans le monde réel : ensemble de capteurs (fixés aux objets, à l'utilisateur ou à l'environnement) ou d'effecteurs (actionneurs physiques, haut-parleur, écrans, ...)
- D'une fonction de conversion de grandeurs du monde physique en valeurs exploitables par le monde numérique et vice/versa : cette fonction est mise en œuvre par des convertisseurs analogique/numérique et numérique/analogique, et par un ensemble d'électronique et/ou de logiciels de contrôle de bas niveau.
- D'une partie logicielle (bibliothèques) qui s'interface avec le reste de l'application informatique et qui peut fournir des niveaux d'abstraction plus ou moins élevés par rapport au dispositif (par exemple un localisateur peut fournir des coordonnées spatiales absolues en mm après calibration).

Quant à la partie physique elle peut être liée (ce qui correspond à la taxonomie de Wendy MacKay – Cf. § II.3.1) :

- A l'environnement : c'est par exemple le cas d'un écran classique, ou d'un microphone posé sur une table,
- A l'utilisateur : comme par exemple la souris, des gants numériques ou encore les casques de visualisation,
- Aux objets comme les capteurs du Flock of Birds ou les haut-parleurs du système Hydra [Buxton 1992].

Dans le cas où le dispositif est lié à l'utilisateur, le retour haptique que procure le dispositif est important dans le contrôle que l'utilisateur a du dispositif. Ainsi une souris peut être facilement lâchée et reprise à la même position, alors qu'une « souris virtuelle » (où une caméra associée à un logiciel de reconnaissance de forme suit l'index de l'utilisateur), n'offre pas la même présence dans le monde réel.

Cette constatation de l'influence potentielle de la partie physique d'un dispositif d'entrée/sortie nous conduit, dans le modèle IRVO (Cf. Chapitre VI), à modéliser les dispositifs d'entrée/sortie en deux parties :

- Une partie physique qui peut être perçue (principalement par la vue et le toucher) par l'utilisateur,
- Une partie qui représente la fonction de conversion réel/virtuel ou virtuel/réel proprement dite et que nous appelons *transducteur* dans le formalisme d'IRVO.

Cette distinction n'est utile que si la perception de la partie physique est réellement importante dans le cadre de l'interaction modélisée. Si ce n'est pas le cas, alors il n'est pas nécessaire de distinguer la partie réelle et le *transducteur* représente alors le dispositif d'entrée/sortie complet (Cf. VI.2.6.2).

### III.4.2 Problème du choix des dispositifs par les concepteurs

Les périphériques d'entrées et de sorties sont très nombreux et il est souvent difficile pour un concepteur d'application de choisir tel ou tel dispositif. Il n'existe aucun périphérique universel qui permettrait de mettre en œuvre toutes les techniques.

Si on analyse les applications existantes, on voit cependant se dégager quelques tendances :

- Si l'application envisagée doit être mobile, alors il est presque nécessaire d'utiliser des ordinateurs portables ou portés (*wearable computer*) couplés à des micro/casques audio si l'on souhaite du son et avec un casque de visualisation semi-transparent si on veut de la vision

augmentée. De plus, il est presque toujours nécessaire d'avoir un système de localisation qui oblige soit de rajouter une caméra (pour les systèmes basés sur la vision artificielle) soit un dispositif *ad hoc* comme un GPS ou un localisateur magnétique. Le résultat est que le pauvre utilisateur se retrouve harnaché comme un astronaute du temps de la mission Apollo, mais comme lui ne profite pas de la faible pesanteur lunaire, le poids de l'équipement compromet sérieusement l'utilisabilité du dispositif. De plus, il ne faut pas craindre le ridicule pour oser se promener dans la rue ce qui fait que les expériences sont en général limitées à un campus universitaire comme Touring Machine (§ II.2.8.3) ou TROC (§ II.2.9.3) !

- Si l'application est fixe, alors la tendance est à dégager l'utilisateur de tout dispositif intrusif et de mettre les périphériques dans l'environnement par exemple en utilisant des vidéoprojecteurs et des caméras (utilisant la vision artificielle) ou directement sur ou dans les objets manipulés. Si l'on cherche une très haute intégration des images réelles et virtuelles cette technique présente quelques limites par rapport aux casques. Par contre, il est facile de mettre en place des systèmes qui permettent un travail à plusieurs et ainsi supporter le travail collaboratif en local. Pour le son, s'il est facile de mettre un haut-parleur fixe, mettre un micro d'ambiance pose de gros problèmes de qualité de prise de son, surtout si on souhaite faire de la reconnaissance vocale. Il faut également penser au retour du son du haut-parleur dans le micro ce qui, dans le cas typique de système d'audioconférences, conduit à des bouclages de type effet Larsen ou écho. Dans ce cas, le combiné micro/casque reste la solution la plus simple sinon il faut installer des dispositifs supprimeurs d'écho qui sont loin d'être satisfaisants pour les utilisateurs car ils ne permettent pas vraiment le *full-duplex* et reproduisent un fonctionnement à l'alternat.

Une troisième voie serait d'avoir des utilisateurs sans équipements portés mais tout en permettant la mobilité. Dans ce cas il faut véritablement rendre l'environnement intelligent et c'est ce que propose les domaines de l'informatique pervasive et de l'intelligence ambiante (AmI=Ambient Intelligence), mais cela nous entraîne au-delà du cadre de ce mémoire.

Pour aider les concepteurs à explorer les différentes solutions, le modèle IRVO, en modélisant les dispositifs d'entrée/sortie sous la forme de transducteurs, fournit un support à la réflexion et à l'étude de solutions alternatives.

### III.5 Bilan du chapitre

Dans ce chapitre nous avons étudié les différents dispositifs d'entrée/sortie existants et qui sont utilisés dans les systèmes de Réalité Mixte. Ils sont nombreux et leurs caractéristiques peuvent fortement influencer les capacités du système de Réalité Mixte à être réellement utilisable par les utilisateurs. Nous avons exploité ces résultats pour le choix des dispositifs d'entrée/sortie dans le projet ACCORéA (Cf. Chapitre VIII).

Cette étude nous a servi également à dégager le concept de *transducteur* (une des entités du modèle IRVO) comme le composant interne au périphérique qui réalise le passage monde réel → monde virtuel (ou l'inverse) et que nous séparons de la partie physique du dispositif dont les caractéristiques peuvent grandement influencer l'utilisabilité du dispositif selon le contexte dans lequel il est utilisé.



# Chapitre IV Modèles pour l'Interaction Homme-Machine

*I have always wished that my computer would be as easy to use as my telephone.  
My wish has come true. I no longer know how to use my telephone.*  
Bjarne Stroustrup

La conception et la réalisation des interfaces sont une tâche complexe et les modèles par leur décomposition en concepts permettent de réduire et maîtriser cette complexité. Les modèles sont la base même d'une approche analytique des systèmes.

Les modèles en IHM se décomposent en deux grandes familles d'après [Beaudouin-Lafon 2000a] :

- Les modèles d'interaction qui sont un ensemble de principes, règles et propriétés qui guident la conception d'une interface. Ils décrivent comment combiner les techniques d'interaction d'une manière significative et cohérente et définissent le « look and feel » de l'interaction d'un point de vue de l'utilisateur. Par exemple la manipulation directe est un modèle générique d'interaction.
- Les modèles architecturaux qui décrivent les éléments fonctionnels dans l'implémentation de l'interface et leurs relations. Les environnements de développement d'interfaces ont généré une grande variété de modèle d'implémentation comme le célèbre MVC.

Les modèles d'interaction sont destinés à la conception des interfaces tandis que les modèles architecturaux sont destinés au développement de l'interface. L'approche basée sur les modèles (MBA = Model-Based Approach) cherche à combler l'espace entre les modèles d'interaction et les modèles architecturaux. Nous les aborderons spécifiquement au Chapitre VII.

Le plan de ce chapitre est donc structuré en suivant cette logique : nous présentons tout d'abord les différents modèles d'interaction avant de présenter les modèles d'architecture. Pour chaque modèle, nous cherchons à évaluer la portée du modèle vis-à-vis de notre domaine d'étude, la Réalité Mixte.

Nous concluons ce chapitre par une synthèse de ces différents modèles qui montrent les limitations que nous y voyons et qui nous conduisent à proposer notre propre modèle au Chapitre VI.

## IV.1 Modèles d'interaction

Nous abordons dans cette section les modèles d'interaction qui sont destinés à la conception des interfaces.

Nous commencerons par la théorie de l'action de Norman, modèle issu de la psychologie qui a eu une influence considérable sur les travaux en IHM depuis une vingtaine d'année. Cette théorie n'est pas un modèle d'interaction mais elle a servi de cadre conceptuel au modèle de la manipulation directe.

Ensuite nous examinons des modèles qui vont au-delà des interfaces WIMP mais sans prendre en compte explicitement la Réalité Mixte : il s'agit du modèle de l'interaction instrumentale de Michel Beaudouin-Lafon et du modèle DPI. Nous examinerons également la théorie de l'activité médiatisée par l'instrument de Pierre Rabardel.

Puis nous considérons des modèles qui prennent explicitement en compte des objets réels comme moyen d'interaction : le modèle MCRpd pour les interfaces tangibles et les modèles de l'interaction physique de Sato.

Nous finissons cette revue par un « modèle » (même si son auteur n'emploie jamais ce mot) dédié à la Réalité Mixte : ASUR et ses différentes versions. Nous verrons aussi au chapitre suivant un autre formalisme permettant l'écriture graphique de scénarios pour les systèmes mixtes collaboratifs et mobiles (Cf. § V.2.1).

Pour chaque modèle, après une présentation générale, nous examinons sa portée et son utilisation potentielle dans le cadre de la Réalité Mixte.

### IV.1.1 Théorie de l'action et manipulation directe

Un modèle de l'interaction venant de la psychologie a été proposé par D. A. Norman [Vacherand-Revel 1992, Norman 1986] et qui a eu un grand retentissement dans le milieu de l'IHM. La théorie de l'action propose de modéliser le processus de réalisation et d'évaluation d'une action en sept phases d'activités de l'utilisateur (Figure 25) :

- Etablissement du but, représentation mentale de l'objectif
- Formation d'une intention, c'est la décision d'agir en vue d'atteindre le but
- Spécification d'une séquence d'action, représentation psychologique des actions qui sont à exécuter sur le système
- Exécution d'une action
- Perception de l'état du système
- Interprétation de l'état du système effectuée sur l'état perçu en fonction des variables psychologiques impliquées dans le but
- Evaluation de l'état du système du point de vue des buts et des intentions. Cela entraîne souvent un nouvel ensemble de buts et d'intentions

Il faut bien préciser que les phases ne sont pas forcément séquentielles : certaines peuvent être sautées, d'autres répétées.

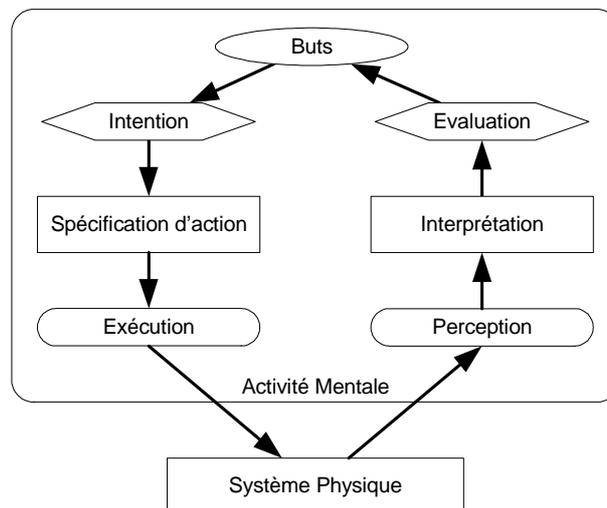


Figure 25 : les sept étapes de l'activité de l'utilisateur (d'après [Norman 1986])

L'interface introduit une distance qui est le résultat de la divergence entre les variables psychologiques de l'utilisateur et l'état réel du système informatique. Cette divergence est représentée par Norman comme deux gouffres qui doivent être comblés (Figure 26) :

- **Le gouffre de l'exécution** représente la distance mentale que l'utilisateur doit parcourir pour passer de ses intentions à l'exécution des actions.

- **Le gouffre de l'évaluation** représente le parcours inverse à celui du gouffre d'exécution. Les informations internes du système traversent l'interface de sortie qui sont perçues, interprétées et évaluées par l'utilisateur.

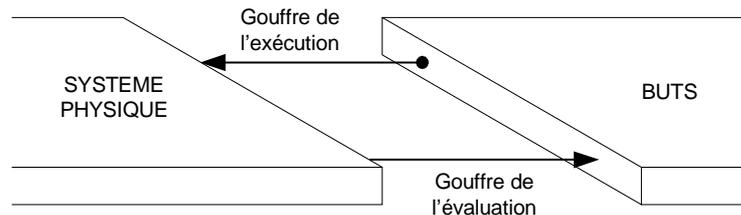


Figure 26 : les deux gouffres de l'exécution et de l'évaluation (d'après [Norman 1986])

Le concepteur du système peut combler les gouffres en partant du côté du système en construisant des interfaces en entrée et en sortie qui correspondent mieux aux besoins de l'utilisateur. L'utilisateur comble les gouffres en créant des plans, des séquences d'actions et des interprétations qui sont plus proches des formes requises par le système physique.

Le concept de **manipulation directe** (qui a été introduit par Ben Schneiderman) est une solution pour combler les gouffres du côté du système. En se rapprochant ainsi des buts de l'utilisateur, ces interfaces réduisent les efforts cognitifs que font les utilisateurs pour atteindre leurs buts d'après Hutchins *et al.* [Hutchins 1986]. Ces auteurs proposent de considérer deux aspects de la notion de « caractère direct » d'une interface :

- **La distance** : plus l'interface franchit une partie des gouffres, moins grande est la distance à combler par les efforts cognitifs de l'utilisateur et plus directe est la sensation résultant de l'interaction. L'interaction étant basée sur un *langage d'interface*, ils distinguent plus particulièrement :
  - La *distance sémantique* qui caractérise la relation entre les intentions de l'utilisateur et la signification des expressions utilisées dans le langage d'interface en entrée et en sortie.
  - La *distance articulatoire* qui caractérise la relation entre la forme physique d'une expression dans le langage d'interface et sa signification.
- **L'engagement direct** : cela caractérise la sensation qu'a l'utilisateur d'agir directement sur les objets du domaine de la tâche en construisant l'interface sur une métaphore du monde modélisé à la différence des interfaces construites sur une métaphore du dialogue ou l'utilisateur est en contact avec des structures linguistiques qui font références aux objets qui le concernent mais qui ne sont pas les objets eux-mêmes.

Un concept important que Norman a développé dans son ouvrage « The Design of Everyday Things » (1989) est celui d'**affordance** qui a été initialement introduit par Gibson en psychologie écologique. L'*affordance* se réfère aux propriétés réelles ou perçues qui déterminent comment un objet peut potentiellement être utilisé. Ainsi une porte *suggère* qu'on l'ouvre, une chaise *invite* à s'asseoir. Dans le contexte des interfaces homme-machine ce concept a été particulièrement étudié surtout sous l'angle des **affordances perçues** [Gaver 1991] : « il faut que les attributs physiques de l'objet soient compatibles avec ceux de l'acteur, que l'information sur ces attributs soit disponible sous une forme compatible avec le système perceptif, et que ces attributs et les actions qu'ils suggèrent soient significatifs pour une culture et un observateur ». L'intérêt principal de l'utilisation d'affordances dans une interface utilisateur est qu'elle devient plus facile à apprendre à utiliser ce qui tend à la rendre « transparente » pour l'utilisateur.

#### *Intérêt de la théorie de l'action pour la Réalité Mixte*

**La théorie de l'action** se place sur le plan de la modélisation de l'utilisateur dans le cadre de son interaction avec l'ordinateur. Ce modèle reste tout à fait valable dans le cadre de la Réalité Mixte,

toutefois il s'agit d'un modèle conceptuel qui ne peut pas être directement utilisé dans la conception.

L'un des principes fondateurs de la Réalité Mixte c'est la réintroduction des artefacts réels dans l'interaction avec l'ordinateur (« Back to the Real World » [Wellner 1993a]). Il serait erroné de penser que si l'on peut manipuler directement ces objets alors on obtient une interface à *manipulation directe* au sens propre et que les gouffres de l'interaction sont comblés. En effet, utiliser des objets réels couvre seulement l'aspect de l'engagement direct : il n'y a plus besoin de métaphore du monde réel puisqu'on manipule le monde réel lui-même. Par contre les distances articulatoire et sémantique peuvent être tout aussi importantes que dans une interface informatique classique.

De plus, les objets réels ne remplacent pas les objets virtuels mais viennent en complément. On se retrouve finalement avec un problème double de manipulation d'objets réels d'une part et de manipulations d'objets virtuels d'autre part, chaque partie présentant ses propres gouffres de l'exécution et de l'évaluation.

Le concept d'affordance garde également tout son sens en Réalité Mixte et génère de nouvelles interrogations. Les objets réels manipulés suggèrent des actions possibles mais ces actions sont-elles bien mises en valeur par le système ? Et surtout si l'action physique déclenche une action informatique, cette action informatique est-elle bien « compatible » avec l'affordance perçue de l'objet physique ?

### IV.1.2 Interaction instrumentale

Michel Beaudouin-Lafon propose un modèle pour les interfaces Post-WIMP qu'il appelle « interaction instrumentale » [Beaudouin-Lafon 2000a]. Les objets de la tâche sont appelés objets du domaine et sont manipulés avec des artefacts informatiques appelés instruments d'interaction. Ces instruments sont des transducteurs entre l'utilisateur et les objets du domaine : l'utilisateur agit sur l'instrument qui transforme les actions de l'utilisateur en commandes. L'utilisateur contrôle ses actions grâce à la réaction de l'instrument et des objets manipulés. Les instruments retournent également des informations (*feedback*) quand la commande est exécutée sur l'objet cible (Figure 27).

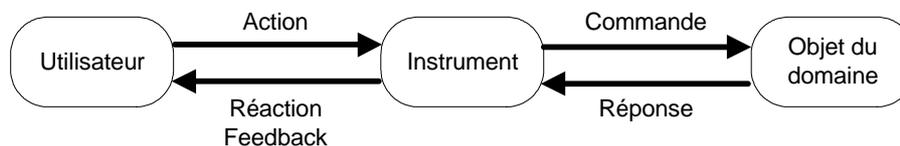


Figure 27 : Modèle de l'interaction instrumentale (adapté de [Beaudouin-Lafon 2000a])

Un instrument décompose l'interaction en 2 niveaux :

- L'interaction entre l'utilisateur et l'instrument, définie comme l'action physique de l'utilisateur sur l'instrument et la réaction de l'instrument,
- L'interaction entre l'instrument et l'objet du domaine, définie comme la commande envoyée à l'objet et la réponse de l'objet que l'instrument peut transformer en feedback à l'utilisateur.

L'instrument est composé de deux parties :

- Une partie physique, le périphérique d'entrée,
- Une partie logique, la représentation de l'instrument dans le logiciel et sur l'écran.

Les instruments sont activés lorsqu'ils sont sous le contrôle de l'utilisateur, c'est-à-dire lorsque la partie physique a été associée à la partie logique. En général, l'instrument est activé en positionnant le pointeur de la souris au-dessus ou en cliquant au préalable dans une palette d'outils et reste activé jusqu'à ce qu'un autre instrument soit activé.

Trois propriétés sont associées aux instruments :

- Le degré d'indirection mesure les décalages spatial et temporel générés par un instrument :

- Le décalage spatial est la distance à l'écran entre la partie logique de l'instrument et l'objet sur lequel il agit : elle peut être très faible (comme pour des poignées de sélection), grande voire arbitraire (comme pour une boîte de dialogue).
- Le décalage temporel est la différence de temps entre l'action physique sur l'instrument et la réponse de l'objet. Un objet peut répondre en temps réel (comme le défilement d'un texte lorsqu'on clique sur une barre de défilement) ou seulement à la fin de l'action (par exemple quand on clique sur le bouton 'OK' d'une boîte de dialogue).
- Le degré d'intégration mesure le rapport entre les degrés de liberté que fournit la partie logique de l'instrument et les degrés de liberté capturés par le périphérique d'entrée. Ce rapport peut être inférieur à 1 (exemple une barre de défilement manipulé par une souris 2D) ou supérieur à 1 (contrôle de 3 angles de rotation avec une souris 2D).
- Le degré de compatibilité mesure la similarité entre les actions physiques de l'utilisateur sur l'instrument et la réponse de l'objet. Il peut être élevé (comme dans le cas du déplacement d'un objet avec la souris) ou faible (défilement du texte vers le haut lorsqu'on déplace le curseur d'une barre de défilement vers le bas).

#### *Intérêt de ce modèle pour la Réalité Mixte*

**L'interaction instrumentale** a pour objectif la conception d'interfaces avancées et fait intervenir explicitement l'utilisateur et sa relation avec le système informatique. Ce modèle introduit la notion d'instrument qui comporte le périphérique physique et la partie logicielle représentée à l'écran. De ce point de vue, ce modèle prend partiellement en compte l'interaction avec le monde réel mais ne considère toutefois pas le cas où les objets du domaine sont réels, ce qui ne permet pas de l'utiliser comme modèle pour la réalité mixte. Cependant, la notion d'instrument est très pertinente et nous verrons que dans notre modèle IRVO, la notion d'outil est très proche.

### **IV.1.3 Modèle DPI (Document, Presentations, Instruments)**

Ce modèle proposé Michel Beaudouin-Lafon et son équipe peut être vu comme l'intégration de l'interaction instrumentale (Cf. ci-dessus) avec une approche centrée sur les documents [Beaudoux 2001]. Il est donc composé de deux métaphores :

- La métaphore de l'instrument exposée ci-dessus ; les auteurs distinguent plus précisément :
  - Des instruments directs qui permettent d'agir directement sur les documents.
  - Des instruments indirects : l'utilisateur agit directement sur l'instrument mais indirectement sur le document.
  - Des instruments de perception : l'utilisateur agit directement sur l'instrument mais n'agit pas sur le document.
- La métaphore du document qui présente 2 facettes :
  - La persistance : un document fournit un support permanent pour son contenu.
  - La présentation : un document a une apparence concrète.

Dans le cas des documents physiques, la persistance et la présentation sont naturellement couplés mais sont séparés dans les documents électroniques : le système de fichiers gère la persistance et les périphériques de sortie, la présentation.

DPI est basé sur une abstraction de présentations multiples qui permet de voir et d'éditer un document à travers n'importe laquelle de ses présentations, ce qui introduit comme pré requis fondamental que les résultats des éditions doivent être visualisés dans toutes les présentations.

Le modèle conceptuel de DPI (Figure 28) est rapproché d'un diagramme action-perception basé sur la théorie de l'action de Norman (Cf. § IV.1.1).

Un modèle fonctionnel destiné au concepteur des interfaces est également proposé. Il réifie les relations du modèle conceptuel en primitives :

- Primitives de persistance : ce sont les propriétés du document et de ses présentations.
- Primitives d'action : elles sont vues comme des services qui sont produits et consommés tout au long de la chaîne action-perception.
- Primitives de perception : ce sont les représentations qui sont la perception par l'utilisateur des propriétés du document.

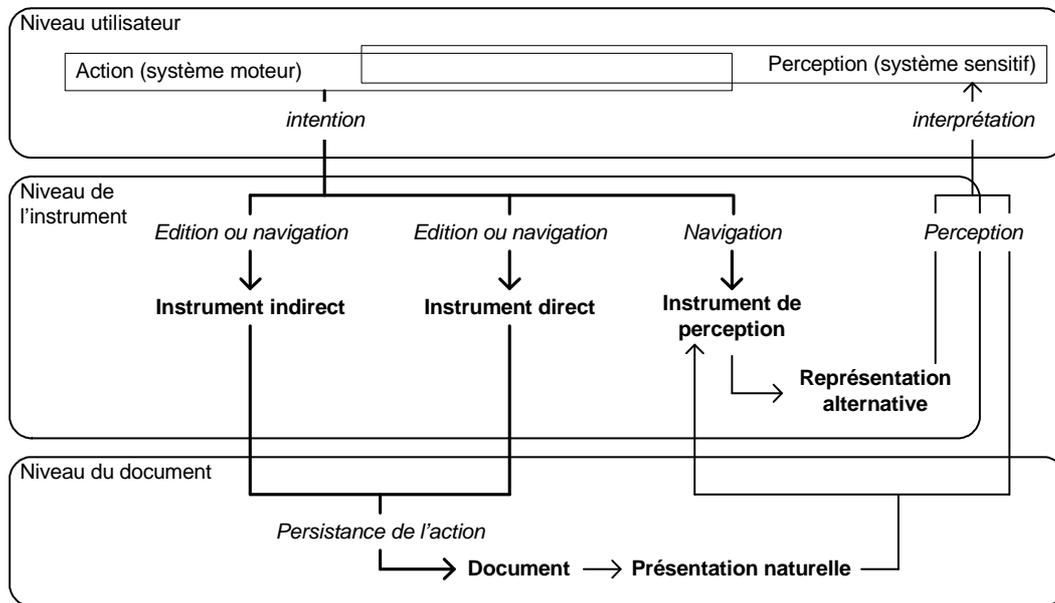


Figure 28 : Modèle conceptuel de DPI (d'après [Beaudoux 2001]).

#### Intérêt de ce modèle pour la Réalité Mixte

Le **modèle DPI**, intègre l'interaction instrumentale avec une approche centrée sur les documents. Ce modèle présente donc les mêmes avantages et inconvénients vis-à-vis de la Réalité Mixte que l'interaction instrumentale, en particulier que seuls des documents numériques sont considérés. En s'appuyant sur la théorie de l'action de Norman, son modèle conceptuel précise le rôle de l'instrument dans une boucle action-perception passant par l'utilisation d'un instrument et d'un document. Ce modèle a été une source d'inspiration forte pour notre propre modèle IRVO (Cf. Chapitre VI) qui présente donc de grandes similarités dans son approche.

#### IV.1.4 Théorie de l'activité médiatisée par l'instrument

Pierre Rabardel propose une « théorie de l'activité médiatisée par l'instrument et de genèse de l'instrument » [Rabardel 2003a, 2003b]. Dans cette approche le sens d'*instrument* est notablement différent car il s'agit d'une extension de la notion d'*outil* au sens de la théorie de la médiation par l'outil d'Engeström issue de la théorie de l'Activité de Leontiev [Bourguin 2000, chap. 3], [Fjeld 2002a].

La théorie de l'Activité postule que toute action est liée à l'environnement matériel et social dans lequel elle se déroule. Selon ce point de vue, les actions sont socialement et physiquement situées, et les circonstances (notamment leur perception par le sujet) conditionnent l'interprétation de l'action. Celle-ci est alors considérée comme une réponse spécifique à cet ensemble de circonstances qui correspondent à des ressources et des contraintes liées à l'environnement.

Le modèle structurel d'Engeström, issu de la théorie de l'Activité [Hasu 2000], propose une modélisation simple du concept d'activité en exprimant la médiation qui existe entre le *sujet* et l'*objet* de l'activité. Cette médiation est représentée par le concept d'*outil*. L'outil influence fortement l'activité, en supportant et en complétant les capacités humaines. « *Outil* » est pris ici dans un sens très général pour englober tout ce qui est utilisé dans le processus de transformation et donc

recouvre aussi bien des outils matériels que des outils intellectuels. Il faut noter que l'*objet* de l'activité correspond à un besoin ou un désir et qui motive l'existence de cette activité.

Dans l'approche de Rabardel, l'*outil* médiateur de l'activité est reconsidéré en *instrument*. La principale différence entre un outil et un instrument est que ce dernier se définit à la fois par rapport à l'objet et au sujet de l'activité. L'instrument se compose donc :

- de l'artefact matériel ou symbolique, produit par d'autres ou par le sujet lui-même,
- des modalités d'usage de cet artefact par le sujet qui sont les *schèmes d'utilisation*.

La Figure 29 présente son modèle de l'activité médiatisé par l'instrument. On peut voir plusieurs types de médiation :

- La médiation entre le sujet et l'objet de l'activité. Elle est similaire à la médiation par l'outil de Engeström. Pierre Rabardel propose de distinguer 2 formes de celle-ci :
  - La *médiation épistémique* dont le but est de connaître l'objet (ses propriétés, ses évolutions, etc.)
  - La *médiation pragmatique* qui concerne les actions sur l'objet (transformation, etc.)
- Les médiations avec les autres sujets, ce qu'il appelle *médiations interpersonnelles*. Elles peuvent aussi être épistémiques ou pragmatiques.
- La médiation du sujet avec lui-même, ce qu'il appelle *médiation réflexive*.

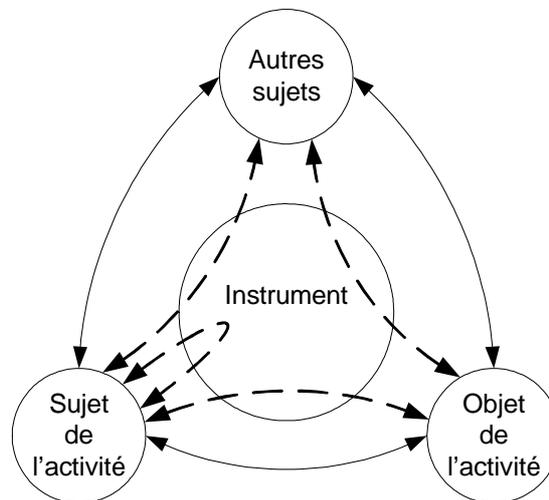


Figure 29 : Modèle de l'activité médiatisée par l'instrument (d'après [Rabardel 2003b])

L'instrument est donc le résultat d'une élaboration progressive qui correspond à l'appropriation de l'outil par l'utilisateur ce que Pierre Rabardel appelle *genèse instrumentale*. Il distingue :

- L'instrumentalisation, qui est en rapport avec l'émergence et l'évolution de l'artefact, ce qui prolonge le concept d'outil.
- L'instrumentation, qui concerne l'émergence et l'évolution des schèmes d'utilisation, ainsi que l'assimilation d'artéfacts nouveaux à des schèmes existants.

#### *Intérêt de ce modèle pour la Réalité Mixte*

L'approche instrumentale de Pierre Rabardel présente un cadre riche et le concept d'instrument séduisant. Cependant, ce modèle provenant des Sciences Humaines est très conceptuel et difficile à appréhender pour des concepteurs de systèmes informatiques. Il serait très intéressant de creuser les relations qui existent entre le concept d'instrument de Rabardel et celui de Beaudouin-Lafon mais cela sort du cadre de cette thèse.

### IV.1.5 Modèle MCRpd

Les interfaces tangibles ont été proposées par Ishii et Ullmer [Ishii 1997, Ullmer 2000]. Comme nous l'avons expliqué dans le paragraphe II.1.6, nous les voyons comme une partie de la Réalité Mixte. Les mêmes auteurs proposent un modèle d'interaction pour décrire les interfaces tangibles, appelé MCRpd, acronyme de *model-control-representation (physical and digital)*, et conçu comme une extension du modèle MVC (Figure 30).

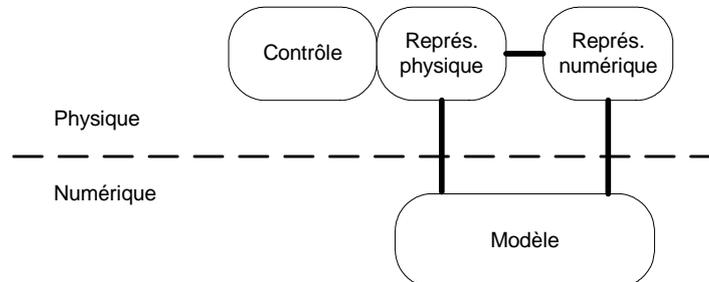


Figure 30 : Modèle MCRpd (d'après [Ullmer 2000])

Les éléments Modèle (M) et Contrôle (C) sont les mêmes que dans MVC (Cf. § IV.2.2.1, ci-après). L'élément Vue (V) de MVC est décomposé en deux parties :

- Les représentations physiques (rep-p) pour les artefacts constituant les éléments physiques des interfaces tangibles,
- Les représentations numériques (rep-d) pour les éléments des interfaces médiatisés par l'ordinateur et qui n'ont pas de parties physiques.

Ce modèle met en avant trois caractéristiques fondamentales des interfaces tangibles :

- La représentation physique est couplée informatiquement à l'information numérique sous-jacente (lien rep-p – modèle).
- Les représentations physiques incorporent des mécanismes pour le contrôle interactif (lien contrôle – rep-p).
- Les représentations physiques sont couplées de manière perceptuelle aux représentations numériques (lien rep-p – rep-d).

#### *Intérêt de ce modèle pour la Réalité Mixte*

Ce modèle est par conception limité aux interfaces tangibles, ce qui couvre tout de même une partie de la Virtualité Augmentée (Cf. § II.1.6). Bien que basé sur MVC (qui est un modèle d'architecture des logiciels et même un framework !), le modèle, tel qu'il est présenté par ses auteurs, est très conceptuel et ils ne donnent pas d'éléments pour le rendre plus opérationnel. En effet, comment le contrôle (qui est dans le modèle MVC d'origine une classe Smalltalk) peut passer dans le monde réel sans remettre en cause le concept sous-jacent ? A contrario, ce modèle partage les mêmes reproches que le modèle MVC, à savoir que la communication entre les différentes composantes est directe ce qui conduit à une mauvaise structuration du logiciel comparativement aux modèles PAC (Cf. § IV.2.2.2) et AMF (Cf. § IV.2.2.3) qui formalisent cette communication à travers la facette « contrôle ».

### IV.1.6 Modèles de l'Interaction physique de Sato

Keiichi Sato *et al.* [Sato 2000] proposent de modéliser les systèmes interactifs qui mélangent des entités du monde physique (qu'ils nomment « espace physique ») et du monde numérique (qu'ils appellent « espace des medias » et correspond à ce que nous appelons « monde virtuel »). A cause de la complexité générée par l'intégration de ces deux espaces, ils proposent de se placer dans une approche centrée sur l'utilisateur pour produire des systèmes qui satisfassent les besoins des utilisateurs tant sur les plans cognitifs, sociaux, culturels que sur le plan fonctionnel.

Ils proposent de modéliser 8 cas d'interaction entre un utilisateur et une interface physique en relation avec les objets, l'environnement et l'espace des médias :

- **Modèle H-P-O** : il s'agit du cas le plus basique de relation entre un objet physique (O) et un utilisateur (H). L'interface physique (P) est par construction incorporée à l'objet lui-même (par exemple un interrupteur pour allumer la lumière).
- **Modèle H-P-O-M** : par rapport au modèle H-P-O les objets incorporent des fonctionnalités informatiques : ainsi l'espace de médias (M) est derrière l'objet et l'augmente ou permet d'accéder à des informations à travers l'objet (exemple d'une feuille de papier avec un tag RFID).
- **Modèle H-P-M-MO** : Dans le cas d'objets multiples (MO) à contrôler en même temps, une forme de support médiatisé (M) est nécessaire pour transmettre le contrôle aux différentes unités qui forment un groupe (exemple d'un cluster de machines pour remplacer une puissante machine unique).
- **Modèle H-P-M-E** : C'est le même cas que ci-dessus mais au lieu de contrôler des objets, il s'agit de contrôler l'environnement (E) comme les lumières, le son ou l'air.
- **Modèle H-E-M** : dans le cas où les fonctions informatique et de réseau sont incluses dans des artefacts ordinaires, l'interaction de l'utilisateur a directement lieu avec ces artefacts sans interface spécialisée. Ce cas, et les deux précédents, recouvrent les cas d'informatique ubiquitaire et pervasive.
- **Modèle H-P-M** : dans ce cas l'interface physique pilote directement les entités de l'espace des médias de manière plus directe que la souris ou le clavier. Ceci est clairement le modèle correspondant aux interfaces tangibles.
- **Modèle H-P-M-RO** : Il s'agit d'interagir avec des objets à distance (RO = Remote Objects) par l'intermédiaire de l'espace des médias. Il peut s'agir par exemple de cas de chirurgie à distance.
- **Modèle H-P-M-P-H** : c'est un cas de communication personne-personne médiatisée. les jeux multi-utilisateurs sont des exemples de ce modèle.

#### *Intérêt de ce modèle pour la Réalité Mixte*

Ce modèle de l'interaction physique de Sato (ou plutôt les huit modèles d'interaction) sont particulièrement pertinents pour distinguer les différents cas d'interaction avec le monde physique et des données numériques. Il n'est cependant guère opérationnel car il n'y a pas de formalisme associé permettant de modéliser une application concrète. Il apparaît en pratique plus comme une taxonomie des différents cas possibles d'interaction avec le monde physique.

### **IV.1.7 Modèle spécifique pour les interfaces de Réalité Mixte**

Dans cette section nous examinons le formalisme ASUR, l'un des formalismes dédié à la modélisation de l'interaction pour la Réalité Mixte. Le second formalisme, sera étudié dans le chapitre suivant car il couvre également le travail collaboratif.

ASUR a connu trois versions que nous présentons successivement avant d'en faire une analyse critique.

#### **IV.1.7.1 Formalisme ASUR**

Le formalisme ASUR est proposé par Dubois [Dubois 2001a, Dubois 2002c, Dubois 2003a]. Il était appelé auparavant OP-a-S (Objet-Personne-adaptateur-Système) [Dubois 2000].

ASUR est un formalisme permettant de décrire les entités intervenant au cours de l'utilisation d'un système de Réalité Mixte dans le contexte de la réalisation d'une tâche donnée.

ASUR comporte 6 composants :

- Le composant  $R_{task}$ , objet du monde réel, qui est l'objet de la tâche,
- Le composant  $R_{tool}$ , objet du monde réel, qui est utilisé comme outil pour réaliser la tâche,
- Le composant  $U$ , qui représente l'utilisateur du système
- Le composant  $S$  qui représente le système informatique (ordinateur)
- Les composants  $A$  qui sont des adaptateurs permettant le passage d'informations entre le monde réel ( $U$ ,  $R_{task}$ ,  $R_{tool}$ ) et le monde informatique ( $S$ ). Il en existe deux sortes :  $A_{in}$ , pour passer du monde réel au monde informatique et  $A_{out}$  pour passer du monde informatique au monde réel.

ASUR propose 3 types de liens entre les composants :

- Des échanges de données unidirectionnels entre 2 composants représentés par une flèche simple entre le composant émetteur et le composant récepteur,
- Des échanges d'énergie entre deux composants représentés par un trait simple (exemple d'un stylo qui écrit sur un livre),
- Des relations représentant la fusion technique de deux objets dans un même dispositif (par exemple fusion d'une caméra et d'un vidéoprojecteur).

La Figure 31 montre un exemple de modélisation de CASPER-v2 avec ASUR. Cette application de chirurgie augmentée a été présentée au paragraphe II.2.6.1. L'objet de la tâche ( $R_{task}$ ) est le patient subissant l'opération de ponction, le chirurgien étant l'utilisateur ( $U$ ). Ce dernier manipule un outil, l'aiguille de ponction ( $R_{tool}$ ), et la perçoit ce qui explique la double flèche entre ces 2 composants. La position de l'aiguille est capturée par le localisateur (adaptateur d'entrée,  $A_{in}$ ) pour fournir la position de l'aiguille au système. Le localisateur suit également les mouvements du casque de visualisation semi-transparent (adaptateur de sortie,  $A_{out}$ ) pour que le système puisse afficher sur celui-ci une image correctement ajustée qui se superpose à la vision du patient.

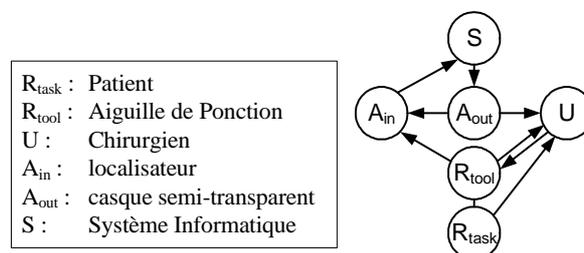


Figure 31 : Description de CASPER-V2 avec ASUR (d'après [Dubois 2001a]).

Les adaptateurs et les objets réels sont caractérisés par :

- Le **sens perceptif** (visuel, auditif, haptique, etc.) **ou actionnel** de l'utilisateur impliqué dans la perception de ce composant ou pour transmettre des données vers ce composant,
- Le **lieu de la perception ou d'action** : c'est l'endroit où l'utilisateur doit porter son attention pour émettre des données ou en percevoir.
- Le **niveau de partage** du composant : c'est la capacité d'un adaptateur ou d'un objet de partager les données avec plusieurs utilisateurs.

Les relations sont quant à elles caractérisées par :

- Le **langage** utilisé pour véhiculer une donnée. ASUR reprend ici les travaux de [Bernsen 1994] sur la multimodalité.
- Le **poids de la relation** qui exprime la pertinence et l'importance des données portées par la relation pour la réalisation de la tâche.
- Le **référentiel de présentation** adopté pour rendre perceptible les données. Le référentiel peut être associé à chaque composant ASUR :  $A_{in}$ ,  $A_{out}$ ,  $R_{tool}$ ,  $R_{task}$ ,  $U$  ou encore une donnée virtuelle ( $S$ ).

### IV.1.7.2 ASUR++

ASUR a été étendu aux systèmes mixtes et mobiles sous le nom de « ASUR++ » [Dubois 2002a, Dubois 2002b, Dubois 2003b]. Il a essentiellement été rajouté une nouvelle relation (notée par une flèche à double trait) entre 2 composants qui exprime une dépendance spatiale qui, si elle est satisfaite, déclenche d'autres actions. Par exemple, si un visiteur s'approche d'un tableau dans un musée augmenté, des informations supplémentaires lui sont fournies à ce moment-là.

Nous ne détaillerons pas plus ces extensions car elles concernent le support de la mobilité, domaine que nous avons choisi de ne pas aborder dans cette thèse.

### IV.1.7.3 ASUR 2004<sup>13</sup>

ASUR vient de connaître une évolution récente dans le cadre de travaux de l'Action Spécifique « systèmes mixtes » du CNRS (AS 153, RTP 16) qui a été présentée à la conférence IHM en septembre 2004 [Dubois 2004b].

Cette évolution tient essentiellement dans le raffinement du composant S qui représente le système informatique tout entier. Il est donc introduit de nouveaux types de composants :

- Le composant  $S_{Tool}$  qui représente une entité numérique dont l'activation entraîne une modification du comportement ou de l'apparence d'une autre entité numérique. Il représente les éléments numériques utiles à l'interaction en entrée (interacteur, pointeur, navigation).
- Le composant  $S_{Presentation}$  qui regroupe les objets du domaine, c'est-à-dire les concepts manipulés par le noyau fonctionnel et donc situés au cœur de l'interaction de l'utilisateur avec le système. Il existe deux types de ces composants :
  - Le composant  $S_{Object}$  qui est le symétrique du  $R_{Object}$  de ASUR. Il constitue l'objet de la tâche. Dans un éditeur de texte, il s'agira par exemple du document électronique à modifier.
  - Le composant  $S_{Info}$  représente un des objets du domaine, pertinent pour la tâche en cours mais ne constituant pas la raison d'être de la tâche. Il s'agit principalement de présenter à l'utilisateur des informations sur l'état interne du système. Il traduit par exemple une variable du domaine ou rend perceptible une entité numérique représentant une entité physique, etc.

Le composant S se trouve donc réduit au système d'exploitation, aux pilotes, aux capacités de stockage et au noyau fonctionnel de l'application : il devient le méta-composant S dans cette nouvelle version qui contient :

- D'une part les composants systèmes mentionnés ci-dessus et qui ne sont pas représentés graphiquement car il n'est pas pertinent de les détailler pour la conception de l'interaction d'un utilisateur avec un support mixte,
- Et d'autre part, les éléments numériques support de l'interaction en entrée ( $S_{Tool}$ ) et les objets du domaine ( $S_{Presentation}$ ) qui sont du type  $S_{Object}$  ou  $S_{Info}$ .

Une nouvelle relation est également ajoutée qui représente un lien entre une entité et sa représentation. Une entité du monde réelle peut représenter une entité du monde numérique : par exemple le cube des *mediaBlocks* [Ullmer 2000] représente une vidéo numérique.

Un exemple de la modélisation du système CASPER-v2 est représenté Figure 32. Par rapport au schéma de la Figure 31 on voit bien apparaître les entités numériques qui correspondent à la modélisation de l'aiguille (position et orientation) et au cône de la trajectoire idéale que doit suivre le chirurgien. Ces trois éléments sont rendus visibles au chirurgien à travers le casque de visualisation

<sup>13</sup> Le nom « ASUR 2004 » a été proposé par Emmanuel Dubois pendant sa présentation orale à la conférence IHM pour pouvoir aisément différencier cette version par rapport aux précédentes mais n'apparaît pas sur l'article dans les actes [Dubois 2004b].

semi transparent (HMD). Le reste du schéma est similaire même si la représentation graphique des entités a évoluée.

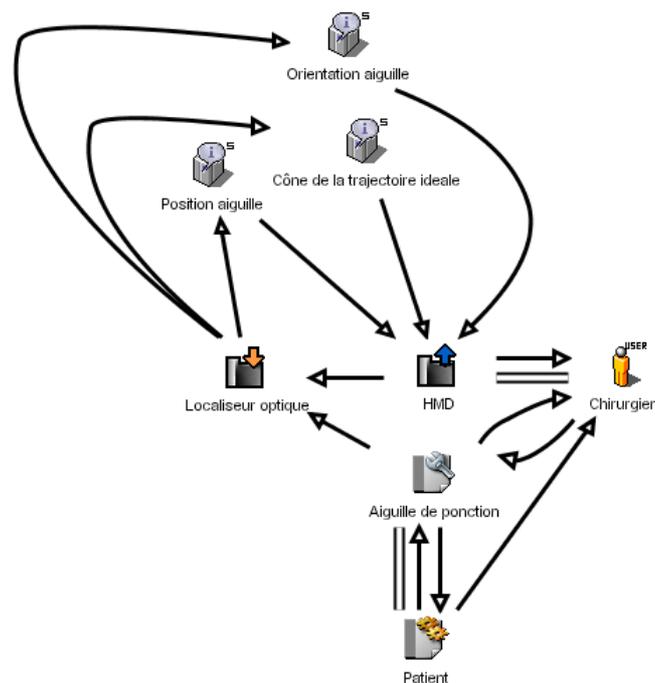


Figure 32 : Représentation de CASPER-v2 avec ASUR 2004 (d'après [Juras 2004]).

#### IV.1.7.4 Critique du modèle ASUR

Le modèle ASUR présente plusieurs intérêts qui ont été montrés par [Dubois 2001b] :

- Il permet de mettre en évidence les différentes entités numériques et physiques pour réaliser une tâche particulière, d'identifier les liens entre ces différentes entités et enfin de caractériser ces entités et ces liens.
- Cette caractérisation permet de raisonner sur l'utilisabilité du système mixte en cours de conception en vérifiant des propriétés ergonomiques.
- Il permet d'aider à la classification des systèmes. En caractérisant les entités du monde réel, il permet de rapidement identifier la classe de système : Réalité Augmentée ou Virtualité Augmentée.

Emmanuel Dubois identifie également trois limitations du modèle ASUR [Dubois 2001b] :

- Les deux types d'augmentation (exécution augmentée et évaluation augmentée, Cf. § II.3.5) ne sont pas distinguables dans les schémas ASUR,
- Les outils et objets virtuels ne sont pas représentés ; ainsi ASUR est d'un intérêt limité pour modéliser des applications de Virtualité Augmentée. Cette limitation n'existe plus depuis la version ASUR 2004 qui introduit justement des outils et objets virtuels.
- Un seul utilisateur est représenté. ASUR ne peut donc prendre en compte les systèmes collaboratifs.

Ces trois limitations ont justement constitué notre point de départ pour bâtir notre modèle IRVO (Cf. Chapitre VI).

## IV.2 Modèles d'architecture

D'après [Coutaz 2001] l'architecture d'un système informatique est un ensemble de structures comprenant chacune des composants, les propriétés extérieurement visibles de ces composants et les relations qu'ils entretiennent. Chaque composant (qui, selon les modèles, peut être un service, un module, un processus, un objet, un agent, etc...) possède donc des propriétés externes (des services fournis, des ressources requises, etc...) qui permettent l'interaction avec les autres composants.

Chaque modèle possède son point de vue sur le système, ce qui fait qu'un seul modèle ne peut couvrir tous les besoins. L'hétérogénéité est donc nécessaire pour englober les différentes approches mais au risque d'être la source d'incompatibilité entre les modélisations car elles ne sont pas orthogonales. On peut faire l'analogie avec les plans en architecture où chaque corps de métier (maçon, électricien, plombier, ...) a son propre plan adapté à ses besoins mais évidemment ils concourent tous à la construction d'un bâtiment unique.

Dans cette section nous allons examiner les modèles d'architecture classiques conçus pour les interfaces du type WIMP (Windows, Icon, Menus and Pointing devices). Ils sont particulièrement nombreux et nous ne citons que les principaux : les modèles en couche (Seeheim et Arch) et les modèles multi-agents (MVC, PAC et AMF, le modèle développé dans notre laboratoire). Ces deux approches sont d'ailleurs conciliables comme dans le modèle PAC-Amodeus et le rapprochement entre AMF et Arch.

Nous présenterons ensuite une adaptation du modèle PAC-Amodeus qui a été faite pour prendre en compte la Réalité Mixte, avant de faire une synthèse générale sur ces modèles.

### IV.2.1 Modèles en couche

Les modèles en couche sont anciens mais toujours d'actualité car ils fournissent un cadre général de référence par rapport auquel la plupart des autres modèles d'interaction se positionnent.

#### IV.2.1.1 Modèle « Seeheim »

Ce modèle a été établi par un « working group » de l'IFIP (International Federation for Information Processing) lors d'un workshop qui s'est tenu en Allemagne à Seeheim en 1983 et publié en 1985 par Pfaff [Reynolds 1997, van der Veer 2001].

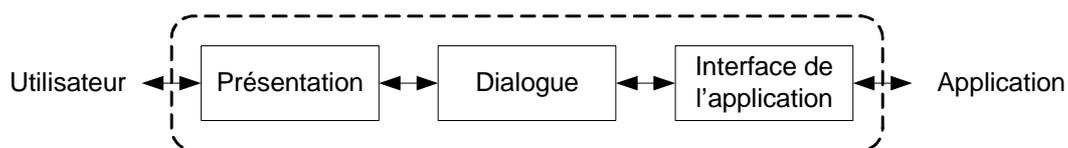


Figure 33 : Modèle Seeheim

Il définit trois couches (Figure 33) :

- Une couche « Présentation » qui définit les aspects lexicaux de l'interaction avec l'utilisateur, et en particulier l'organisation de l'écran et la disposition du clavier.
- Une couche « Dialogue » qui définit les aspects syntaxiques de l'interaction et qui est responsable des aspects dynamiques de l'application.
- Une couche « Interface de l'application » (Application Interface) qui est une couche d'adaptation du noyau non-interactif de l'application pour la conformer au style d'interaction envisagé.

Une variante de ce modèle est présentée par [Coutaz 2001] où est présenté un « Noyau Fonctionnel » à la droite du Dialogue, donc en englobant l'interface de l'application et l'application elle-même.

### IV.2.1.2 Modèle « Arch »

Le modèle Arch [UIMS 1992] hérite du modèle Seeheim et en constitue un raffinement. Il reprend le contrôleur de dialogue de Seeheim comme centre de l'arche (Figure 34). Du côté du « pilier » applicatif, il distingue 2 niveaux :

- Le « noyau fonctionnel » de l'application qui gère les données internes et leur rémanence.
- « L'adaptateur du noyau fonctionnel » qui a les fonctionnalités de l'interface d'application de Seeheim.

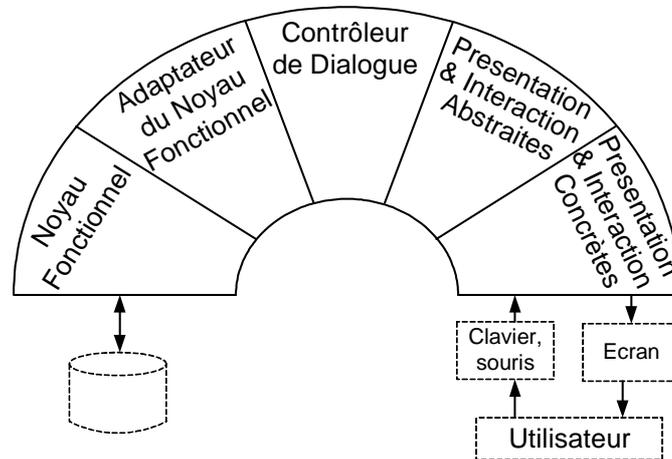


Figure 34 : Modèle « Arch » (adapté de [UIMS 1992])

Du côté du « pilier » interface il décompose l'interface utilisateur en 2 niveaux :

- Un niveau abstrait, comportant les objets de présentations et d'interaction indépendants de l'interface concrète et donc des dispositifs d'interaction avec l'utilisateur.
- Un niveau concret, lié à une « toolkit » précise gérant le dialogue avec l'utilisateur (et donc lié à des dispositifs d'entrée/sortie précis).

Dans le modèle original [UIMS 1992] ces 2 couches sont appelées « Presentation Component » et « Interaction Toolkit Component » ce qui peut porter à confusion en laissant sous-entendre qu'une couche est dédiée aux sorties (affichage à l'écran) alors que l'autre se consacre aux entrées (souris, clavier) ce qui est évidemment faux.

Ainsi, on rencontre dans la littérature d'autres termes ; par exemple [Coutaz 2001] a rebaptisé ces couches plus justement :

- Présentation logique,
- Présentation physique.

Mais le mot « physique » ne nous semble pas approprié car, dans le contexte de la Réalité Mixte l'usage de ce mot peut prêter à confusion avec les objets physiques. Nous préférons donc nommer explicitement ces deux couches (Cf. Figure 34) :

- Présentation et Interaction abstraites,
- Présentation et Interaction concrètes.

## IV.2.2 Modèles multi-agents

Afin de structurer de manière plus fine les IHM, de nombreux modèles multi-agents ont été proposés comme MVC (Model-View-Controller), PAC (Présentation-Abstraction-Contrôle), ou AMF (Agents multi-facettes).

Tous ces modèles se basent sur la notion d'agents logiciels qui communiquent entre eux. Un agent est un système de traitement de l'information qui se distingue par [Coutaz 2001] :

- Un jeu d'opérations,
- Des mécanismes d'entrées/sorties,
- Un état interne.

En IHM, il s'agit d'agents « simples » et non pas d'agents intelligents au sens de l'Intelligence Artificielle. Leur comportement est donc essentiellement dicté par les interactions de l'utilisateur (directement ou indirectement pas la communication avec les autres agents) : on parle donc généralement d'agents réactifs pour les distinguer des autres qui sont appelés agents cognitifs [Tarpin-Bernard 1997].

Ces agents sont généralement structurés en facettes fonctionnelles qui décomposent et structurent l'agent en définissant un vocabulaire d'éléments conceptuels (agents, facettes, communication par message, etc...) et une certaine sémantique précisant les règles de composition et de communications entre agents et les rôles que les agents ont dans le modèle.

#### IV.2.2.1 MVC

Le modèle MVC (Model-View-Controller) a été développé à l'origine comme modèle de programmation des interfaces pour le langage Smalltalk [Burbeck 1992, Coutaz 2001]. Il présente une décomposition d'une application en agents comportant 3 « facettes » :

- La facette « Contrôleur » gère les entrées provenant du clavier et de la souris,
- La facette « Vue » est chargée de l'affichage des données à l'écran (sous forme graphique et/ou textuelle)
- La facette « Modèle » gère le comportement et les données de l'application, répond aux demandes d'information sur son état (en général en provenance de la Vue) et répond aux demandes de modifications de l'état (en général en provenance du contrôleur).

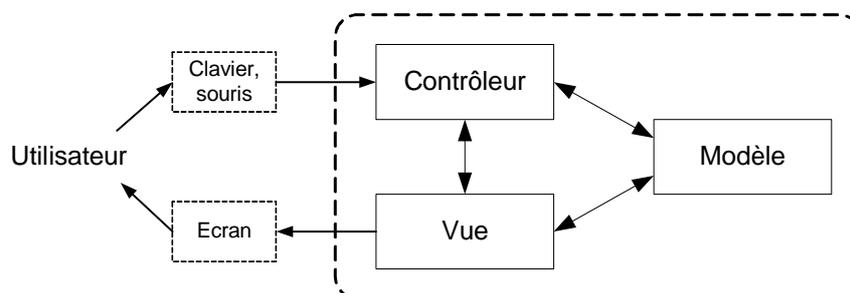


Figure 35 : Modèle MVC

#### IV.2.2.2 PAC

Le modèle PAC a été développé par Joëlle Coutaz [Coutaz 2001]. Il modélise la structure d'un logiciel interactif sous la forme d'une hiérarchie d'agents (Figure 36). Chaque agent est responsable d'un aspect spécifique d'une fonctionnalité de l'application et il possède 3 facettes :

- La facette Présentation (P) modélise le comportement visible de l'agent. Elle concerne à la fois les entrées et les sorties. La présentation interprète les événements résultant des actions physiques de l'utilisateur via des dispositifs d'entrée, et engendre via des dispositifs de sortie des actions physiques,
- La facette Abstraction (A) maintient le modèle de données sous-jacent à l'agent. Elle constitue le Noyau fonctionnel de l'agent au sens de Seeheim ou Arch,
- La facette Contrôle (C) a deux rôles :
  - Elle relie les facettes P et A de l'agent. Tout échange entre ces facettes A et P passe obligatoirement par la facette C ;
  - Elle gère la communication avec les autres agents PAC (parents ou enfants de la hiérarchie) par l'intermédiaire de la facette C de ces agents.

La facette Contrôle a donc des fonctions d'arbitrage (synchronisation et coordination entre agents mais aussi entre les facettes A et P de l'agent) et de traduction (transformation de formalismes entre A et P qui utilisent des systèmes de représentation différents).

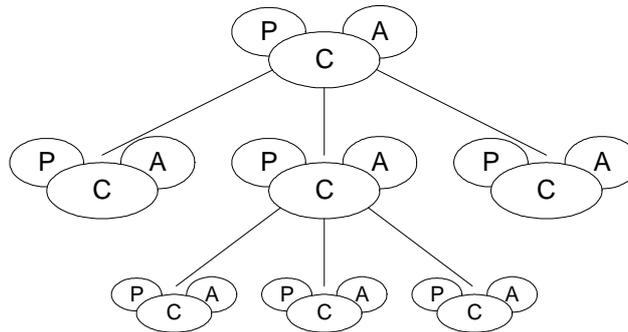


Figure 36 : Modèle PAC

### IV.2.2.3 AMF

AMF a été développé à l'École Centrale de Lyon par Kamel Ouadou [Ouadou 1994]. Ce modèle a été complété et étendu au travail collaboratif par Franck Tarpin-Bernard [Tarpin-Bernard 1997, Tarpin-Bernard 1999].

Il part sur une critique du modèle PAC :

- La décomposition Abstraction/Présentation est généralement insuffisante pour des applications complexes. Des fonctionnalités se retrouvent mélangées dans des composants trop macroscopiques alors qu'elles relèvent de thématiques différentes.
- La structure de la facette Contrôle est peu formalisée. Or cette facette est la clef de voûte des modèles d'architecture.

AMF reprend donc les mêmes facettes de base P, A, C mais permet d'en ajouter de nouvelles pour des besoins plus particuliers ce qui permet un découpage plus fin des agents interactifs et de répondre à la première critique. On peut avoir de nouvelles facettes provenant d'une décomposition fine de la facette contrôle, de la duplication des facettes classiques (par exemple, plusieurs facettes présentation), ou pour modéliser de nouveaux aspects des agents (par exemple, facettes « Aide » ou « Erreur », modélisation de l'utilisateur, gestion de la coopération, etc.) Chaque facette se voit donc attribué un rôle précis ce qui conduit à une meilleure modélisation. Cela permet également la constitution de « patterns » d'autant plus réutilisables qu'ils sont stéréotypés.

Pour répondre à la deuxième critique, AMF propose un formalisme complet pour modéliser la facette Contrôle et représenter les échanges entre les différentes facettes de l'agent et entre les différents agents. La facette Contrôle se trouve ainsi précisément modélisée dans son comportement. Il est donc facile de passer du modèle au logiciel lui-même en faisant l'instanciation de ce modèle dans le contrôleur du dialogue. Ceci est accompli par la notion de « moteur AMF » (Cf. § IV.2.3.2). Le formalisme graphique utilisé pour modéliser la facette contrôle peut ainsi être vu comme un véritable langage de programmation visuelle du contrôleur de dialogue de l'application elle-même.

Comme dans PAC, les agents sont organisés de manière hiérarchique selon des règles de composition. Au plus haut niveau se trouve un agent racine (appelé généralement « Application »). La composition peut s'exprimer soit sous forme d'emboîtement (Figure 37a), soit sous forme de graphes (Figure 37b) utile pour les schémas complexes.

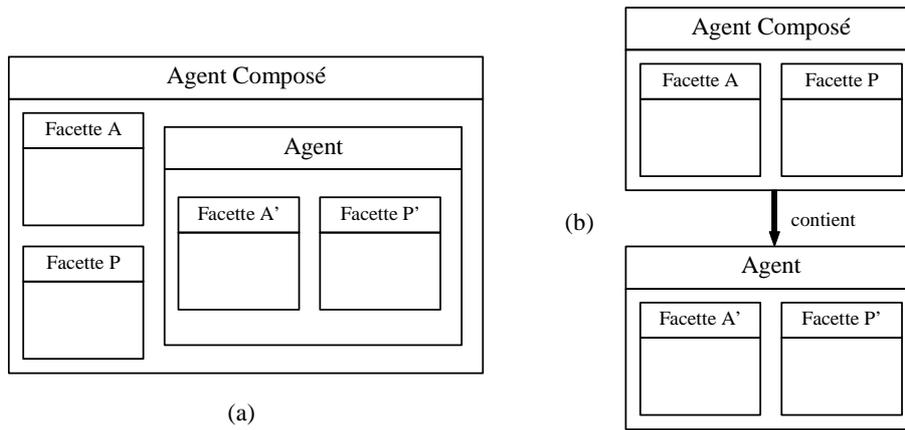


Figure 37 : Représentation de la composition d'agents AMF (d'après [Tarpin-Bernard 1997])

Modélisation du contrôle

Cette modélisation est basée sur 2 concepts :

- Les **ports de communication**, au niveau des facettes, représentent les services qui sont offerts par la facette (ports en entrée) et ceux qui sont nécessaires (ports en sortie). Un troisième type de port est proposé (port en entrée/sortie) qui correspond à une service offert dont l'activation se conclut par l'invocation d'un autre service distant (Figure 38).
- Les **administrateurs de contrôle**, au niveau de l'agent, connectent les ports de communication entre eux. Un administrateur de contrôle (Figure 39) joue trois rôles :
  - un **rôle de connexion** qui consiste à gérer les relations logiques pouvant exister entre les ports de communication qui lui sont attachés (sources ⇔ cibles) ;
  - un **rôle comportemental** qui exprime les règles d'activation de l'administrateur, c'est-à-dire sous quelles conditions et à quel moment les messages émis par les ports sources seront transmis aux ports cibles ;
  - un **rôle de traduction** qui consiste à transformer les messages émis par les ports sources en messages compréhensibles par les ports cibles.

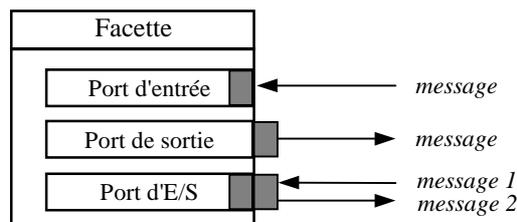


Figure 38 : Représentation des ports de communication (d'après [Tarpin-Bernard 1997])

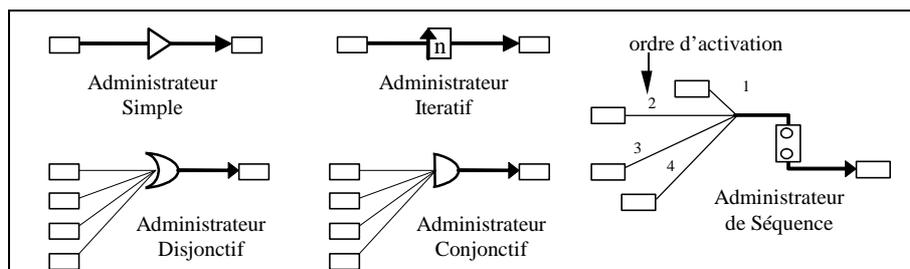


Figure 39 : Représentation schématique des administrateurs (d'après [Tarpin-Bernard 1997])

Les “facettes” *Contrôle* sont différentes des autres facettes car elles n'ont pas véritablement d'existence propre et ne présentent pas de ports de communication. En effet, une facette *Contrôle* est constituée des “administrateurs de contrôle” qui sont gérés par l'agent lui-même.

#### Exemple de modélisation AMF

La Figure 40 présente la modélisation d'une application d'Agenda [Tarpin-Bernard 1999] : elle gère des rendez-vous (qu'il est possible de créer, modifier, déplacer et supprimer) et un calendrier qui peut être visualisé selon 4 modes (jour, semaine, mois, année). Les agents rendez-vous sont à instanciations multiples ce qui est représenté sur le schéma par une bordure en relief. Nous ne détaillerons pas plus ici cet exemple qui est complètement décrit dans [Tarpin-Bernard 1999].

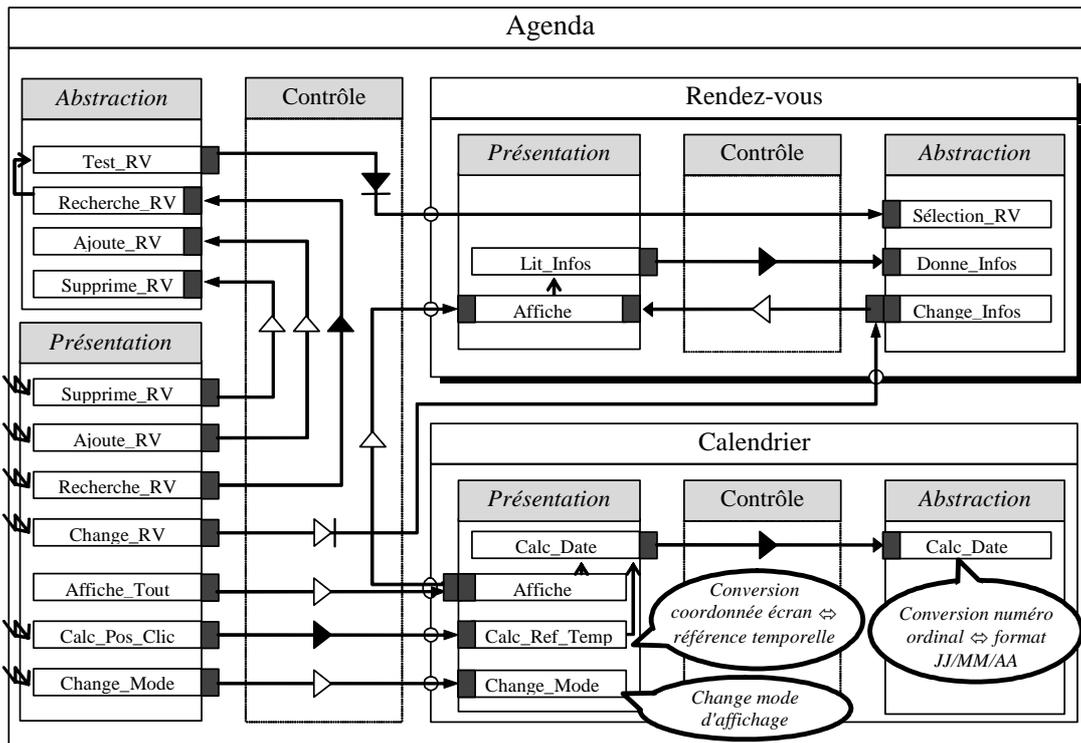


Figure 40 : Modélisation AMF d'un logiciel de gestion d'agenda (d'après [Tarpin-Bernard 1999])

### IV.2.3 Modèles Hybrides

Les approches en couche et multi-agents ne sont pas incompatibles mais plutôt complémentaires comme le montre le modèle PAC-Amodeus ou l'association entre AMF et Arch.

#### IV.2.3.1 PAC-Amodeus

Le modèle hybride PAC-Amodeus (Nigay, 1994, cité dans [Coutaz 2001]) reprend :

- D'une part le modèle en 5 couches Arch,
- D'autre part, structure le Contrôleur de Dialogue de ce modèle en agents PAC.

Ainsi la facette Abstraction d'un agent PAC communique avec un objet conceptuel de l'ANF (ou du NF si l'ANF n'existe pas). Symétriquement, via la facette P, l'agent communique avec un objet de présentation de la couche Composants Techniques de Présentation (Figure 41).

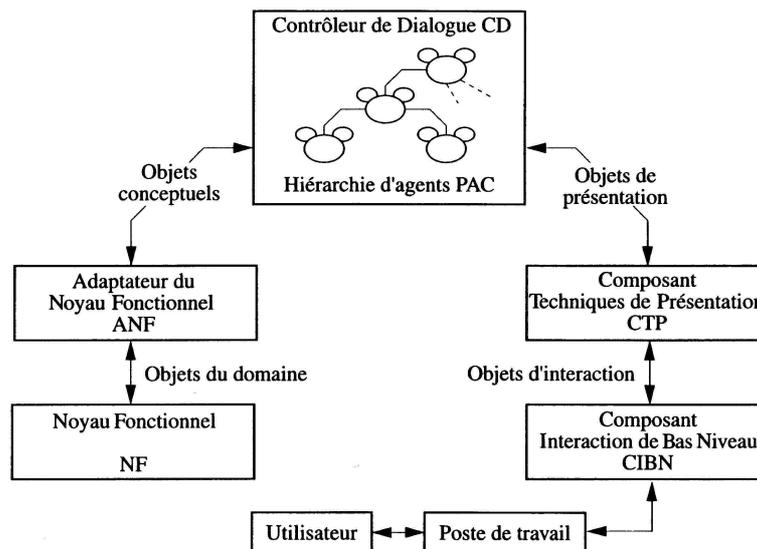


Figure 41 : Modèle Hybride PAC-Amodeus (d'après [Coutaz 2001])

#### IV.2.3.2 AMF et Arch : AMF comme contrôleur de dialogue

AMF a aussi été rapproché du modèle Arch avec cependant une mise en correspondance différente de PAC-Amodeus puisque AMF couvre les 3 couches centrales de l'arche comme le montre la Figure 42 et non le seul contrôleur de dialogue [Samaan 2004]. Dans ce cas la facette abstraction des agents AMF a valeur d'adaptateur du noyau fonctionnel et la facette présentation représente le niveau abstrait de l'interface, le niveau concret se trouvant dans la présentation de l'application.

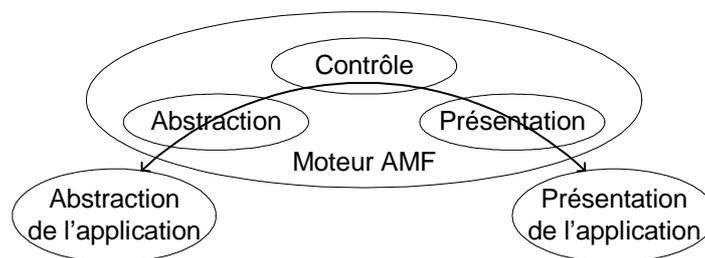


Figure 42 : Les couches de l'implémentation de AMF

Ce rapprochement d'AMF avec Arch permet d'aller au-delà d'un simple modèle d'architecture. En effet, le modèle AMF est directement instancié en contrôleur de dialogue et le « moteur AMF » pilote donc le dialogue de l'application [Vaisman 2002]. Le formalisme graphique utilisé pour modéliser la facette contrôle peut ainsi être vue comme un véritable langage de programmation visuelle du contrôleur de dialogue de l'application elle-même. Il devient donc possible de pouvoir adapter le comportement du dialogue au sein même de l'application ce qui permet d'envisager que les adaptations légères soient directement réalisées par l'utilisateur final.

#### IV.2.3.3 PAC-Amodeus modifié pour les systèmes mixtes

Le modèle ASUR a été rapproché du modèle PAC-Amodeus (Cf. § IV.2.3.1) dans [Dubois 2001b]. Le principe est d'avoir autant de branches du « pilier réel » de l'arche du modèle PAC-Amodeus qu'il existe d'adaptateur entre le monde réel et le monde informatique (Figure 43). Ainsi les deux couches CTP et CIBN du modèle PAC-Amodeus sont répliqués, ce qui assure une indépendance entre les différentes boîtes à outil requises pour s'interfacer avec les différents adaptateurs.

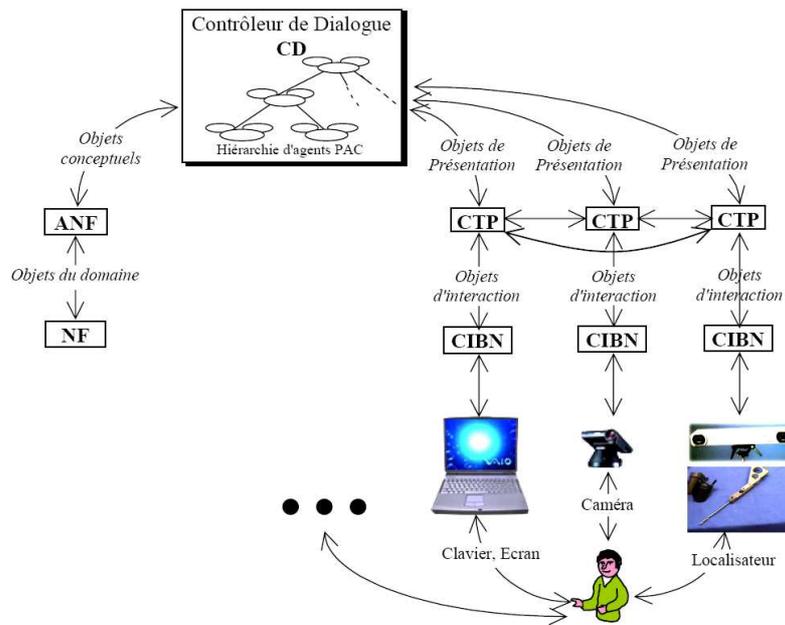


Figure 43 : Adaptation du modèle PAC-Amodeus aux systèmes mixtes (d'après [Dubois 2001b]).

Dans cette adaptation, les différents CTP peuvent communiquer entre eux. En effet, il n'est pas toujours nécessaire de remonter au contrôleur de dialogue certaines actions automatiques qui n'intéressent que les couches de présentation. Par exemple, dans le système Karma (Cf. § II.2.4.1), pour afficher des informations à l'utilisateur qui augmentent sa perception de l'imprimante, il est nécessaire de localiser le casque de visualisation semi transparent ainsi que l'imprimante par un dispositif *ad hoc* afin de pouvoir aligner l'information avec l'objet réel en permanence. Cette mise à jour de l'affichage n'est faite que pour compenser les mouvements de l'utilisateur qui se déplace autour de l'imprimante et n'a aucune incidence sur le comportement interne de l'application.

#### IV.2.3.4 Synthèse sur la portée des modèles d'architecture logicielle pour la Réalité Mixte

Tous les modèles d'architecture classiques, que ce soient les modèles en couches, les modèles multi-agents ou les modèles hybrides, se focalisent sur la modélisation du logiciel, les aspects liés au matériel, tant du côté du noyau fonctionnel (accès aux mémoires de masse, au réseau, etc.) que du côté de l'interface utilisateur (dispositifs d'interaction en entrée et en sortie) sont externes au modèle.

En effet, ils assument tous plus ou moins implicitement une interface utilisateur basée sur le couple clavier/souris en entrée et un écran graphique en sortie ainsi qu'un environnement multifenêtre. Cependant, dès que l'on considère des interfaces d'une autre nature, telle que celles proposées par la Réalité Mixte, ces modèles se montrent dans l'impossibilité de décrire complètement l'interface car certaines parties leur échappent, en particulier tous les aspects liés aux objets réels et en particulier l'interaction de l'utilisateur avec ces objets réels qui ne sont pas prises en compte avec ces modèles. Ils gardent bien sûr tout leur intérêt pour la modélisation du dialogue dans le monde informatique mais doivent donc être complétés sur les aspects liés aux artefacts réels.

Cette limitation a été identifiée par [Dubois 2001b] qui a proposé une extension du modèle PAC-Amodeus en ajoutant pour chaque dispositif d'entrée/sortie (localisateur, caméra, etc.) des piliers à l'arche du modèle qui se connectent sur le contrôleur de dialogue qui joue maintenant un véritable rôle de chef d'orchestre du système.

Une autre adaptation de PAC-Amodeus pour les systèmes mixtes collaboratifs et mobiles a été proposée par [Renevier 2004c] qui est présentée au chapitre suivant, paragraphe V.3.5.

### IV.3 Bilan du chapitre

Ce chapitre a abordé les différents modèles pour l'interaction Homme-Machine selon deux axes différents mais complémentaires, les modèles d'interactions qui guident la conception des interfaces, et les modèles d'architectures qui décrivent les éléments fonctionnels pour l'implémentation de l'interface.

Pour les modèles d'interaction, nous avons vu que la **manipulation directe** restait un principe toujours applicable aux systèmes mixtes. La théorie de l'action qui fournit un cadre conceptuel à ce modèle reste donc pertinente pour analyser les systèmes de Réalité Mixte.

L'**interaction instrumentale** et le **modèle DPI** proposent le concept d'instrument qui nous semble particulièrement intéressant car il considère explicitement la partie réelle du dispositif d'entrée et fait donc un pas vers la Réalité Mixte. Le **modèle MCRpd** ne concerne que les interfaces tangibles, ce qui couvre tout de même une partie de la Virtualité Augmentée. Cependant, il reste à un niveau conceptuel et sa filiation avec MVC quelque peu problématique. Le **modèle de l'interaction physique** de Sato est particulièrement pertinent pour distinguer les différents cas d'interaction avec le monde physique mais lui aussi n'est guère opérationnel car il n'y a pas de formalisme associé permettant de modéliser une application concrète.

Le **modèle ASUR** apparaît, quant à lui, parfaitement adapté à la Réalité Mixte et propose un formalisme pour concevoir des systèmes mixtes. Il présente cependant 3 limitations (réduites à 2 par la dernière version) et ce sont ces limitations qui ont constituées notre point de départ pour proposer le modèle IRVO présenté au Chapitre VI.

Pour les modèles d'architecture, nous avons vu qu'ils ignoraient les objets réels et que les dispositifs d'interaction étaient implicitement pris en compte. Une solution dérivée de PAC-Amodeus a été proposée par [Dubois 2001b] pour pallier à ces limitations. Nous proposons au Chapitre VII nos propres propositions d'extension du modèle AMF-C pour prendre en compte les dimensions liées aux objets réels dans la modélisation de l'architecture du système.

Ce chapitre était volontairement limité à l'Interaction Homme-Machine dans ses aspects mono-utilisateur. Le chapitre suivant introduit le TCAO (Travail Collaboratif Assisté par Ordinateur) puis se focalisera sur les incidences du travail collaboratif sur les aspects liés à l'Interaction Homme-Machine.



# Chapitre V Le Travail Collaboratif Assisté par Ordinateur

*The real danger is not that computers will begin to think like men,  
but that men will begin to think like computers.*  
Sydney J. Harris (1917 – 1986)

Dans ce chapitre, nous abordons le Travail Collaboratif Assisté par Ordinateur (TCAO).

Nous ne prétendons pas avoir ici une approche globale de ce vaste champ de recherche et nous nous focalisons principalement sur les problèmes nouveaux que soulèvent les collecticiels du point de vue de l'interaction homme-machine par rapport aux logiciels mono-utilisateur.

Nous commençons ce chapitre par une définition du TCAO et un rappel des principaux concepts associés. Ensuite, en suivant la même logique que dans le chapitre précédent, nous présentons les différents modèles d'interaction avant de présenter les modèles d'architecture. Pour chaque modèle, nous évaluons la portée du modèle vis-à-vis de notre domaine d'étude, la Réalité Mixte.

Nous concluons ce chapitre par une synthèse critique de ces différents modèles.

## V.1 Présentation du TCAO et définitions

La notion de « travail collaboratif » est très large. On peut, en effet, regrouper sous ce terme toutes les activités humaines faisant intervenir plus d'une personne simultanément ou non. Si ces activités sont supportées par des outils informatisés alors on parle de Travail Collaboratif Assisté par Ordinateur (TCAO) ou Computer Supported Cooperative Work (CSCW), les outils eux-mêmes étant appelés collecticiels ou synergiciels (*groupware*) [Ellis 1991].

Les domaines d'applications sont particulièrement vastes et couvrent des activités diverses comme la conception de produits, la maintenance, l'enseignement, les relations commerciales, les jeux, etc. Bien que ces activités ne poursuivent pas les mêmes objectifs, elles possèdent des caractéristiques communes qui doivent être prises en compte par les systèmes informatiques destinés au support du TCAO. Ainsi, ils doivent permettre aux utilisateurs de réaliser une tâche collaborative à partir de leur poste de travail que ce soit un ordinateur fixe, un portable un ordinateur de poche, voire un téléphone mobile (notion de « TCAO capillaire » [David 2003b, David 2003c]).

Ellis [Ellis 1991] définit un collecticiel comme : « un système informatique qui assiste un groupe de personnes engagé dans une tâche commune (ou but commun) et qui fournit une interface à un environnement partagé ». Il est important de noter dans cette définition que le partage des données est explicite ce qui exclut les systèmes multi-utilisateurs dans lequel le partage de données est implicite et transparent [Tarpin-Bernard 1997].

Dans un collecticiel on cherche à pouvoir réunir différentes personnes éventuellement distantes géographiquement (bureaux voisins, dans une autre ville ou pays) et ne travaillant pas nécessairement en même temps (rythmes différents, décalages horaires, etc.). Il s'agit donc de gérer, via le système informatique, la participation de plusieurs personnes qui ne sont présentes que virtuellement (dans l'espace et/ou le temps). Le premier objectif des collecticiels est donc l'abolition des dimensions espace et temps [David 2001a].

Les principaux objectifs des collecticiels identifiés par [David 2001a] sont :

- Obtenir des gains de performance,
- Capitaliser des connaissances,
- Améliorer des temps de réponse,
- Partager des compétences,
- Faciliter le travail à distance.

Dans les collecticiels, on considère généralement les concepts suivants :

- Groupe et acteurs : chaque individu qui participe à l'activité est appelé acteur. Un groupe est un ensemble d'acteurs travaillant sur un même domaine.
- Rôle : dans un groupe, à un instant donné, chaque acteur joue un rôle. Ce rôle est caractérisé par un nombre de droits vis-à-vis des autres acteurs et des données partagées. Ce rôle peut évoluer au cours du temps et un acteur peut aussi avoir plusieurs rôles.
- Vue : Il s'agit de la perception par un acteur des données partagées. Cette vue peut être publique (accessible aux autres utilisateurs), privée (accessible seulement au propriétaire), semi-privée, etc.
- Informations partagées : ce sont les données partagées à travers le système par les différents acteurs. La granularité des informations gérées par le système peut être grossière (un document entier) ou très fine (un mot).

### V.1.1 Taxonomies

La première taxonomie, désormais classique, est proposée par Johansen [Ellis 1991] qui classe les collecticiels selon une matrice à 2 dimensions (Figure 44) :

- L'axe spatial : le travail collaboratif a-t-il lieu sur le même lieu physique ou à distance sur des lieux différents,
- L'axe temporel : le travail est-il synchrone ou asynchrone.

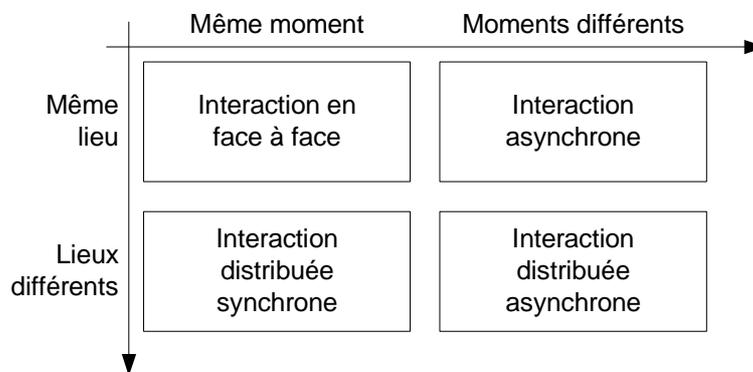


Figure 44 : Taxonomie espace-temps des collecticiels (d'après [Ellis 1991])

Cette matrice a été critiquée par Karsenty [Karsenty 1994] et de nombreuses modifications ont été proposées en particulier par Dix ou Grudin [David 2001a].

Parmi les différentes taxonomies proposées qui affinent la notion de « synchronisme » nous retiendrons les 4 niveaux de coopération de Tarpin-Bernard *et al.* [David 1996, Tarpin-Bernard 1997] :

- La **coopération asynchrone** : les différents participants interagissent en échangeant des données et en travaillant à des moments différents,
- La **coopération en session** : les différents participants travaillent en même temps mais de façon autonome. Ils peuvent communiquer mais ne peuvent pas partager visuellement les objets de la discussion.
- La **coopération en réunion** : les participants travaillent et communiquent simultanément tout en partageant les objets de leurs travaux et de leurs discussions. Ils ont des rôles

attribués en fonction du but de la réunion et l'organisation de leurs interventions est régie par un mécanisme de type « tour de parole »

- La **coopération étroite** : les participants travaillent, communiquent et interagissent en temps réel sur tous les objets partagés du projet. Les conséquences de leurs interventions sont directement gérées et rendues visibles au niveau des objets manipulés.

D'autres classifications ont été proposées comme celle basée sur les types d'applications classées par catégories fonctionnelles [Karsenty 1994] :

- La messagerie électronique.
- Les éditeurs partagés.
- Les conférences et réunions assistées par ordinateur.
- Les systèmes d'aide à la décision.
- Les systèmes de coordination.

### V.1.2 Le trèfle fonctionnel des collecticiels

Ce modèle a été proposé initialement par Ellis *et al.* [Ellis 1994]. Il propose un découpage fonctionnel en 3 parties :

- Le modèle ontologique qui décrit les objets et les opérations que le système fournit,
- Le modèle de coordination modélise les activités de participants et leurs relations,
- Le modèle d'interface utilisateur spécifie l'interface entre le système et les utilisateurs mais aussi l'interface entre les utilisateurs.

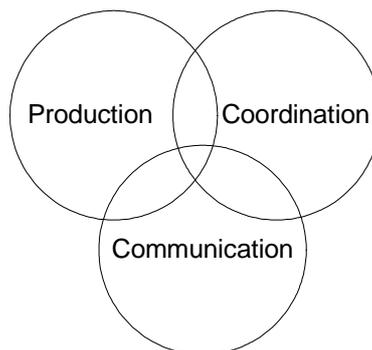


Figure 45 : Le trèfle fonctionnel des collecticiels.

Cette décomposition est critiquable car l'interface utilisateur se trouve mêlée au niveau fonctionnel. Au contraire, l'interface utilisateur ne constitue pas un 3<sup>ème</sup> aspect fonctionnel mais est orthogonale à toutes les fonctionnalités du collecticiel. Salber *et al.* [Salber 1995] ont donc proposé une variante de ce modèle (Figure 45) :

- L'espace de communication. Il s'agit de permettre la communication interpersonnelle des acteurs. Ces échanges ne sont pas interprétés par la machine. Ils peuvent revêtir différentes formes :
  - Messagerie électronique,
  - Chat, messagerie instantanée,
  - Visiophonie, visioconférence,
  - Etc.
- L'espace de production (similaire au modèle ontologique d'Ellis). Il recouvre toutes les fonctionnalités liées à la construction et à l'accès aux informations partagées entre les différents acteurs. En général, les outils sont spécifiques au domaine d'application.
- L'espace de coordination (similaire au modèle de coordination d'Ellis). Cet espace recouvre les aspects de suivi de l'activité des autres acteurs ainsi que l'utilisation des ressources partagées. Il s'agit de définir des rôles, des droits d'accès, des sous-groupes, de distribuer les

tâches, d'en suivre leur accomplissement, etc. Ceci peut être réalisé « à la main » par la communication directe des acteurs, mais, pour rendre plus efficace cette activité de coordination, il est important de fournir des outils adaptés, les outils de workflow en constituant les représentants les plus sophistiqués.

Bertrand David [David 2001a] propose une variante de ce modèle, le modèle « 3C » (Figure 46) :

- En renommant « espace de conversation » ce qui est appelé l'espace de communication ci-dessus, ceci pour bien distinguer les communications interindividuelles des mécanismes de communication au sens informatique du terme.
- En faisant ressortir l'espace de communication comme étant le support aux activités collaboratives de production et de coordination.

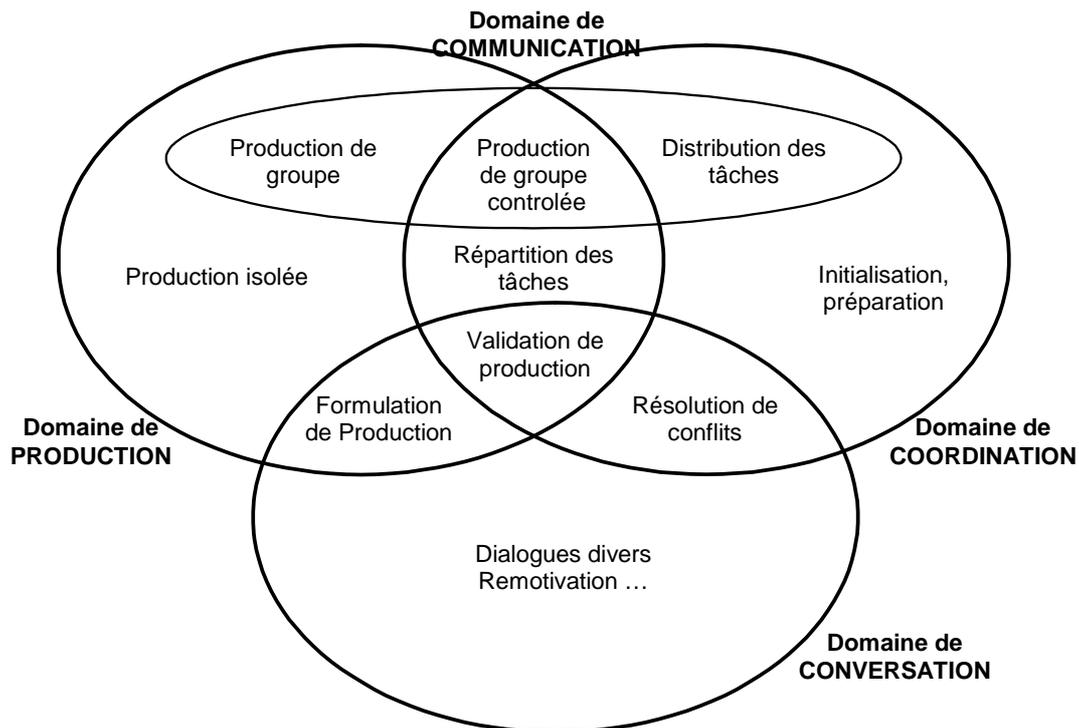


Figure 46 : Modèle « 3C » et les tâches (d'après [David 2001a])

### V.1.3 IHM et les collecticiels

Les collecticiels soulèvent de nombreux problèmes du point de vue de l'interaction homme-machine car l'interaction y est plus complexe que dans les applications mono-utilisateur. En effet, non seulement le système doit permettre à l'utilisateur d'accomplir sa tâche « comme avant » mais il doit aussi permettre de prendre en compte le contexte : d'autres acteurs agissent sur les mêmes données et il faut communiquer et se coordonner avec eux.

Ainsi la priorité est donnée à l'interface multi-utilisateurs. Au-delà de l'interface homme-machine traditionnelle, il s'agit d'évoluer vers une interface homme-machine-homme. Pour être efficace, elle doit respecter plusieurs concepts comme [David 2001a] :

- Le WYSIWIS (What You See Is What I See) qui garantit une identité de vue à chaque acteur. Cependant, le WYSIWIS strict (mise à jour immédiate de toutes les vues) n'est pas forcément souhaitable car cela peut être perturbant pour les utilisateurs et il s'agit le plus souvent d'un WYSIWIS relâché qui retarde la mise à jour des vues à des instants propices.
- La rétroaction de groupe qui permet à chaque acteur de savoir ce que font les autres.
- Le télépointeur qui est un outil de télé-désignation manipulé par un acteur à la fois qui s'adresse au groupe.

Plus globalement, toutes ces considérations sont regroupées sous le concept d'*awareness*, c'est-à-dire la conscience de la présence des autres [Tang 2001] et des moyens permettant d'avoir cette conscience de groupe.

Le contrôle et le partage des données constituent une autre priorité. En effet l'utilisation de données communes, surtout dans des systèmes synchrones, génère des conflits lors d'accès concurrents. On se retrouve confronté à des mécanismes tels que ceux mis en œuvre dans les bases de données réparties et qui doivent assurer le contrôle des droits d'accès aux données, le contrôle de la concurrence et le contrôle de la cohérence des données. Cependant, les mécanismes classiques de verrouillage de données peuvent conduire à l'impossibilité pour un acteur d'accéder aux données ce qui doit être clairement signalé à l'utilisateur par un écho visuel. Il est donc important d'avoir une granularité de gestion des informations assez fine et adaptée à la fréquence des accès aux données partagées. Pour éviter des blocages dus à des objets verrouillés par d'autres utilisateurs, de nombreux collecticiels font le choix d'une approche optimiste en proposant des procédures de réparations en cas de conflit avéré des actions des utilisateurs.

Dans les points importants liés à l'IHM nous citerons encore la prise en compte des différents terminaux utilisés qui peuvent aller d'une station de travail sophistiquée à un téléphone mobile avec un mini-écran. Cette notion de travail collaboratif dans un contexte de nomadisme a été baptisée « Travail Collaboratif Capillaire » par notre équipe [David 2003b, David 2003c] par analogie avec les vaisseaux sanguins pour montrer que le travail collaboratif irrigue l'ensemble des terminaux en ramifications de plus en plus fines.

Dans la suite de ce chapitre, nous présentons les modèles pour l'Interaction Homme-Machine qui prennent en compte les nouvelles contraintes apportées par le travail collaboratif. Comme pour le Chapitre IV, nous examinerons successivement les modèles d'interaction puis les modèles d'architecture en faisant à chaque fois une analyse de la portée de ces modèles dans le cadre de la Réalité Mixte.

## V.2 Modèles d'interaction

Nous nous limiterons ici à ne présenter qu'un seul « modèle » d'interaction qui aborde l'interaction dans le contexte de la Réalité Mixte et qui prend en compte le travail collaboratif. Il prend également en compte la mobilité des acteurs et des dispositifs, mais cet aspect ne sera pas particulièrement développé car la mobilité a été écartée de notre champ de recherche.

### V.2.1 Formalisme de Renevier pour les systèmes mixtes collaboratifs et mobile

Récemment Philippe Renevier a proposé un formalisme pour la conception de systèmes mixtes collaboratifs et mobiles [Renevier 2004b, 2004c]. Ce formalisme a été conçu pour permettre l'écriture sous forme graphique de scénarios projetés dans une approche de conception basée sur des scénarios qui sera présentée au Chapitre VII, § VII.4.2.

Cette notation permet de représenter les utilisateurs et les objets, leurs relations spatiales ainsi que la création ou la destruction d'objet.

Les utilisateurs sont représentés graphiquement sous forme d'un triangle (Figure 47a). Il existe plusieurs variantes graphiques pour représenter les cas où l'utilisateur est représenté (physiquement ou numériquement) et s'il est fixe.

Les objets sont représentés sous la forme d'un cercle, s'il s'agit de l'objet de la tâche, et d'un carré, s'il s'agit d'un outil (Figure 47b). Il existe plusieurs variantes pour représenter les objets physiques et numériques et s'ils sont originaux où si ce sont des représentants d'objets (double cercle ou double carré).

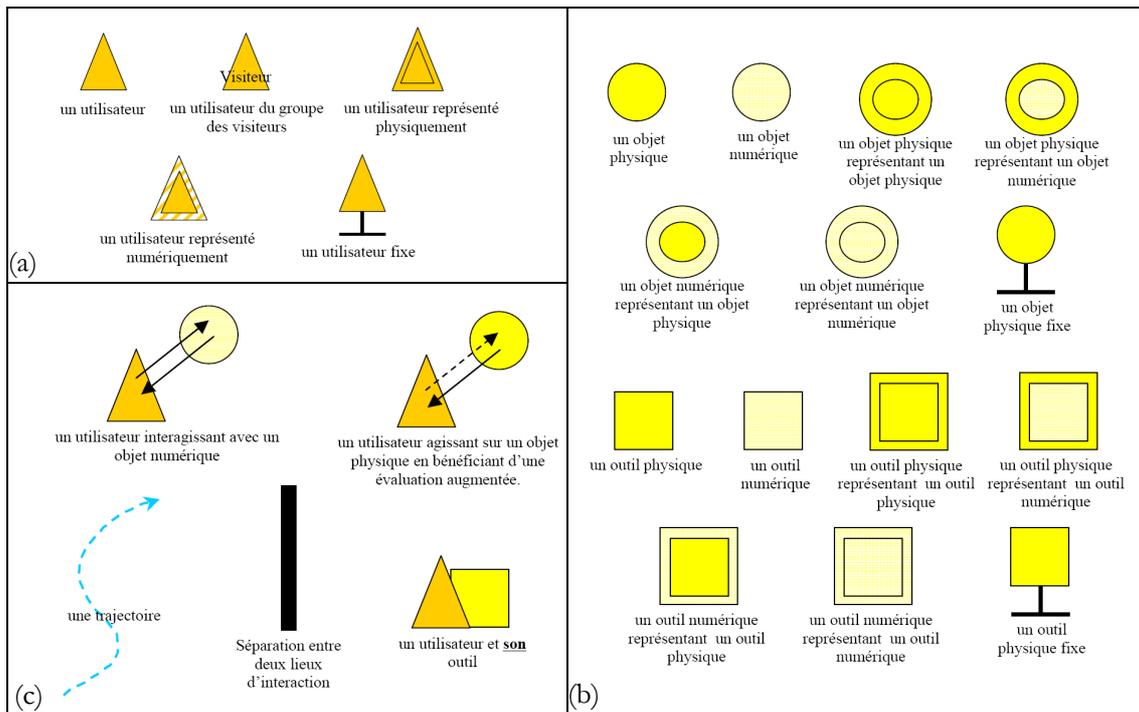


Figure 47 : Formalisme de Renevier : (a) représentation des utilisateurs ; (b) représentation des objets ; (c) représentation des interactions et relations spatiales (d'après [Renevier 2004c]).

Les interactions et les relations spatiales sont représentées en Figure 47c :

- Les interactions mixtes entre un utilisateur et un objet sont représentées par une double flèche (ou une simple flèche si elle est à sens unique). Les actions de l'utilisateur sans intervention du système informatique sont représentées par une flèche pointillée.
- L'adjacence de deux entités exprime une relation d'association. Les deux entités forment un tout qui se déplace ensemble.
- Un mouvement spécifique qui a un impact sur l'interaction est représenté par des trajectoires. Sinon le mouvement libre des utilisateurs autour des objets est implicite.
- La séparation entre 2 lieux d'interaction est représentée par une barre épaisse ce qui traduit le fait qu'un utilisateur ne peut agir directement sur une entité qui se trouve au-delà de la séparation.

Enfin, il existe deux relations particulières :

- Une flèche double ( $\Rightarrow$ ) représente la création d'une nouvelle entité tandis qu'une flèche double barrée représente sa destruction. Cela permet de traduire la création/destruction des augmentations entre le monde physique et le monde numérique.
- Lorsqu'une entité originale et une ou plusieurs de ses représentations sont sur le même schéma elles sont reliées par un trait entre ces entités.

La Figure 48 présente un exemple basé sur l'application Navicam+, extension collaborative de Navicam [Rekimoto 1995].

L'objet (cercle) est un tableau dans un musée et est fixe (symbolisé par le T renversé). Dans un premier temps (étape A) un expert agit sur le tableau (flèche dirigée vers l'objet) en définissant l'augmentation (flèche dirigée vers l'utilisateur) qu'il associe au tableau. Ce mécanisme de définition de l'augmentation est représenté par la flèche double. Dans un deuxième temps (étape B) un visiteur observe l'œuvre d'art et en même temps l'augmentation associée (représentée par le K barré) qui sont les informations qu'avait laissés l'expert auparavant.

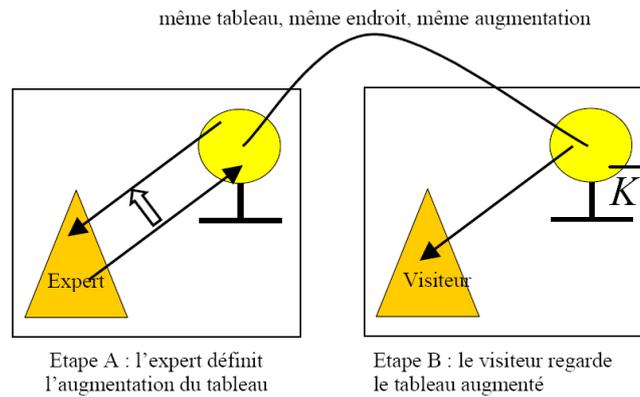


Figure 48 : Exemple de représentation de scénario de Navicam+ (d'après [Reneier 2004c])

### Commentaires sur ce formalisme

Ce formalisme permet une représentation plus concise qu'un scénario textuel et est plus simple ce qui favorise l'exploration des solutions en phase de conception d'après [Renevier 2004c].

Ce formalisme se plaçant à un niveau abstrait, ne détaille pas l'interaction concrète de l'utilisateur avec le système et en particulier les dispositifs d'entrée/sortie utilisés. D'ailleurs, Philippe Renevier propose dans une deuxième phase de conception de passer à des scénarios concrets où il s'agit de réaliser les scénarios abstraits en choisissant les modalités d'interaction. A ce stade, il propose de changer de formalisme et d'utiliser le formalisme ASUR (Cf. § IV.1.7) pour modéliser le système dans cette phase de scénarios concrets.

Le formalisme d'IRVO que nous présentons au chapitre suivant présente plusieurs similarités avec ce formalisme mais en se plaçant au même niveau que ASUR donc à un niveau concret de la spécification de l'interface. La comparaison entre les trois formalismes est faite au paragraphe VI.3.3.

## V.3 Modèles d'architecture

Les modèles d'architecture des logiciels collaboratifs sont, pour la plupart, une évolution des modèles d'architectures logicielles des applications mono-utilisateur présentées au chapitre précédent (§ IV.2). Ces modèles sont nombreux et nous avons choisi de ne présenter que les plus importants.

Nous examinons tout d'abord le modèle de Dewan qui est un modèle en couche adapté aux collecticiels et qui peut être vu comme une évolution/généralisation du modèle Arch (Cf. § IV.2.1.2).

Nous examinons ensuite l'évolution du formalisme PAC-Amodeus en PAC\* pour prendre en compte le trèfle du collecticiel, puis sa fusion avec le modèle de Dewan pour produire le modèle Clover. Nous présentons également l'évolution du formalisme AMF en AMF-C de manière plus approfondie car il est utilisé dans nos travaux. Nous terminons par la dernière évolution en date du modèle PAC-Amodeus pour prendre en compte les systèmes mixtes collaboratifs et mobiles.

Nous concluons cette section par une discussion sur le positionnement de ces modèles par rapport à la Réalité Mixte.

### V.3.1 Modèle de Dewan

Prasun Dewan a proposé un modèle d'architecture en couche (Figure 49) pour les logiciels collaboratifs [Dewan 1998] que l'on peut voir comme une extension/généralisation du modèle Arch (Cf. § IV.2.1.2). Au niveau le plus bas, c'est-à-dire le plus proche de l'utilisateur, la couche 1 gère les

entrées/sorties avec l'utilisateur. Chaque Couche N+1 représente une abstraction des interacteurs présents dans la Couche N, de la même manière que la facette A d'un agent PAC ou AMF est l'abstraction de la facette P. Les abstractions de la couche N+1 envoient des événements de sortie à leurs interacteurs de la couche N et reçoivent des événements d'entrée de ces derniers. Au plus haut niveau se trouve la « couche sémantique » qui gère les abstractions de l'application (appelés les « objets sémantiques ») ce qu'on peut assimiler au Noyau Fonctionnel du model Arch (Cf. § IV.2.1.2)

Dans ce modèle, certaines couches sont partagées (couches L à R+1) : elles contiennent des objets qui sont partagés par tous les utilisateurs. D'autres couches sont répliquées (couches R à 1) et elles contiennent des objets qui sont la propriété d'un seul utilisateur. La dernière couche partagée est appelée « couche de base » et la première répliquée est appelée « point de branchement ». Ainsi, dans le cas le plus général ce modèle se présente donc comme un « arbre de couches ». Dans le cas particulier d'une application mono-utilisateur le modèle se simplifie en une architecture simple en couche à la manière de Seeheim ou Arch (Cf. § IV.1).

Le modèle fait apparaître des communications horizontales entre couches répliquées qui représentent des événements collaboratifs et ce qui permet de mettre en œuvre le concept d'*awareness*. Ceci peut être accompli de la couche 1 (Station de travail), à la couche L (dans le cas de noyaux fonctionnels répliqués) ou à toute autre couche intermédiaire selon l'architecture choisie.

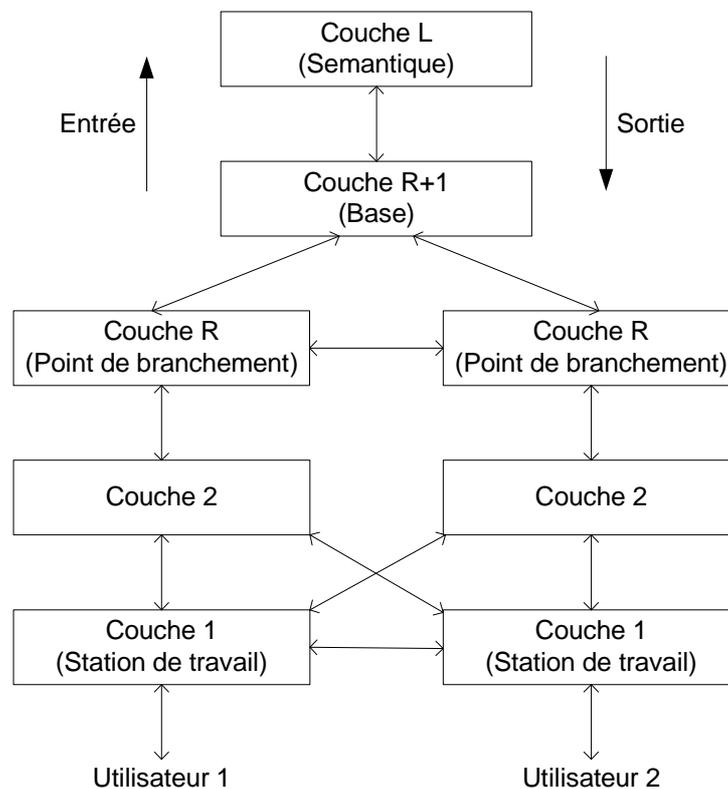


Figure 49 : Modèle de Dewan (d'après [Dewan 1998])

### V.3.2 Modèle PAC\*

PAC\* est une extension de PAC-Amodeus (Cf. § IV.2.3.1) orientée vers les collecticiels où chaque facette est décomposée selon les 3 dimensions du trèfle du collecticiel : coproduction, coordination et communication [Calvary 1997, Coutaz 2001].

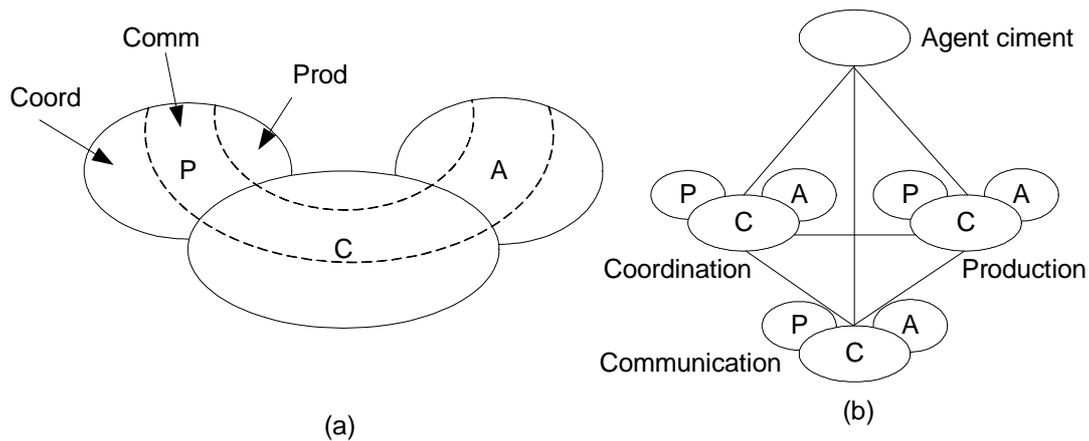


Figure 50 : Modèle PAC\* (d'après [Calvary 1997])

Chaque facette de l'agent PAC (Cf. § IV.2.2.2) est décomposée en 3 parties en suivant les 3 aspects du trèfle du collecticiel (Cf. § V.1.2). La Figure 50a représente une vue « compacte » de cette décomposition. De manière alternative, l'agent PAC\* peut être décomposé en 3 agents PAC, chacun dédié à l'un des aspects du modèle en trèfle. Les 3 agents communiquent entre eux directement et, via un « agent ciment » (dépourvu de facettes P et A), communiquent avec d'autres agents PAC\* (Figure 50b). Il existe encore d'autres manières de structurer les agents PAC\* [Calvary 1997].

Par rapport au modèle PAC-Amodeus, le contrôleur de dialogue est peuplé d'agents PAC\* au lieu d'agent PAC. Un rapprochement avec le modèle de Dewan a également été proposé pour PAC\* [Coutaz 2001].

### V.3.3 Modèle « Clover »

Le modèle Clover, proposé par Yann Laurillau *et al.* [Laurillau 2002a, 2002b, 2002c], reprend les principes de structuration des modèles de Dewan et PAC\* (Figure 51) :

- Il instancie le modèle de Dewan à un modèle en 6 couches correspondant aux 5 couches du modèle Arch où :
  - le Noyau Fonctionnel est divisé en 2 couches : Un « Trèfle Fonctionnel » (TF) partagé et un « Trèfle Fonctionnel » répliqué.
  - les autres couches du modèle Arch sont répliquées pour chaque utilisateur.
- Les 2 couches correspondantes au Noyau Fonctionnel sont structurées en 3 composants, Production, Communication et Coordination correspondant au modèle en trèfle.
- Le contrôleur de dialogue est structuré en hiérarchie d'agents PAC\* conformément à ce modèle et au modèle PAC-Amodeus dont il dérive.

Un « métamodèle Clover » est également proposé [Laurillau 2002a, 2002b, 2002c]. Il suit la même structure générale que le modèle de Dewan, à savoir que le nombre de couches n'est pas fixé et que le point de branchement entre couches partagées et couches répliquées peut être choisi à tout niveau. Chaque couche est structurée en 3 composants suivant le modèle en trèfle, mais cette décomposition n'est pas obligatoire (ou peut être partielle) et n'est en général pas présente dans les couches les plus basses proches du matériel.

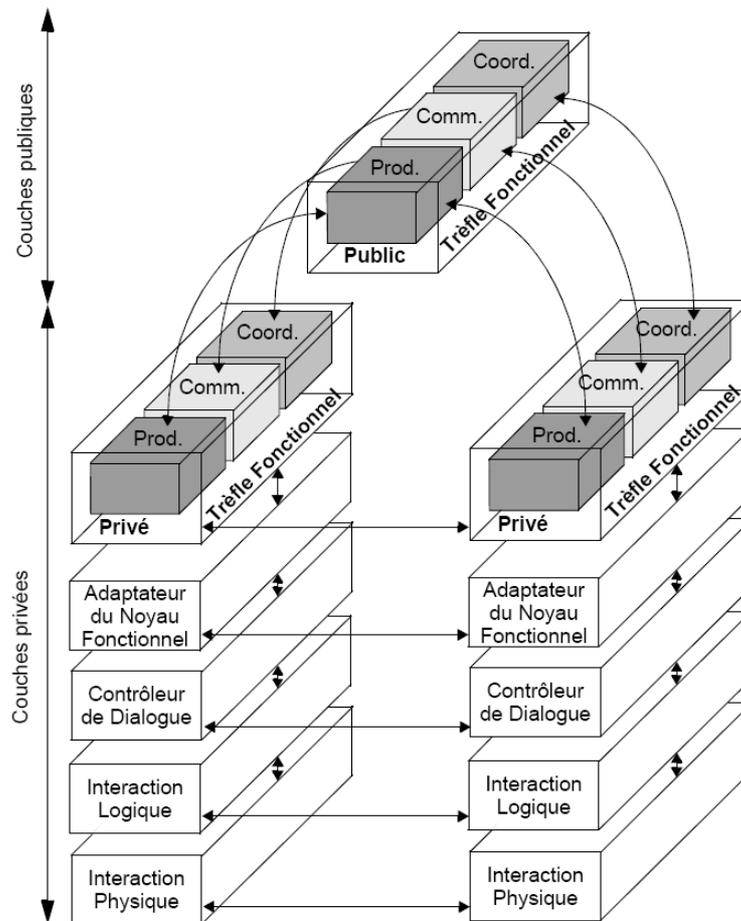


Figure 51 : Modèle Clover (d'après [Laurillau 2002c])

### V.3.4 Modèle AMF-C

AMF-C (AMF Coopératif) est une extension du modèle AMF (Cf. § IV.2.2.3) au travail collaboratif proposé dans notre équipe par Franck Tarpin-Bernard [Tarpin-Bernard 1997].

A la base, le logiciel est toujours structuré en agents AMF. Les utilisateurs interagissent avec les agents selon 2 approches différentes :

- Soit ils interagissent tous avec un seul agent. Celui-ci doit être **fragmenté** en plusieurs morceaux répartis sur les différents postes de travail et éventuellement sur un ou plusieurs serveurs.
- Soit ils interagissent chacun avec leur propre copie de l'agent qui est donc **répliqué** sur chaque poste de travail.

Dans les deux cas le système doit connaître l'état et la localisation des éléments dispersés dans le réseau (fragments ou copie d'agent).

#### V.3.4.1 AMF-C fragmenté

Le principe est donc de fragmenter les agents sur les différents postes de travail ou serveurs. Une frontière naturelle de fragmentation est bien sûr la facette de l'agent AMF-C. Ainsi, une fragmentation classique repose sur un découpage de l'agent :

- Sur le serveur, les facettes Contrôle et Abstraction,
- Sur chaque client, une facette Présentation. Il y a donc autant de facettes Présentation qu'il y a d'utilisateurs.

On voit donc ici l'un des intérêts du modèle AMF qui permet l'ajout d'un nombre quelconque de facettes ce qui permet de multiplier les différentes « vues » pour les utilisateurs : ces facettes peuvent être identiques, ce qui permet un WYSIWIS strict ou différentes en fonction des rôles des utilisateurs.

La Figure 52 donne un exemple simple (c'est quasiment un « pattern ») d'un agent AMF-C fragmenté. Un nouvel administrateur permet le verrouillage pour gérer la concurrence d'accès à l'objet.

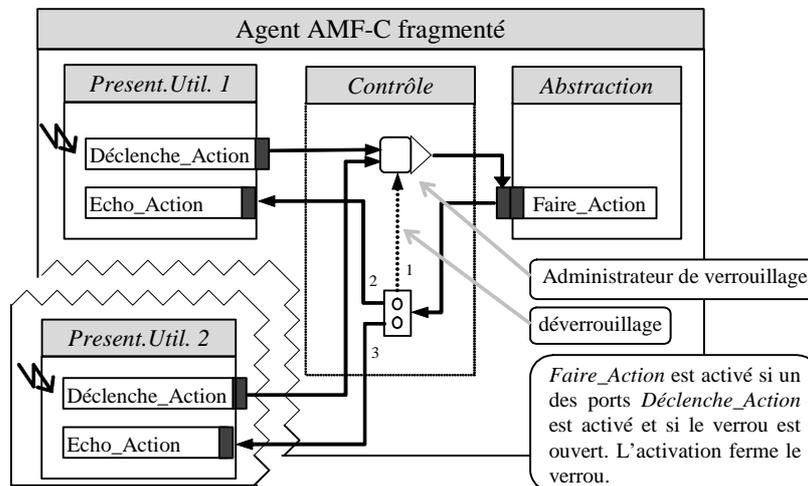


Figure 52 : Exemple d'agent AMF fragmenté (d'après [Tarpin-Bernard 1998])

La fragmentation des facettes n'est pas limitée aux facettes présentation mais peut tout aussi bien s'appliquer aux facettes contrôle et abstraction. Cependant, lorsqu'il y a trop de facettes réparties, il semble préférable de ne plus fragmenter les agents mais de les répliquer.

#### V.3.4.2 AMF-C répliqué

La réplification est basée sur les notions d'agent de référence et d'agent local (Figure 53) :

- Dans le premier cas (Figure 53a), il existe un agent de référence (typiquement placé sur un serveur) partagé par  $n$  participants : chaque participant interagit avec un agent local qui est sur le poste de travail. L'avantage de cette architecture est la simplicité et la régularité mais elle est rigide et peut provoquer des goulots d'étranglement.
- Dans le deuxième cas (Figure 53b), l'agent de référence peut être « virtuel » : les agents locaux placés sur les postes de travail communiquent directement entre eux à travers le réseau. C'est le cas typique d'une architecture pair à pair sans serveur central. En pratique, au lieu d'un véritable agent de référence réparti, c'est l'un des agents locaux qui peut être « élu dynamiquement » agent de référence pour la durée d'une session. Cette architecture permet une meilleure tolérance aux pannes et des temps de réponse meilleurs mais est plus complexe à gérer.

Dans les 2 cas, les agents locaux doivent signifier toute action (création, modification, destruction) aux autres agents locaux frères (directement ou via l'agent de référence) pour que tous les agents soient dans le même état. On retrouve donc les problèmes classiques de programmation concurrente avec pose/libération de verrous sur les objets. Pour mettre en place ces mécanismes, AMF-C s'appuie sur une plateforme de groupware générique [Tarpin-Bernard 1997], [Delotte 2003]. Le lien est assuré par des administrateurs *ad hoc* qui permettent la sélection/libération d'objet et la validation des actions des utilisateurs. C'est donc au niveau de la plateforme que sont implémentées les stratégies d'accès optimiste ou pessimiste aux objets sans que la structure de l'agent AMF-C soit modifiée.

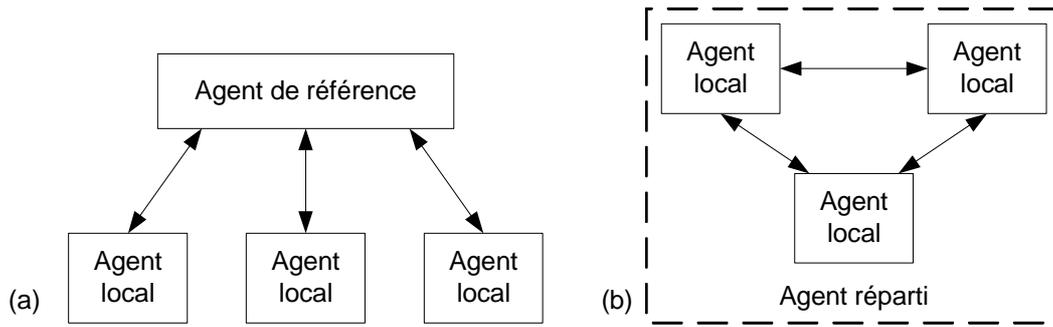


Figure 53 : AMF-C répliqué ; (a) gestion centralisée ; (b) gestion répartie (d'après [Tarpin-Bernard 1997])

La Figure 54 donne un exemple de 2 agents AMF-C répartis. Une nouvelle facette « Distant » est ajoutée qui permet de faire l'écho local des actions des autres utilisateurs. La facette « Accès », quant à elle, permet de désactiver la facette « Présentation » en fonction des droits effectifs de l'utilisateur. Une dernière facette « Adaptation locale » permet de changer les stratégies locales des administrateurs et ainsi de mettre en place facilement diverses formes de WYSIWIS relâché.

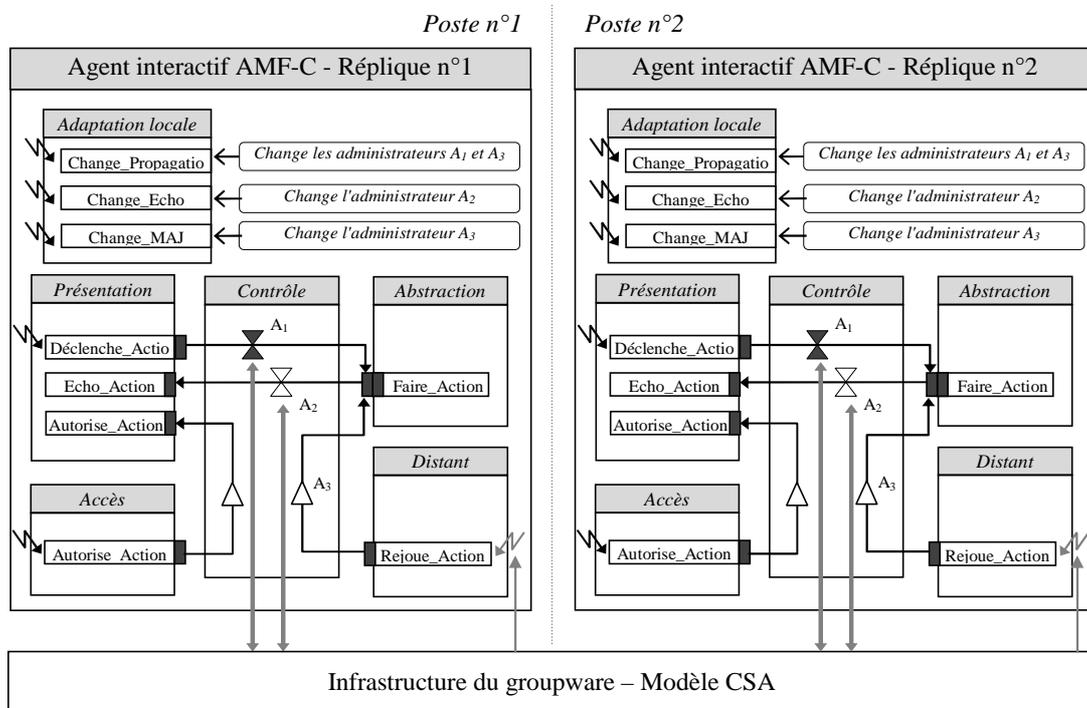


Figure 54 : Exemple d'agents AMF-C répartis (d'après [Delotte 2003]).

### V.3.5 PAC-Amodeus modifié pour les systèmes mixtes collaboratifs et mobiles

Philippe Renevier propose une extension du modèle PAC-Amodeus pour modéliser l'architecture des systèmes mixtes collaboratifs et mobiles [Renevier 2003c].

D'une part, pour les aspects liés au travail collaboratif, il suit les extensions proposées par Laurillau dans le modèle Clover (Cf. § V.3.3) où le Noyau Fonctionnel de l'application est fractionné en deux parties : un Noyau Fonctionnel commun partagé par tous les utilisateurs et autant de NF privés qu'il y a d'utilisateurs.

D'autre part, pour les aspects liés à la mobilité, il faut prendre en compte les informations relatives à la localisation de l'utilisateur (essentiellement sa position et son orientation). D'une manière plus

générale il s'agit de prendre en compte le contexte qui génère trois flux d'informations contextuelles (Figure 55) :

- Les données contextuelles peuvent être des données à part entière, des propriétés des objets numériques et/ou physiques manipulés. C'est par exemple la création d'une annotation numérique d'un objet physique à un endroit précis : la localisation est une propriété de l'augmentation qui doit être enregistrée dans le Noyau Fonctionnel. L'auteur propose de faire aboutir ce flux au niveau de l'Adaptateur du Noyau Fonctionnel (flux 1).
- Le contexte des utilisateurs détermine leurs activités et de ce fait leurs interactions. Par exemple on peut envisager, dans une application d'aide à la visite d'une ville, que le système propose des interfaces différentes lorsqu'on est à l'extérieur (pour trouver son chemin comme avec la *Touring Machine* de Feiner – Cf. § II.2.8.3) ou à l'intérieur des bâtiments (pour visiter un musée comme dans le musée augmentée de Rekimoto – Cf. § II.2.8.1). L'auteur propose donc de traiter ce flux au niveau du Contrôleur de Dialogue (flux 2).
- Les données contextuelles peuvent être utilisées comme des données en entrée lors des interactions. Ce flux a été identifié par Emmanuel Dubois [Dubois 2001b] (Cf. § IV.2.3.3) et concerne les deux couches Présentation Physique et Présentation Logique. Par exemple le mouvement de l'utilisateur permet d'accéder aux informations comme si l'utilisateur manipulait la souris (flux 3).

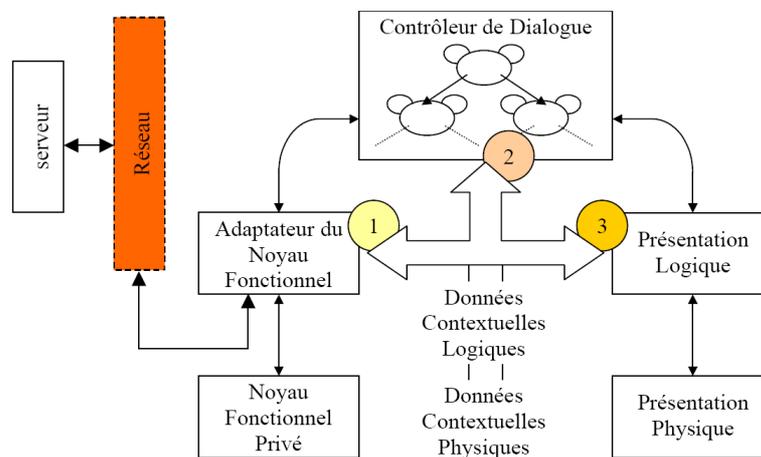


Figure 55 : Modèle PAC-Amodeus modifié et les 3 flux de données contextuelles (d'après [Renevier 2003c]).

### V.3.6 Synthèse sur les modèles d'architecture des collecticiels et la Réalité Mixte

Nous avons vu que les modèles d'architecture pour les collecticiels comme PAC\*, Clover, ou AMF-C sont des extensions des modèles d'architectures vus au chapitre précédent. Les critiques que nous avons formulées pour ces modèles restent valables à savoir que tous les aspects liés aux objets réels, et en particulier l'interaction de l'utilisateur avec ces objets réels, ne sont pas pris en compte et ces modèles doivent donc être complétés sur les aspects liés aux artefacts réels.

Ce problème a été identifié par Philippe Renevier [Renevier 2004c] qui propose une solution pour la prise en compte du contexte dans l'architecture PAC-Amodeus. Cela permet effectivement de prendre en compte les objets réels qui interviennent dans l'interaction (par exemple, leur localisation) mais aussi plus généralement l'environnement et les aspects liés à la mobilité.

Cependant, la solution proposée qui permet à des données de « shunter » le Contrôleur de Dialogue dans le cas du flux 1, nous semble discutable car elle remet en cause le statut central du contrôleur de dialogue. Toutefois cela demanderait une analyse approfondie des avantages et des inconvénients de cette solution par rapport au respect du modèle en couche de l'arche.

## V.4 Bilan du chapitre

Dans ce chapitre, après avoir rappelé brièvement la définition du Travail Collaboratif Assisté par Ordinateur (TCAO) et de ses principaux concepts, nous avons abordé l'incidence du TCAO sur l'Interaction Homme-Machine. Nous avons plus particulièrement examiné les modèles d'IHM qui prennent en compte les nouvelles contraintes apportées par le travail collaboratif, en adoptant le même plan qu'au Chapitre IV.

Pour les modèles d'interaction, nous avons examiné le **formalisme de Renevier** qui permet de représenter pour les systèmes mixtes collaboratifs et mobiles, les utilisateurs, les objets, et leurs relations spatiales ainsi que la création ou la destruction d'objets. Ce formalisme est développé initialement pour représenter des scénarios abstraits sous une forme graphique plus simple et plus concise que son équivalent textuel, ce qui favorise l'exploration des solutions en phase de conception. Cependant ce formalisme se place à un niveau abstrait et ne détaille pas l'interaction concrète de l'utilisateur avec le système et en particulier les dispositifs d'entrée/sortie utilisés, ce qui amène son auteur à proposer un changement de formalisme en reprenant le formalisme ASUR (Cf. § IV.1.7) pour modéliser le système dans cette phase d'expression des scénarios concrets. Notre formalisme IRVO, que nous présentons au chapitre suivant, présente plusieurs similarités avec ce formalisme mais en se plaçant au même niveau que ASUR donc à un niveau plus concret de la spécification de l'interface.

Pour les modèles d'architecture des collecticiels, nous avons vu que, comme les modèles d'architecture dont ils sont issus, ils ignorent les objets réels et que les dispositifs d'interaction sont implicitement pris en compte. Pour pallier à ces limitations, une extension de PAC-Amodeus a été proposée par [Renevier 2004c].

Nous voyons au § VII.3.3.2 que nous proposons également d'étendre le modèle d'architecture AMF-C, sur lequel nous nous appuyons, en ajoutant une facette aux agents pour prendre en compte les aspects liés au monde réel.

# Chapitre VI IRVO : Notre proposition pour la modélisation de l'IHM dans le contexte de la Réalité Mixte

*Entia non sunt multiplicanda praeter necessitatem.*  
[Les entités ne devraient pas être multipliées sans nécessité]  
Guillaume d'Occam (v. 1285 – 1349)

Nous avons passé en revue, au Chapitre IV, les modèles dédiés à la modélisation des interfaces Homme-Machine et, au Chapitre V, leurs extensions pour la prise en compte du TCAO.

Parmi tous ces modèles, seuls deux apparaissent aujourd'hui comme prenant en compte explicitement la Réalité Mixte :

- Le formalisme ASUR pour les systèmes mixtes et mobiles,
- Le formalisme de Renevier pour les systèmes mixtes collaboratifs et mobiles.

Dans ce chapitre, nous présentons notre propre modèle de l'interaction pour les systèmes mixtes et collaboratifs. Au début de nos travaux, seul existait le formalisme ASUR dans sa version initiale (§ IV.1.7.1) et notre proposition partait de certaines limitations identifiées dans ASUR que nous voulions résoudre. De manière concomitante, le modèle ASUR a évolué et le formalisme de Renevier est apparu ce qui donne un terrain de discussion nouveau à nos travaux propice à une émulation réciproque.

D'après [Beaudouin-Lafon 2000a] un modèle de l'interaction doit posséder :

- des **capacités descriptives** : on doit être capable de décrire des applications existantes et de nouvelles applications.
- des **capacités comparatives** : pour fournir des métriques afin de comparer des conceptions alternatives (opposé aux capacités *prescriptives* qui décident a priori ce qui est bon ou mauvais).
- des **capacités génératives** : pour faciliter la création de nouvelles techniques d'interaction.

Nous étudierons les capacités descriptives et comparatives d'IRVO mais les capacités génératives sont beaucoup plus difficiles à évaluer.

Ce chapitre est donc organisé de la manière suivante. La première section est consacrée à la présentation du modèle IRVO et de ses objectifs. Dans la section VI.2 nous présentons de manière détaillée le formalisme graphique d'IRVO. Nous analysons les capacités descriptives d'IRVO dans la section VI.3 et ses capacités prédictives dans la section VI.4.

## VI.1 Présentation du modèle IRVO

### VI.1.1 Sources du modèle

Notre travail de modélisation a trouvé ses origines dans de nombreuses sources.

La première influence provient du modèle de l'interaction instrumentale de Beaudouin-Lafon (Cf. § IV.1.2) qui apporte des éléments pour la prise en compte du monde réel dans l'interaction grâce à

la notion d'instrument qui se décompose en une partie physique et une partie logique représentée à l'écran. Ce modèle ne considère toutefois pas le cas où les objets du domaine sont réels, ce qui ne permet pas de l'utiliser comme modèle pour la réalité mixte. Malgré tout, cette notion d'instrument a fortement influencé notre modèle IRVO dans lequel la notion d'outil mixte (une partie réelle et une partie virtuelle couplée) est très proche.

Le modèle DPI (Cf. § IV.1.3) reprend la notion d'instrument dans un modèle plus large où les objets du domaine sont considérés sous la métaphore du document. Le modèle conceptuel de DPI, qui est rapproché d'un diagramme action-perception basé sur la théorie de l'action de Norman, a également été une source importante du modèle IRVO.

Notre point de départ originel principal pour la constitution d'IRVO a été la constatation des limitations du modèle ASUR identifiées par Emmanuel Dubois dans sa thèse en juillet 2001 [Dubois 2001b] :

- Les deux types d'augmentation (exécution augmentée et évaluation augmentée, Cf. § II.3.5) ne sont pas distinguables dans les schémas ASUR,
- Les outils et objets virtuels ne sont pas représentés ; ainsi ASUR est d'un intérêt limité pour modéliser des applications de Virtualité Augmentée,
- Un seul utilisateur est représenté. ASUR ne peut donc prendre en compte les systèmes collaboratifs.

Nous avons apporté des solutions à ces trois problèmes dans IRVO [Chalon 2004a, Chalon 2004b] :

- Pour distinguer les 2 types d'augmentation :
  - D'une part nous représentons explicitement les actions des outils sur les objets, qu'ils soient réels ou virtuels, ce qui permet d'identifier aisément les cas d'exécution augmentée,
  - D'autre part nous représentons explicitement la fusion perceptuelle d'objets réels et virtuels par les utilisateurs ce qui fait apparaître immédiatement les cas d'évaluation augmentée.
- En représentant tant les objets réels que virtuels, ce qui permet de représenter les cas de réalité augmentée et de virtualité augmentée.
- En permettant la représentation de plusieurs utilisateurs et de leurs interactions tant avec le système qu'avec les autres utilisateurs ce qui permet de prendre en compte le travail collaboratif.

Depuis la thèse d'Emmanuel Dubois et en concomitance à nos travaux, le modèle ASUR a connu plusieurs évolutions, dont la dernière en date (septembre 2004), ASUR 2004 (Cf. § IV.1.7.3), introduit des outils et objets virtuels et résout une des limitations identifiées. Cependant, l'aspect multi-utilisateur n'est toujours pas présent dans ASUR 2004.

De même, et toujours en concomitance à nos travaux, Philippe Renevier dans sa thèse [Renevier 2004c] a développé un formalisme pour représenter les systèmes mixtes collaboratifs sur supports mobiles que nous avons présenté au paragraphe V.2.1. Ce formalisme permet de représenter à peu près tout ce que IRVO peut représenter et même plus puisqu'il prend en compte les supports mobiles. Toutefois il se place à un niveau plus abstrait de l'interaction et ne considère pas les dispositifs d'entrée/sortie dans ses représentations.

En comparaison avec ces nouveaux formalismes, IRVO possède encore des caractéristiques propres qu'aucun des 2 autres formalismes ne prend en compte. Une comparaison entre les 3 formalismes est faite au § VI.3.3.

### VI.1.2 Objectifs

L'objectif principal d'IRVO est de **modéliser l'interaction entre le (ou les) utilisateur(s) et le système de RM en explicitant les outils et objets (tant réels que virtuels) mis en jeu et leurs relations.**

IRVO est conçu pour être utilisé dans la phase de conception du système et reflète l'utilisation future du système par l'utilisateur. Il ne se préoccupe donc pas de l'architecture du logiciel et n'intervient pas dans la phase de développement du logiciel. Par contre, le passage entre ces deux phases est étudié dans le cadre d'une approche basée sur les modèles (MBA) au chapitre suivant.

Pour obtenir un niveau de détail suffisant pour la conception du système, **IRVO se place au niveau des tâches élémentaires d'interaction**, c'est-à-dire qu'il se place au niveau des feuilles de l'arbre de tâches (Cf. § VII.2). Il s'agit donc d'un niveau très concret où l'on considère les actions élémentaires de l'utilisateur avec le système de Réalité Mixte. Pour une application donnée, il existe donc *a priori* autant de modélisations IRVO que de tâches élémentaires. Toutefois, nous montrons au chapitre suivant (§ VII.2.3) que IRVO peut se placer à un niveau d'abstraction plus élevé en s'appuyant sur l'arbre de tâches.

IRVO a deux finalités :

- Modéliser des applications existantes dans le but de les analyser, les comparer, et les classer. En modélisant les applications existantes, il est ainsi possible d'établir un catalogue d'exemples dans lequel on peut ensuite puiser lors de la conception de nouveaux systèmes.
- Modéliser de nouvelles applications dans le but d'explorer plusieurs solutions tant sur le plan des dispositifs techniques d'interaction que sur le plan des critères ergonomiques.

### VI.1.3 Principales entités et relations d'IRVO

Le modèle IRVO distingue 3 catégories principales d'entités (Figure 56a) :

- **l'utilisateur (U)**, ou plus généralement les utilisateurs pour des systèmes collaboratifs,
- les **objets** manipulables et/ou perceptibles par les utilisateurs ; ils sont distingués en 2 sous-catégories selon l'usage qui en est fait par l'utilisateur :
  - **l'objet du domaine de la tâche (O)** : c'est l'objet sur lequel se porte l'intérêt de l'utilisateur pour la tâche. Il s'agit soit d'un objet physique, soit de données informatiques représentées sous une forme externe perceptible (objet virtuel). L'utilisateur agit sur ces objets par l'intermédiaire des outils (ou directement s'il n'y a pas d'outil).
  - les **outils** ( $T=tool$  en anglais<sup>14</sup>) : ils participent à la réalisation de la tâche en permettant à l'utilisateur d'agir sur d'autres objets (d'autres outils ou l'objet du domaine de la tâche).
- le **modèle interne** de l'application (M) qui représente l'application informatique privée des objets virtuels. En se basant sur le modèle Arch (§ IV.2.1.2), le modèle interne correspond donc au Noyau Fonctionnel, à l'Adaptateur de Noyau fonctionnel, au Contrôleur de Dialogue et à la Présentation Abstraite.

L'utilisateur est en relation avec les objets mais pas directement avec le modèle interne. Le modèle interne est lui en relation avec les objets pour connaître leur état et éventuellement les modifier.

---

<sup>14</sup> Afin de ne pas avoir de problèmes d'adaptation pour présenter nos travaux dans les conférences internationales (donc en langue anglaise), nous avons choisi de développer la notation IRVO en utilisant systématiquement des abréviations anglaises. Plus prosaïquement, cela résout le problème qui se pose en français en prenant l'initiale des mots car outil et objet ont la même initiale !

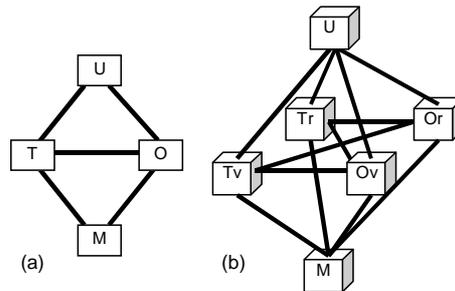


Figure 56 : Les principales entités et leurs relations.

Si l'utilisateur se trouve clairement dans le monde réel et le modèle interne de l'application dans le monde virtuel, dans le cas de la Réalité Mixte, les outils peuvent être réels ou virtuels, et peuvent agir sur des objets du domaine réels ou virtuels. La Figure 56b montre les 4 cas possibles de relation entre un outil réel (Tr) ou virtuel (Tv) et un objet réel (Or) ou virtuel (Ov). Ces cas ne sont pas exclusifs car la frontière entre les mondes réels et virtuels peut passer à travers un outil ou un objet mixte.

Nous définissons les objets réels et virtuels ainsi (d'après [Milgram 1994a]) :

- Les **objets réels** sont des objets qui ont une **existence concrète et objective** et qui peuvent être observés directement,
- Les **objets virtuels** sont des objets qui existent par essence ou effet mais non formellement ou concrètement; pour être vus, ils doivent être simulés.

*Prise en compte du travail collaboratif*

Les schémas de la Figure 56 montrent les principales relations et les entités dans le cas de l'interaction d'un utilisateur. IRVO permet de représenter simultanément l'interaction de plusieurs utilisateurs ce qui permet de prendre en compte le travail collaboratif. Par exemple la Figure 57 montre deux cas d'interactions multi-utilisateurs :

- dans la Figure 57a les utilisateurs partagent les mêmes outils et les mêmes objets du domaine ; il s'agit d'un cas de travail collaboratif co-localisé.
- dans la Figure 57b les utilisateurs utilisent leurs propres outils et objets du domaine : cela peut tout aussi bien représenter du travail collaboratif co-localisé ou à distance. Dans cette figure, il se peut qu'un objet comme O1 soit réel (et augmenté) et que les autres (O2 et O3) soient virtuels. Ainsi les utilisateurs U2 et U3 peuvent voir les résultats des actions de U1 et peuvent agir avec des outils T2 et T3 (réels ou virtuels), la cohérence de l'ensemble étant assurée par le modèle interne M.

Bien sûr, de nombreux autres cas « intermédiaires » peuvent être envisagés : par exemple, un seul objet du domaine de la tâche partagé et plusieurs outils individuels, etc.

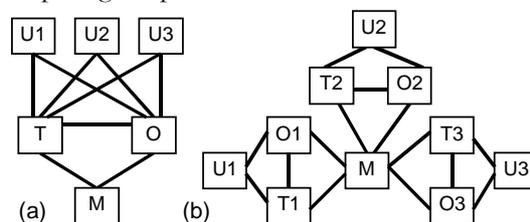


Figure 57 : Exemples de modélisations du travail collaborative : a: outils et objets partagés; b: outils et objets propres à chaque utilisateur.

*Passage monde réel – monde virtuel*

IRVO modélise des entités qui se trouvent soit dans le monde réel (utilisateurs et objets) soit dans le monde virtuel (objets et modèle interne). Pour passer d'un monde dans l'autre il faut donc des dispositifs d'entrée/sortie qui permettent l'échange d'informations entre les deux mondes. Dans le modèle IRVO nous les modélisons sous la forme de **transducteurs** et ce sont les seules entités qui se trouvent « à cheval » sur la frontière séparant les mondes réel et virtuel. Ce sont des entités secondaires dans le modèle IRVO car elles ne participent pas directement à l'interaction entre les utilisateurs et le système de Réalité Mixte. Cependant, les caractéristiques de ces transducteurs (et surtout leurs limitations) ont une influence sur la qualité de l'interaction au point de la rendre difficilement acceptable voire totalement impossible dans certains cas ; ainsi, même si leur présence semble secondaire dans une première phase d'analyse, elle devient incontournable pour une modélisation fine d'un système de Réalité Mixte en phase de conception afin de pouvoir explorer plusieurs solutions techniques alternatives.

## VI.2 Formalisme graphique d'IRVO

Nous présentons dans cette partie le formalisme graphique du modèle IRVO. Il représente les entités du modèle par une notation graphique qui fait apparaître certaines propriétés de ces entités. Les relations entre ces entités sont également caractérisées et représentées graphiquement sous forme de flèches.

### VI.2.1 Frontières

Les frontières (Figure 58) ne sont pas des entités mais elles permettent de représenter des propriétés des entités de manière topologique : il suffit qu'une entité se trouve dans un lieu délimité par des frontières pour quelle possède l'une des valeurs possible de la propriété représentée par la frontière.

Deux types de frontières sont représentés :

- La frontière monde réel/monde virtuel,
- Les frontières entre différents lieux du monde réel.

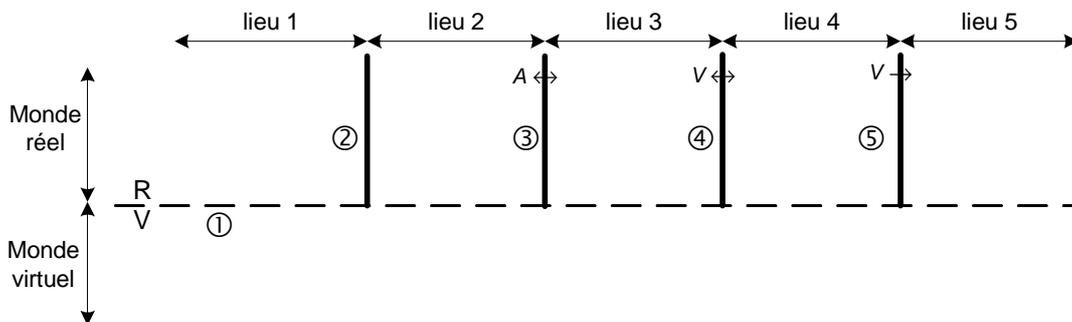


Figure 58 : Représentation des frontières.

#### VI.2.1.1 Frontière entre les mondes réel et virtuel

Nous représentons la frontière entre les mondes réel et virtuel par une ligne horizontale pointillée avec, de part et d'autre de la ligne, les initiales 'R' et 'V' (item ①, Figure 58). La nature (réelle ou virtuelle) des entités est donc simplement précisée par leur placement par rapport à cette frontière. La frontière est représentée en ligne pointillée car elle est « aisément » franchissable avec l'aide de transducteurs (Cf. § VI.2.6).

### VI.2.1.2 Frontières entre différents lieux du monde réel

Nous représentons la frontière entre différents lieux du monde réel (Figure 58, item ②) par une ligne verticale pleine pour montrer qu'elle est infranchissable. Cependant, si 2 lieux sont « assez proches » l'un de l'autre, on peut avoir des cas de franchissement partiels :

- Les personnes peuvent communiquer entre elles par la voix sans se voir (item ③, Figure 58) ; par exemple elles sont de chaque côté d'une machine qui les cache visuellement. La flèche double et la lettre A signifient que des flux audio (A) peuvent franchir la frontière.
- Les personnes peuvent se voir (par exemple à travers une vitre) mais sans pouvoir se parler (item ④, Figure 58). La flèche double et la lettre V signifient que des utilisateurs peuvent voir (V) à travers la frontière.
- L'item ⑤ de la Figure 58 montre une représentation d'un miroir sans teint. La flèche simple et la lettre V signifient qu'une personne se trouvant dans le lieu 5 peut voir les objets se trouvant dans le lieu 4, c-à-d qu'une flèche provenant du lieu 4 peut arriver sur le canal 'V' d'un utilisateur du lieu 5 (Cf. § VI.2.3 sur la modélisation des canaux de communication des utilisateurs).

### VI.2.1.3 Frontières dans le monde virtuel

Nous considérons qu'il n'y a pas de frontière dans le monde virtuel car les réseaux fournissent un environnement ubiquitaire appelé communément CyberSpace. Bien sûr, cela n'est pas totalement transparent aux utilisateurs, car les retards apportés essentiellement par les délais de transmission peuvent être perceptibles et même gênant en particulier pour les transmissions audio.

## VI.2.2 Caractéristiques communes aux entités

Les entités possèdent plusieurs propriétés communes qui sont représentées graphiquement sur le schéma par :

- Un **symbole graphique**, dépendant du type de l'objet.
- Une **étiquette** (placée à l'extérieur du symbole, en haut à droite) qui précise le type de l'objet (U=utilisateur, T=outil, O=objet du domaine, M=modèle interne, S=senseur, E=effecteur). Cette étiquette doit être unique dans un schéma pour pouvoir faire référence sans ambiguïté à n'importe quelle entité. Ainsi, s'il existe plusieurs entités du même type dans le schéma, elles devront être numérotées (par exemple : U1, U2, U3 pour 3 utilisateurs).
- Un **nom** (placé à l'intérieur du symbole, en haut au centre) qui doit être significatif pour la tâche considérée (par exemple l'utilisateur U est nommé « Chirurgien » pour un système d'opération médicale assistée par ordinateur).
- Une **propriété de mobilité** absolue ou relative (Cf. § VI.2.8) (placée à l'intérieur du symbole, en bas à droite). Cette indication est facultative.
- Sa **nature** (réelle ou virtuelle) représentée par sa position topologique sur le schéma par rapport à la frontière réel/virtuel (Cf. § VI.2.1.1). Seuls les transducteurs ont la capacité d'être « à cheval » sur la frontière.
- Son **lieu physique** (pour les entités du monde réel seulement) représentée par sa position topologique sur le schéma par rapport aux frontières délimitant les lieux (Cf. § VI.2.1.2)

À chaque entité est également associé un ensemble de propriétés propres dont seules quelques unes sont représentées graphiquement sur le schéma afin de ne pas le surcharger. Cependant, dans les évolutions futures du formalisme, il est envisagé de pouvoir en ajouter certaines sur la représentation graphique si cela s'avère utile.

### VI.2.3 Utilisateurs

Un utilisateur est représenté comme en Figure 59.

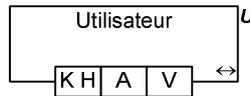


Figure 59 : Représentation d'un utilisateur.

En plus des propriétés communes à toutes les entités, un utilisateur est principalement caractérisé par les canaux de communication qu'il peut utiliser ; nous considérons :

- Le canal Visuel (V) :
  - en entrée : il s'agit du sens de la vue,
  - en sortie : bien que rarement utilisé, la direction du regard peut être détectée par un oculomètre (ou *eye-tracker*) et servir d'information de commande pour le dispositif informatique.
- Le canal audio (A) :
  - en entrée : sens de l'ouïe,
  - en sortie : la voix (parole, chant, ...).
- Le canal kinesthésique/haptique (KH) :
  - en entrée : sens du toucher, par la main bien sûr, mais aussi par l'ensemble du corps (perception de la température, du vent, etc...)
  - en sortie : manipulation, geste, ... En général, les gestes sont effectués par la main dominante mais on peut aussi envisager l'utilisation des 2 mains (interaction bi-manuelle), des pieds, des mouvements de la tête, des mimiques du visage, etc... La position du corps dans l'espace peut aussi constituer une possibilité d'interaction.

Nous nous limitons à ces sens principaux. En effet, le goût et l'odorat pourraient être envisagés si des périphériques d'entrée/sortie *ad hoc* existaient au-delà de simples prototypes de recherche. Dans cette éventualité nous pourrions donc rajouter :

- Le canal T (comme *Taste* en anglais) : en entrée seulement : sens du goût.
- Le canal S (comme *Smell* en anglais) : en entrée seulement : sens de l'odorat.

Il est difficile de parler de ces canaux en sortie car, même si le corps humain émet en permanence des odeurs qu'il est tout à fait possible de détecter par des capteurs biochimiques, il n'est pas possible de le faire de manière consciente et volontaire et cette émission est donc de portée limitée comme moyen d'interaction (Cf. § III.1).

Il existe encore d'autres dispositifs d'interaction, encore très expérimentaux : les interfaces cerveau-ordinateur (BCI – Brain-Computer Interfaces, Cf. § III.3.4). Le principe est de capter les signaux EEG (électro-encéphalogrammes) des utilisateurs à partir d'électrodes posées sur leur scalp afin de détecter les états mentaux et s'en servir pour commander des robots ou des claviers virtuels par exemple [Millán 2004]. Pour prendre en compte ces dispositifs nous pourrions donc ajouter un canal supplémentaire 'B' en sortie (B=*Brain*).

Nous verrons plus loin que les relations entre l'utilisateur et les autres entités partent et arrivent sur les canaux ce qui précise la nature de l'échange d'information.

Nous avons donc choisi de caractériser l'utilisateur dans IRVO par ses capacités de communication vers le monde extérieur soit par son action, soit par la perception de l'environnement par ses 5 sens, ce qui est suffisant dans une première approche pour décrire l'interaction avec un système mixte. La nature des échanges est précisée plus finement par les caractéristiques des relations soit d'action, soit de perception qui partent ou arrivent à l'utilisateur (Cf. § VI.2.7.8). Pour une modélisation plus

complète de l'utilisateur on se reportera aux nombreux modèles de l'utilisateur que proposent les approches basées sur les modèles que nous aborderons au Chapitre VII.

Dans le modèle IRVO, les utilisateurs sont supposés toujours être dans le monde réel ce qui ne permet pas, par exemple, de représenter des avatars contrairement au formalisme de Renevier (Cf. § V.2.1) qui permet de modéliser des utilisateurs représentés numériquement.

## VI.2.4 Objets

### VI.2.4.1 Objets réels et virtuels

Les objets sont représentés graphiquement par un rectangle. Les outils et les objets du domaine sont simplement distingués par une étiquette (T ou O) placée au coin haut-droit du rectangle (Figure 60a et b).

En plus, comme nous l'avons indiqué, les objets sont les seules entités d'IRVO qui peuvent être réelles ou virtuelles ; il y a 2 manières redondantes de préciser cela :

- Dans l'étiquette, en la suffixant par 'r' ou 'v' ce qui donne 4 cas possibles : Tr, Tv, Or, Ov.
- Dans la position de l'objet par rapport à la frontière entre le monde réel et virtuel (Cf. § VI.2.1.1).

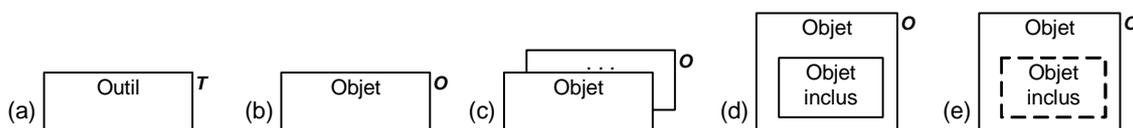


Figure 60 : Représentations des outils et des objets.

Il est possible de représenter des collections d'objets du même type sous la forme d'un empilement (Figure 60c). Cette représentation implique que tous les objets partagent les mêmes propriétés.

Lorsqu'un objet est un composant d'un autre objet, il peut être représenté comme en Figure 60d. Cette représentation implique :

- Que l'objet inclus est lié de manière fixe à l'objet englobant par défaut (en effet, on verra au paragraphe VI.2.8.2 que des notations peuvent préciser d'autres possibilités).
- Que si l'objet englobant est perceptible (c-à-d visible, touchable, audible, etc...) par un utilisateur, l'objet inclus est lui aussi perceptible dans les mêmes conditions. Dans le cas contraire, l'objet inclus est dessiné en ligne pointillée pour signifier qu'il est masqué par l'objet englobant (Figure 60e).

Les représentations de la Figure 60c à la Figure 60e sont bien sûr également applicables aux outils. Par exemple, on peut représenter une palette d'outils graphiques composée de plusieurs outils : 'crayon', 'pinceau', 'gomme', etc...

D'autres caractéristiques non représentées graphiquement sont en particulier la taille et le poids de l'artefact ce qui peut être important si l'utilisateur doit porter ou déplacer l'objet pour accomplir la tâche. De même, sa composition détaillée (matière, texture) et sa structure (rigide/souple) peuvent aussi intervenir. Dans la plupart des cas le nom de l'objet ainsi que le contexte peut apporter suffisamment de renseignement : par exemple, un outil 'stylo' est sans difficulté manipulable par tout utilisateur alors qu'un objet 'table' n'est en principe pas déplaçable sauf si le contexte est l'agencement de bureaux paysager.

Pour les outils et les objets virtuels, seules les caractéristiques nécessaires à l'accomplissement de la tâche ont besoin d'être modélisées. Ainsi, une table sera caractérisée uniquement par ses paramètres géométriques pour une application de dessin alors que son poids, sa matière, sa texture, etc. seront importants pour une application d'agencement de bureaux paysager.

### VI.2.4.2 Objets mixtes

Les objets mixtes sont « à cheval » sur les mondes réels et virtuels. Il serait donc simple de les représenter traversés par la frontière R/V. Cependant, cette approche est trop globale et n'apporterait pas grand-chose pour la conception du système.

Ainsi, nous choisissons de représenter un objet mixte dans le formalisme d'IRVO par 2 entités : un objet réel (étiqueté 'Or' pour un objet du domaine ou 'Tr' pour un outil) et un objet virtuel (étiqueté 'Ov' pour un objet du domaine ou 'Tv' pour un outil), l'ensemble étant compris dans un groupe dessiné en trait pointillé et étiqueté 'O' ou 'T' (Figure 61).

Cette représentation permet de bien distinguer les différentes composantes de l'objet mixte et donc de les modéliser complètement indépendamment, mais aussi de bien faire ressortir que l'objet mixte est essentiellement une abstraction qui n'a pas d'existence tangible mais « existe » surtout par la perception que les utilisateurs en ont (Cf. paragraphe VI.2.7.7). Un groupe n'est donc pas une entité mais une relation qui exprime le lien entre deux entités (ou plus), d'où sa représentation en traits pointillés.

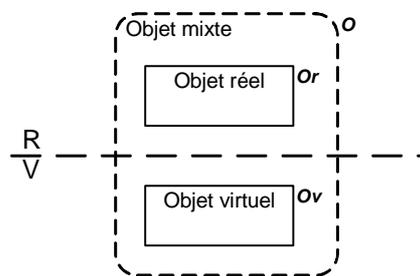


Figure 61 : Représentation d'un objet mixte

De manière plus précise, nous distinguons les deux cas (Cf. la taxonomie de Dubois – § II.3.5) :

- La Réalité Augmentée où les objets réels sont augmentés par des données virtuelles : nous étiquetons dans ce cas l'objet réel 'Or' et l'augmentation virtuelle 'O+v'. L'objet complet est donc  $O=Or+v$ .
- La Virtualité Augmentée où les objets virtuels sont augmentés par des objets physiques : nous étiquetons dans ce cas l'objet virtuel 'Ov' et l'augmentation réelle 'O+r'. L'objet complet est donc  $O=Ov+r$ .

Ces notations montrent bien qu'un objet augmenté n'est pas à 50 % réel / 50 % virtuel. Il s'agit soit d'objets réels complétés par un peu de virtuel, soit d'objets virtuels complétés par un peu de réel.

#### Cas des outils

Dans IRVO, nous considérons que ceci est applicable aux outils qui sont également des objets avec les notations correspondantes :

- Des outils réels augmentés  $Tr+v$ , qui sont composés d'un outil réel  $Tr$  et de l'augmentation virtuelle  $T+v$ ,
- Des outils virtuels augmentés  $Tv+r$ , qui sont composés d'un outil virtuel  $Tv$  et d'une augmentation réelle  $T+r$ .

Dans le cas particulier des outils, on peut distinguer deux types d'augmentation :

- Une augmentation de la perception de l'outil par l'utilisateur, comme tout objet,
- Une augmentation de son action : l'outil agit simultanément dans les deux mondes, c'est-à-dire que la partie réelle ( $Tr$ ) a une action sur les objets du monde réel, et que la partie virtuelle ( $Tv$ ) a une action sur les objets du monde numérique

Les deux types d'augmentations ne sont bien sûr pas exclusifs et un outil peut donc être « doublement » augmenté (action et perception).

### Autres cas d'augmentation

Les augmentations expriment un lien entre des objets réels et virtuels. Elles peuvent aussi exprimer des liens entre des objets réels locaux et des objets réels distants perçus par des moyens de transmissions numériques (Trevisan *et al.* les appellent objets numériques réels [Trevisan 2003]). Ainsi les objets mixtes peuvent aussi être « à cheval » sur une frontière physique séparant deux lieux du monde réel. Un exemple de telle augmentation est le cas du DoubleDigitalDesk que nous détaillons au § VI.3.1.3.

Renevier dans son formalisme (§ V.2.1) détaille tous ces cas en distinguant le ‘caractère original’ d’un objet. Ainsi, d’une part des objets physiques peuvent représenter d’autres objets physiques ou des objets numériques et, d’autre part, les objets numériques peuvent représenter des objets physiques ou d’autres objets numériques. Ce dernier cas n’existe pas véritablement dans le modèle IRVO car nous avons aboli la distance dans le monde virtuel ce qui fait qu’il est possible de percevoir le même objet virtuel quelque soit la distance.

### Création et destruction des augmentations

Dans son formalisme, Renevier exprime graphiquement la création et la destruction des augmentations (par une flèche double), en distinguant deux cas :

- Un mode actif, où l'utilisateur crée dynamiquement des associations entre des objets réels ou virtuels
- Un mode passif, où c'est le système informatique qui gère les liens entre les deux mondes.

Dans IRVO, seul le premier cas est représentable, car il s'agit d'une tâche explicite réalisée par l'utilisateur et il y a interaction avec le système. Le deuxième cas, ne faisant pas intervenir l'utilisateur, n'a pas vocation à être représenté par le modèle IRVO.

## VI.2.5 Modèle interne

Le modèle interne (M) est représenté comme en Figure 62.

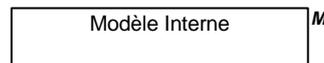


Figure 62 : Représentation du Modèle interne de l'application.

Cette entité représente le modèle comportemental du logiciel et gère directement les outils et les objets virtuels. Il capte également (à l'aide des senseurs, Cf. § VI.2.6) l'état des outils et objets réels et peut également agir sur eux par des effecteurs. Le modèle interne maintient donc la cohérence entre les différents artefacts et est responsable de l'augmentation des objets. C'est, par exemple, en suivant la position spatiale d'un objet physique (à l'aide d'un localisateur) qu'il peut mettre à jour l'objet virtuel de telle manière que l'utilisateur perçoive toujours l'augmentation virtuelle correctement alignée avec l'objet réel.

Le modèle interne assure également la rémanence des objets virtuels entre deux sessions de travail en enregistrant l'état courant sur un support permanent.

Concrètement, en faisant le parallèle avec le modèle Arch (Cf. § IV.2.1.2), le modèle interne englobe les 4 couches suivantes :

- Le Noyau Fonctionnel de l'application,
- L'adaptateur du Noyau Fonctionnel,
- Le Contrôleur de Dialogue,
- La Présentation & Interaction Abstraites.

Les objets virtuels représentés dans le modèle IRVO apparaissent donc comme les entités de la couche Présentation & Interaction Concrètes.

Dans un système purement réactif, le modèle interne a donc un rôle passif vis-à-vis de l'interaction de l'utilisateur avec le système. Cependant, il peut être aussi actif dans des dispositifs utilisant des agents cognitifs qui ont une capacité de traitement propre et qui agissent directement sur les objets (et éventuellement les outils) indépendamment de l'utilisateur.

Dans un système multi-utilisateurs, le modèle interne est aussi responsable de la propagation à tous les utilisateurs des modifications apportées par un utilisateur. Eventuellement, selon le niveau de relâchement du WYSIWIS ou les rôles de différents utilisateurs, ces modifications ne sont pas forcément rendues perceptibles au niveau des objets du domaine (virtuels ou réels).

Le modèle interne **est toujours présent même s'il n'est pas nécessairement représenté dans les schémas** car il n'est pas directement impliqué dans la boucle action-perception de l'utilisateur avec les objets. Il n'y a qu'un exemple parmi les applications que nous avons recensées, le DoubleDigitalDesk (Cf. § II.2.1.3), où il n'y a pas d'application informatique et donc pas de modèle interne.

## VI.2.6 Transducteurs

Pour communiquer entre les mondes réel et virtuel, l'information provenant du monde réel doit être transformée en données numériques, ce qui est réalisé par des **senseurs**<sup>15</sup> (Figure 63a) ; l'opération inverse est réalisée par des **effecteurs** ou **actuateurs** (Figure 63b). Les transducteurs sont toujours placés à la frontière entre les mondes réel et virtuel. Leurs représentations sur un schéma sont facultatives car ils n'interviennent pas directement dans l'interaction à un premier niveau de modélisation. Cependant, les caractéristiques des transducteurs (et surtout leurs limitations) ont une influence directe sur l'interaction au point de la rendre impossible dans certains cas ; ainsi, pour une modélisation fine d'un système de Réalité Mixte en phase de conception, il devient essentiel de les prendre en compte afin de pouvoir explorer plusieurs solutions techniques alternatives.

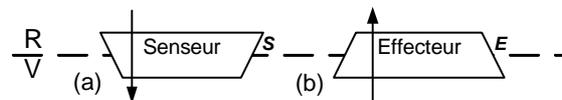


Figure 63 : Représentations des transducteurs.

### VI.2.6.1 Caractéristiques des transducteurs

La première caractéristique d'un transducteur est l'opération de transformation qu'il effectue. Un transducteur intervient dans une relation U—T ou U—O mais aussi T—O pour passer du monde réel au monde virtuel ou l'inverse. C'est la raison pour laquelle, sur la représentation graphique, les relations traversent les transducteurs pour montrer qu'ils effectuent « juste » une modification de la nature de la relation (passage réel/virtuel) sans changer les caractéristiques mêmes de la relation (Cf. § VI.2.7.6).

Comme toute opération de mesure, cette opération de conversion n'est évidemment pas parfaite. Par exemple, pour un localisateur spatial, la résolution, la précision, la justesse, la fidélité, etc. sont des paramètres essentiels à la caractérisation fine du transducteur car ils auront un impact direct sur l'utilisabilité du système complet.

Physiquement ils peuvent être liés (Cf. taxonomie de Wendy MacKay – § II.3.1) :

- aux objets ou aux outils (par exemple un capteur d'accélération),
- à l'utilisateur (par exemple un gant ou un casque de visualisation)
- ou encore ils peuvent être fixés à l'environnement (par exemple, un vidéoprojecteur ou un clavier).

<sup>15</sup> Nous reprenons les termes de *senseurs* et d'*effecteurs* tels que Stephen Ellis *et al.* les définissent pour décrire les dispositifs d'entrées-sorties dans les environnements virtuels [Ellis 1997, p. 170].

Cette information est donnée de manière synthétique sur la représentation graphique par le symbole de contrainte spatiale placé en bas à droite comme pour les autres entités (Cf. § VI.2.8).

La notion de transducteur est très proche de celle d'adaptateur du modèle ASUR (Cf § IV.1.7.1) et en général les deux concepts peuvent être considérés comme équivalents. Cependant, les adaptateurs d'ASUR sont les dispositifs d'entrée/sortie eux-mêmes alors que les transducteurs représentent la fonctionnalité. Cette discussion est reprise ci-dessous au § VI.2.6.2.

#### *Caractéristique de partage des transducteurs :*

Une deuxième caractéristique des transducteurs (et que nous ne représentons pas de manière explicite dans le formalisme graphique) est la possibilité de partage par plusieurs relations, c-à-d le nombre « d'utilisateurs » simultanés du transducteur. Par exemple, un capteur d'accélération fixé à un objet ne peut mesurer l'accélération que de **ce seul objet**. Par contre, un capteur de position basé sur une analyse d'image prise par une caméra, peut mesurer la position géométrique de **plusieurs objets simultanément**. Pour les effecteurs, nous avons les mêmes cas possibles : un casque de visualisation est foncièrement individuel (mais peut afficher simultanément l'image de plusieurs objets virtuels), alors que l'image affichée par un vidéo projecteur peut être vue par plusieurs utilisateurs en même temps. Sauf cas très particuliers, un transducteur lié à une autre entité n'est pas partageable. Seuls les transducteurs liés à l'environnement sont éventuellement partageables si d'autres contraintes géométriques ne limitent pas le nombre d'utilisateurs potentiels.

Cette caractéristique est aussi présente dans le modèle ASUR.

#### *Caractéristique du temps de traversée d'un transducteur*

Une caractéristique primordiale des transducteurs est leur temps de traversée. Il s'agit du délai introduit par le dispositif d'entrée ou de sortie pour effectuer la conversion entre le monde réel et le monde virtuel (ou l'inverse pour un effecteur). Il s'agit tant des retards introduits par l'électronique de mesure que par les logiciels de traitement.

Par exemple, dans le cas de la souris ces délais sont faibles car l'électronique et le logiciel sont simples. Par contre, pour des systèmes de localisation 3D basés sur des caméras et un logiciel de reconnaissance de formes, le délai apporté par l'algorithme de traitement peut devenir très grand, de l'ordre de la seconde. Dans ce cas, ce délai peut devenir perceptible à l'utilisateur et compromettre totalement l'utilisabilité du système. Par exemple, si l'on considère un système de Réalité Mixte où un objet réel en mouvement est localisé par un tel dispositif puis augmenté par un objet virtuel projeté en temps réel par un vidéoprojecteur, l'augmentation semblera toujours être « à la traîne » de l'objet et l'effet de fusion perceptuelle annihilé. Si cet effet peut être considéré comme « simplement gênant » dans l'exemple précédent, il peut être totalement rédhibitoire si c'est l'utilisateur qui déplace l'objet virtuel car la boucle sensori-motrice est totalement perturbée et il sera impossible à l'utilisateur d'atteindre la cible.

Comme il est impossible de retarder le monde réel (!), ce n'est qu'en réduisant les délais apportés par les transducteurs (et donc en changeant de technologie) que l'on peut améliorer le système d'où l'importance de comparer plusieurs solutions alternatives dans la phase de conception.

#### **VI.2.6.2 Distinction entre transducteur et dispositif d'entrée/sortie :**

Tel que défini, un transducteur semble similaire à un dispositif d'entrée/sortie, ce qui est vrai dans la plupart des cas simples. Cependant, la notion de transducteur est associée à **la fonction de conversion** de valeurs physiques en données numériques (pour un senseur) et de données numériques en grandeurs physiques (pour un effecteur). Comme nous l'avons expliqué au § III.4.1, un dispositif d'entrée/sortie est composé d'un transducteur et d'une partie réelle qui peut aussi être considérée comme un outil au sens de notre modèle. Ainsi, il nous semble important de séparer les

deux notions dans notre modèle car, du point de vue de l'utilisateur, le retour haptique que procure le dispositif peut être important dans le contrôle que l'utilisateur a du dispositif.

Ceci constitue une différence fondamentale entre les transducteurs du modèle IRVO et les adaptateurs du modèle ASUR qui sont la représentation complète des dispositifs d'entrée/sortie.

Prenons par exemple le cas d'un périphérique de pointage simple comme la souris. Si nous voulons le modéliser complètement avec IRVO, nous obtenons le schéma de la Figure 64a. Ainsi la souris elle-même est la partie réelle de l'outil de pointage qui est manipulée par l'utilisateur. Les mouvements sont transmis au pointeur (Tv) grâce au senseur (S). Le pointeur (Tv) est donc spatialement lié à la souris (Tr) ce qui est montré par la symbolisation "Tr x" même si ce n'est que partiellement vrai (il suffit de soulever la souris de la table pour cela !). La perception de cet outil par l'utilisateur est donc effectuée de 2 manières différentes : d'une part, de manière haptique par le retour tactile et, d'autre part, de manière visuelle par la visualisation du pointeur sur l'écran (et aussi par la vision de la souris sur la table). Cependant, seule la perception visuelle du pointeur sur l'écran est réellement utile pour pouvoir se servir de la souris, la perception haptique étant secondaire (en tout cas pour un utilisateur expérimenté). Finalement, seule la partie virtuelle de l'outil (c-à-d le pointeur) est véritablement perçue comme étant l'outil et nous pouvons donc utiliser une représentation simplifiée présentée en Figure 64b en fusionnant la souris (Tr) et le senseur (S) et en négligeant les autres perceptions.

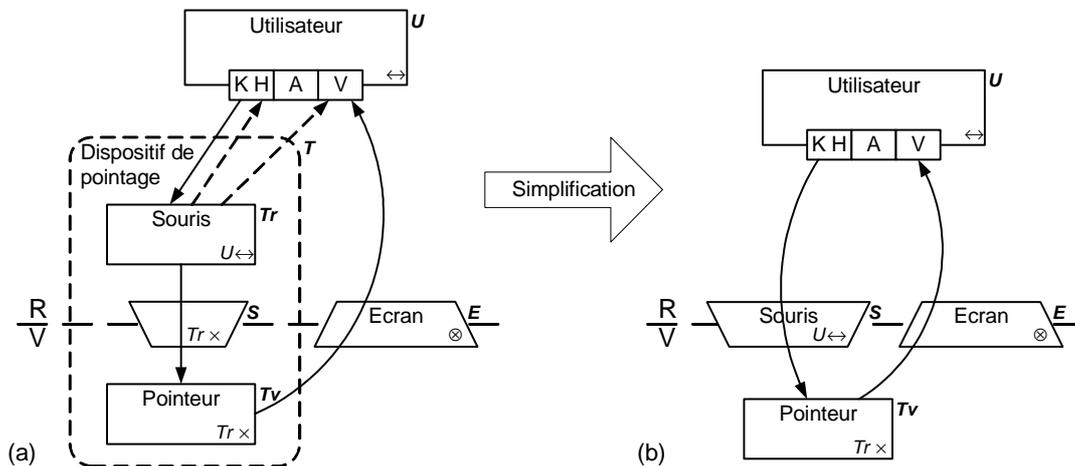


Figure 64 : Représentation d'un dispositif de pointage simple avec IRVO. (a) : modèle complet ; (b) : représentation simplifiée en fusionnant l'outil réel et le transducteur.

## VI.2.7 Relations

Les relations représentent les échanges d'information entre deux entités. Elles sont visuellement représentées par des flèches qui peuvent comporter des annotations textuelles pour préciser certaines propriétés.

### VI.2.7.1 Relations partant ou aboutissant à un utilisateur vers un outil ou un objet

Les relations associées à un utilisateur sont :

- Des perceptions du monde extérieur par l'utilisateur : elles sont représentées par des flèches qui aboutissent à l'utilisateur. Elles aboutissent sur un des canaux de l'utilisateur (KH, A ou V) ce qui renseigne sur le sens perceptif mis en jeu (relations ②, ③, ④, Figure 65). Comme discuté au paragraphe VI.2.3, les canaux "T" (comme *taste*) et "S" (comme *smell*) peuvent être ajoutés au symbole de l'utilisateur si cela s'avère nécessaire.
- Des actions de l'utilisateur sur le monde extérieur ; elles sont représentées par des flèches qui partent de l'utilisateur. A nouveau, le canal (KH, A ou V) d'où part la flèche renseigne sur le type d'action de l'utilisateur (relation ①, Figure 65). Pour le canal KH, en particulier si

plusieurs relations en partent, il peut être important de préciser sur le lien le type d'action ou l'organe impliqué (par exemple 'main droite' / 'main gauche' pour une interaction bi-manuelle — Cf. § VI.3.1.5)

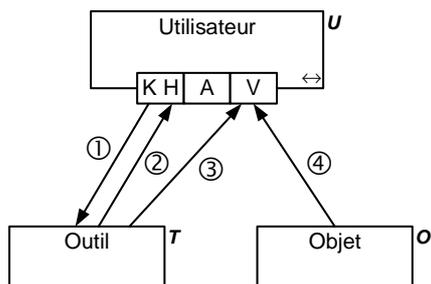


Figure 65 : Relations entre un utilisateur et les objets.

### VI.2.7.2 Relations entre utilisateurs

Il est tout à fait possible d'avoir des flèches partant d'un utilisateur et aboutissant à un autre utilisateur.

Il existe 2 cas :

- L'un des 2 utilisateurs est passif dans l'interaction (il a le statut d'objet au sens d'IRVO) ; on retombe dans le cas précédent et il peut s'agir, par exemple, de la perception visuelle d'une personne par une autre (Figure 66, relation ①), ou d'une action physique d'un utilisateur sur un autre (Figure 66, relation ②). Dans ce cas la flèche part (resp. aboutit) d'un canal (KH, A ou V) de l'utilisateur 'actif' et arrive (resp. part) directement de l'utilisateur 'passif'.
- Les 2 utilisateurs ont un rôle symétrique ; il s'agit alors d'une communication inter-individuelle. Par exemple une conversation peut être modélisée par une flèche bidirectionnelle entre les 2 canaux 'A' des 2 utilisateurs (Figure 66, relation ③).

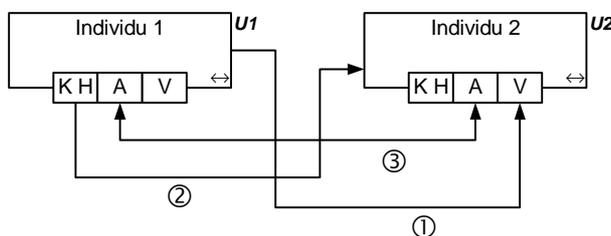


Figure 66 : Représentation de relations interindividuelles

Le but d'IRVO n'est pas de faire une modélisation fine de ces actions ou interactions entre individus. En effet, par défaut, si 2 utilisateurs sont présents dans le même lieu (sans frontière physique), alors toute forme de communication est implicitement possible. Il n'est donc intéressant de les représenter que pour insister sur une interaction interpersonnelle qui est importante pour la tâche considérée.

Par contre, si les deux utilisateurs sont dans 2 lieux différents (donc séparés par une frontière physique), la représentation de la relation interindividuelle qui passe à travers le monde virtuel par des transducteurs *ad hoc*, permet de représenter une communication médiatisée par un système de télécommunication audio-visuelle (Figure 67). En général, si cela ne prête pas à confusion, les transducteurs peuvent être omis comme dans l'exemple du DoubleDigitalDesk (Cf. § VI.3.1.3).

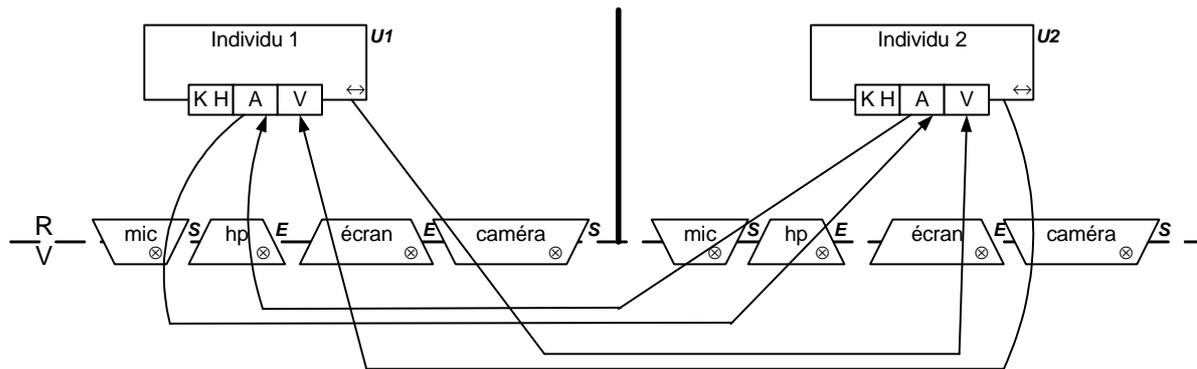


Figure 67 : Exemple de représentation d'une communication médiatisée avec IRVO

### VI.2.7.3 Relations entre les outils et les objets du domaine

Entre des outils et des objets, une relation représente une action d'un outil sur un objet du domaine. Même si le sens de la flèche montre une action outil  $\rightarrow$  objet, une réaction de l'objet sur l'outil est implicitement présente ; cette réaction peut être perçue en retour par l'utilisateur à travers la perception de l'outil (exemple : l'appui d'un stylo sur une surface est perçue par l'utilisateur à travers le stylo comme une résistance au mouvement).

### VI.2.7.4 Relations entre le modèle interne, les outils et les objets du domaine

Le modèle interne est lié directement aux outils et objets du domaine virtuels et indirectement aux outils et objets du domaine réels par l'intermédiaire des transducteurs (Cf. § VI.2.5). Les relations sont généralement bidirectionnelles car il s'agit aussi bien de capturer l'état des objets que de les modifier pour refléter :

- Soit des changements des objets du domaine suite à l'action des outils,
- Soit des changements dus aux actions des autres utilisateurs à distance dans un système collaboratif,
- soit des changements suite à des décisions prises par des agents cognitifs.

Concrètement, le modèle interne est lié à toutes les autres entités. Cependant, pour ne pas alourdir les schémas il est normal de ne représenter que les relations qui sont vraiment importantes pour la tâche modélisée.

### VI.2.7.5 Relations principales et secondaires

Les relations sont normalement représentées en traits continus. Seules les relations intervenant directement dans l'interaction entre l'utilisateur et le système sont représentées. Cependant, des relations qui sont peu importantes vis-à-vis de la tâche actuelle peuvent être représentée en traits pointillés. Par exemple le feedback haptique d'une souris n'est pas important pour la déplacer car le retour est principalement assuré en regardant le pointeur sur l'écran (Cf. Figure 64). En général, pour ne pas surcharger le schéma, il est plus simple de les omettre.

Cette notion est également une caractéristique présente dans ASUR [Dubois 2001b] sous le nom de poids et qui peut prendre 4 valeurs : aucun, périphérique, faible, élevé. Par contre, cette caractéristique n'est pas visualisée sur la notation graphique.

Une notion similaire est aussi prise en compte par [Graham 2000] pour décrire les entités d'un système de réalité mixte sous le nom d'attention reçue (*attention received*) : cette notion spécifie l'attention que l'utilisateur porte actuellement à une entité selon une échelle à 3 valeurs : aucune, périphérique et élevée.

### VI.2.7.6 Statut des transducteurs vis-à-vis des relations

Les transducteurs sont traversés par les relations pour montrer qu'ils changent la nature de l'information (réel / virtuel) mais pas la signification et donc ces transducteurs ne participent pas directement à la boucle d'interaction. Par exemple les mouvements d'un stylo sur des pages sont transformés en coordonnées x et y dans le monde virtuel qui sont acceptées en entrée d'un objet virtuel.

### VI.2.7.7 Fusions de relations

Dans le cas de relation symétriques, comme la conversation entre 2 individus (Figure 66, relation ③), on peut fusionner les 2 flèches aller et retour par une flèche bidirectionnelle.

L'autre cas utilise une entité particulière pour exprimer la fusion de la perception d'un objet et de son 'augmentation'. Les 2 relations de perception des composantes réelles et virtuelles par l'utilisateur sont fusionnées par le symbole  $\oplus$  ce qui exprime la perception « augmentée » de l'objet mixte par l'utilisateur (Figure 68).

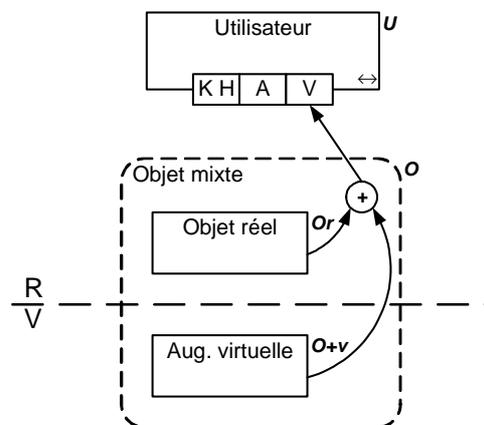


Figure 68 : Représentation de la perception d'un objet mixte par un utilisateur

### VI.2.7.8 Caractéristiques détaillées des relations

Par leur représentation graphique les relations sont caractérisées par :

- Un émetteur : identifié par le talon de la flèche.
- Un récepteur : identifié par la pointe de la flèche.
- Un médium, pour les relations issues ou aboutissant à U grâce à la précision du canal : graphique (V), sonore (A), kinesthésique/haptique (KH), voire gustatif (T) et olfactif (S) si nécessaire.
- La nature implicitement donnée par l'émetteur ou le récepteur : action ( $U \rightarrow x$ ,  $T \rightarrow O$ ), perception ( $x \rightarrow U$ ) ; par analogie on peut aussi parler de perception pour les relations  $x \rightarrow M$  et actions pour  $M \rightarrow x$ .
- Une importance relative : relation principale (trait plein) ou secondaire (trait pointillé).
- Une physicalité/virtualité de l'information véhiculée : par le fait que la flèche est dans le monde réel ou virtuel (notons qu'elle peut être dans les 2 mondes à la fois en traversant un transducteur).
- Une annotation textuelle facultative qui peut préciser la nature de l'émetteur (par exemple 'main droite' pour une relation  $U \rightarrow T$ ), ou le type d'information échangée (par exemple 'position 2D' pour une relation  $T \rightarrow O$ ).

Pour caractériser plus finement les relations, nous nous appuyons en particulier sur les travaux sur la multimodalité et en particulier les résultats du projet ESPRIT AMODEUS-2 [Berssen 1994a, 1994b, 1995].

Une modalité est définie comme « un mode ou un moyen de représenter l'information aux êtres humains ou aux machines sous une forme intersubjective réalisée physiquement » [Berssen 1995]. Il s'agit d'une représentation externe qui se distingue d'une représentation interne propre au système cognitif des êtres humains, ou encore de la représentation interne au système informatisé. Cette information est concrétisée sous une forme physique en un ou plusieurs médiums : graphique, acoustique, haptique et kinesthésique. Les médiums olfactifs et gustatifs sont envisagés mais non pris en compte faute de dispositifs d'interaction réellement utilisables au-delà de prototypes de recherche.

Chaque modalité est caractérisée par une ou plusieurs propriétés orthogonales (selon Berssen) :

- Linguistique/non linguistique : cette propriété représente l'existence ou non d'une langue ou d'un langage (ou plus généralement d'un système structuré de symboles) supportant la communication.
- Analogique/non analogique : cette propriété caractérise une notion de similitude avec la réalité, c-à-d que l'information est représentée sous une forme analogique ou métaphorique.
- Arbitraire/non arbitraire : cette propriété exprime que la correspondance entre signifiant et signifié existe au préalable à la conception du système et est connue par l'utilisateur.
- Statique/dynamique : cette propriété précise si les données échangées évoluent dans le temps ou non. Les événements cycliques (comme le clignotement d'un curseur) sont cependant considérés comme statiques.

Dubois [Dubois 2001b], qui caractérise les relations de la même manière, propose également d'ajouter la dimension spatiale du langage qui peut être unidimensionnel, bidimensionnel (2D) ou tridimensionnel (3D).

Ces caractéristiques peuvent être ajoutées de manière facultative dans le schéma sous forme d'annotation textuelle placée sur les flèches des relations, si le concepteur en voit l'intérêt. Nous pensons cependant qu'il vaut mieux garder ces détails en dehors de la représentation graphique car ils compromettent trop la lisibilité du schéma.

## VI.2.8 Propriété de mobilité des entités

### VI.2.8.1 Propriété de mobilité absolue

Les entités (U, T, O, E, S) peuvent être mobiles ou fixes dans l'environnement de l'accomplissement de la tâche. Il s'agit donc de contraintes de mobilité imposées aux entités par construction. Ces contraintes sont représentées par un symbole facultatif placé en bas à droite de l'entité (Figure 69).

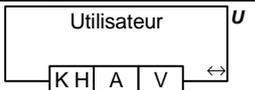
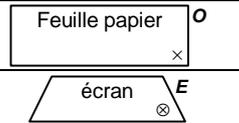
Symbole	Signification	Exemple
↔	L'entité peut bouger (ou être déplacé) pendant l'exécution de la tâche	
×	L'entité est immobile	
⊗	Toujours (c-à-d pendant toutes les tâches de l'application)	

Figure 69 : Représentation des contraintes spatiales.

Utilisés seuls, ces symboles expriment une contrainte vis-à-vis de l'environnement de l'accomplissement de la tâche ; le référentiel est donc :

- soit fixe : par exemple celui de la pièce de travail où se déroule la tâche,
- soit mobile pour une application embarquée : par exemple le référentiel d'un véhicule qui est lui-même en déplacement par rapport à un référentiel plus global.

Ces symboles ne constituent qu'un résumé des situations les plus courantes et nous jugeons que cette information est suffisante sans surcharger le schéma par des symboles trop riches. Cependant, à l'usage nous pourrions affiner cette représentation si cela s'avère nécessaire.

La nature exacte du mouvement (nombre de degrés de liberté en translation et en rotation, amplitude du mouvement, etc...) doit donc être précisée dans le modèle détaillé des entités en dehors du schéma.

Par exemple dans [Mansoux 2004] les auteurs considèrent que les degrés de liberté peuvent prendre 4 valeurs : stationnaire, translation, rotation, libre. Ces degrés de liberté sont exprimés dans un référentiel qui peut être l'environnement ou un autre objet du système.

### VI.2.8.2 Propriété de mobilité relative entre deux entités

Tous les symboles présentés ci-dessus peuvent être précédés par l'étiquette d'une autre entité : cette entité sert de référentiel pour la contrainte qui n'est plus absolue par rapport à l'environnement mais relative à cette entité. Par exemple, en Figure 70a, un casque de visualisation (HMD en anglais) est porté par l'utilisateur et se déplace avec lui : il est donc fixe par rapport à U mais pas par rapport à l'environnement. Dans la Figure 70b, le stylo est tenu par l'utilisateur mais n'est pas lié à U de manière rigide (d'où le symbole ' $\leftrightarrow$ ').



Figure 70 : Représentation des contraintes géométriques relatives.

Ces contraintes relatives peuvent également être vues comme des relations entre les objets d'une autre nature que les relations liées à l'interaction vues au paragraphe VI.2.7. Nous avons délibérément choisi de ne pas les représenter par des flèches (même graphiquement différentes) sur le schéma pour ne pas les confondre et surtout pas le surcharger par des relations secondaires vis-à-vis de l'objectif principal d'IRVO.

Un concept similaire existe dans le modèle ASUR (à partir de la version ASUR++ [Dubois 2003b]) où un trait double permet d'exprimer les liaisons fixes entre deux entités, ou encore la proximité ou l'inclusion.

Dans le formalisme de Renevier (Cf. § V.2.1) la proximité physique ou la liaison est directement exprimée graphiquement par une juxtaposition des deux entités.

### VI.2.8.3 Valeurs implicites de la propriété de mobilité

Nous avons vu que cette propriété de mobilité est facultative. Dans ce cas, il n'y a pas de valeur implicite par défaut, l'information est simplement absente du schéma.

Cependant, dans le cas des objets inclus (Figure 60d et e), s'il n'y a pas de symbole explicite sur le schéma, il faut considérer que les objets inclus sont implicitement liés de manière rigide (" $\otimes$ ") à l'objet englobant. Dans le cas contraire, il faudra donc explicitement mettre le symbole *ad hoc* éventuellement précédé de l'étiquette de l'entité englobante servant de référentiel.

### VI.2.9 Résumé du formalisme d'IRVO

Le tableau de la Figure 71 résume, de manière synthétique, les principaux éléments du formalisme graphique d'IRVO.

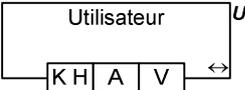
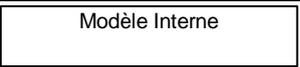
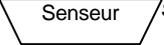
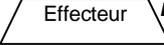
		Symbole graphique	Description	
<b>Entités</b>	Utilisateur (U)		§ VI.2.3	
	Objet	Outil (T) : - Outils réels : Tr - Outils virtuels : Tv		§ VI.2.4
		Objet du domaine (O) : - Objets réels : Or - Objets virtuels : Ov		
	Modèle interne (M)			§ VI.2.5
	Transducteur (sur frontière R/V)	Senseur (S)		§ VI.2.6
Effecteur (E)				
<b>Relations</b>	Echange de données entre entités	Action de l'utilisateur	Flèche partant d'un canal (KH, A, V) d'un utilisateur U	§ VI.2.7.1
		Perception de l'utilisateur	Flèche aboutissant à un canal (KH, A, V) d'un utilisateur U	§ VI.2.7.1
		Communication entre utilisateurs	Flèche entre deux utilisateurs	§ VI.2.7.2
		Action (et réaction) de l'outil	Flèche entre T et O	§ VI.2.7.3
		Relations entre modèle interne et objets	Flèches entre M et T, et entre M et O	§ VI.2.7.4
	Mobilité relative ou absolue		↔, ×, ⊗ (dans l'entité)	§ VI.2.8
	Objet mixte (représente le lien entre les parties réelles et virtuelles de l'objet)	Outil mixte		§ VI.2.4.2
Objet du domaine mixte				
Opérateur de fusion perceptuelle		⊕	§ VI.2.7.7	
<b>Propriétés</b>	Frontière entre mondes réel et virtuel (représentation de la nature des entités)	$\frac{R}{V} \text{ --- }$	§ VI.2.1.1	
	Frontière dans le monde réel (représentation du lieu des entités)		§ VI.2.1.2	

Figure 71 : Tableau résumant le formalisme d'IRVO

## VI.3 Utilisation d'IRVO pour la description des systèmes mixtes

Nous examinons dans cette section les utilisations du modèle IRVO pour décrire des dispositifs de Réalité Mixte existants dans un but d'analyse et de compréhension de ces systèmes. Tout d'abord nous étudions en détail quatre exemples tirés de la littérature et représentatifs du domaine de la Réalité Mixte, plus un cinquième, qui appartient au domaine des interfaces classiques, et qui montre qu'IRVO est aussi utilisable à la description de ce type d'interfaces.

Dans un deuxième temps, nous systématisons notre démarche en l'appliquant aux 44 applications de la littérature que nous avons recensées au Chapitre II, ce qui nous permet de bâtir une

taxonomie des systèmes de Réalité Mixte plus précise que la simple dichotomie Réalité Augmentée/Virtualité Augmentée.

Enfin, nous comparons les capacités descriptives d'IRVO avec celles des deux autres formalismes dédiés à la modélisation des systèmes mixtes : ASUR et le formalisme de Renevier.

### VI.3.1 Exemples de modélisation avec IRVO

Afin de montrer l'usage qui peut être fait d'IRVO dans la modélisation de systèmes de Réalité Mixte nous nous proposons de faire de la « rétroconception » en prenant quelques exemples de la littérature et en les modélisant avec IRVO.

#### VI.3.1.1 Modélisation de CASPER-v2

Nous avons présenté cette application de chirurgie augmentée développée par Emmanuel Dubois [Dubois 2001b] au Chapitre II (§ II.2.6.1). La modélisation ASUR de CASPER-v2 a également été présentée au paragraphe IV.1.7.1.

Nous rappelons brièvement que ce dispositif permet d'assister un chirurgien qui doit réaliser une ponction péri-cardiaque, opération risquée car l'aiguille de ponction doit arriver très près du cœur du patient sans léser des organes vitaux comme les poumons ou le cœur lui-même. L'application est donc une aide au guidage pour la mise en place de l'aiguille de ponction.

La modélisation que nous proposons ici correspond à une version hypothétique où, en plus de la visualisation de la trajectoire idéale sous forme de cône, la partie cachée de l'aiguille est aussi représentée en continuité de celle-ci dans le casque de visualisation porté par le chirurgien. Pour différencier cette version du dispositif CASPER-v2 initial, nous la baptisons CASPER-v2+.

La modélisation de CASPER-v2+ avec notre modèle IRVO est présentée Figure 72. On voit l'action du chirurgien sur le support d'aiguille, outil réel (Tr) qu'il introduit dans le corps du patient, objet réel de la tâche (Or). Il perçoit directement la résistance à la pénétration par sa main (retour kinesthésique) et par son regard il perçoit visuellement le support et la partie de l'aiguille qui est encore visible.

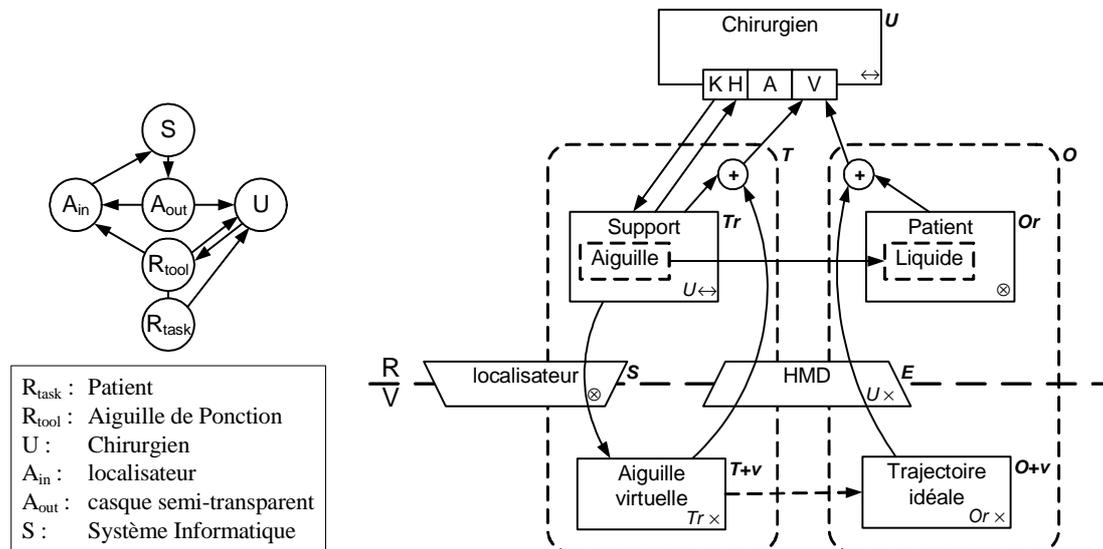


Figure 72 : Modélisation de CASPER-v2+ avec IRVO (à droite) Le modèle ASUR du même dispositif est reproduit à gauche (Cf. Figure 31).

Le système, par l'intermédiaire d'un localisateur à 6 degrés de liberté (senseur S), suit les mouvements du support de l'aiguille et peut ainsi en déduire à tout instant la position réelle de l'aiguille et donc gérer une représentation virtuelle de celle-ci (augmentation virtuelle de l'outil,

$T+v$ ). Par ailleurs le système gère également une représentation virtuelle du patient ; en pratique, il s'agit d'une trajectoire idéale représentée sous forme de cône de sécurité (augmentation virtuelle  $O+v$ ) et qui a été déterminée avant l'opération par le chirurgien à l'aide d'échographies. L'ensemble des augmentations est perçu par le chirurgien par l'intermédiaire d'un casque de visualisation semi-transparent (HMD qui a valeur d'effecteur, E) et qu'il porte (symbolisé par U X). Le symbole  $\oplus$  montre la fusion perceptuelle qui est opérée par le dispositif : l'outil réel est augmenté par une aiguille virtuelle ( $T=Tr+v$ ) qui est perçu comme un tout, et l'objet réel de la tâche est augmenté par la trajectoire idéale ( $O=Or+v$ ) qui est perçu comme un tout également.

L'action se situe clairement dans le monde réel et pour être précis l'outil est l'aiguille de ponction et l'objet véritable de la tâche, le liquide péri-cardien : cette action est représentée par une flèche normale sur le schéma. Cependant l'aiguille comme le liquide sont invisibles au chirurgien car ils sont cachés dans le corps du patient et c'est pourquoi nous les avons représentés en traits pointillés sur le schéma.

A l'inverse, l'aiguille virtuelle et la trajectoire idéale sont visibles au chirurgien (c'est pour cela qu'elles ont été construites !) mais l'action entre cet outil virtuel et l'objet du domaine virtuel est représentée par une flèche en traits pointillés car elle est secondaire vis-à-vis de la tâche. En pratique cela permet de déclencher une alarme lorsque l'aiguille s'approche trop près de la limite du cône.

Nous remarquons que, comme le patient est réputé fixe pendant l'opération (!), il n'est pas utile d'utiliser le localisateur pour suivre la position de l'objet réel. Le localisateur est par contre très pratique pour calibrer le système au début de l'opération pour aligner la trajectoire idéale avec le patient, mais cela n'apparaît pas sur le schéma puisque il s'agit d'une autre tâche (qui pourrait éventuellement être modélisée avec IRVO). On pourrait cependant envisager de localiser le patient aussi pendant l'opération, par mesure de sécurité, pour s'assurer qu'il n'y a pas de dérive dans le système.

Par contre, le localisateur sert pendant toute la durée de l'opération pour suivre la position du casque de visualisation qui est porté par le chirurgien (et donc mobile). Ce suivi n'est volontairement pas représenté dans ce schéma car cela ne fait pas partie de l'interaction personne – système mais se trouve à un niveau technique : il s'agit concrètement de projeter un modèle 3D (aiguille + trajectoire) sur un écran 2D.

IRVO permet donc de modéliser tous les aspects liés à l'interaction entre l'utilisateur et le système : l'action de l'utilisateur sur l'outil ainsi que la perception des différents outils et objets. Il met bien en évidence les artefacts mis en œuvre tant réels que virtuels ainsi que les relations entre ces différents artefacts. Il fait également ressortir la nature mixte de l'outil et de l'objet de la tâche par l'utilisation du rectangle en traits pointillés et la perception « fusionnelle » de ces artefacts par l'utilisateur grâce au symbole  $\oplus$ .

### VI.3.1.2 Modélisation de Audio Notebook

Nous avons présenté cette application [Stifelman 1996] au Chapitre II, § II.2.1.6. Nous avons choisi cette application car c'est l'une des rares à proposer une augmentation sonore d'un objet réel. Rappelons brièvement que ce dispositif maintient une synchronisation entre les pages d'un carnet sur lequel l'utilisateur prend des notes lors d'une conférence et un enregistrement audio réalisé en même temps.

Dans la Figure 73 nous représentons la modélisation IRVO de ce dispositif pour la **tâche de relecture des notes**. Lorsque l'utilisateur relit ses notes, il tourne les pages de son carnet, l'objet réel de la tâche ( $Or$ ). La page est détectée par un dispositif *ad hoc* (S) ce qui déclenche la lecture de l'enregistrement audio ( $O+v$ ) associé à cette page ; cet enregistrement audio est aussi perçu par l'utilisateur par l'intermédiaire du haut-parleur (E) simultanément à la lecture de la page. L'objet

physique est donc bien augmenté par l'enregistrement audio mais il n'y a pas fusion de la perception comme dans l'exemple précédent mais ici utilisation de deux modalités (sonore et visuelle).

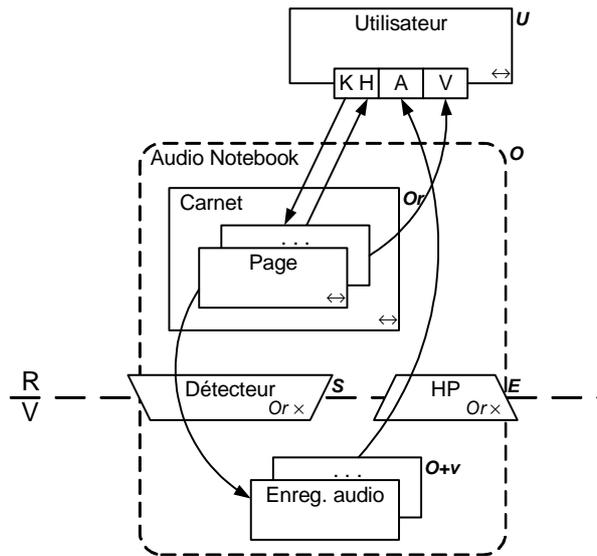


Figure 73 : Modélisation de la tâche « relecture des notes » de Audio Notebook avec IRVO.

Il n'y a pas d'outil dans cette tâche puisque l'utilisateur feuillette directement les pages du carnet avec sa main. Une autre vision serait de dire que le carnet est l'outil et que l'objet véritable de la tâche est le texte manuscrit. Cependant, ce genre de discussion pointilleuse ne nous semble pas fondamental ; en effet, cela ne change rien aux artefacts qui sont mis en œuvre dans ce dispositif mais seulement l'étiquette associée.

### VI.3.1.3 Modélisation du DoubleDigitalDesk

Nous avons présenté le DoubleDigitalDesk [Wellner 1993b] au Chapitre II, § II.2.1.3. La modélisation avec IRVO est donnée Figure 74. Ce dispositif permet le travail collaboratif à distance entre deux utilisateurs (U1 et U2). La distance est représentée dans le schéma par le fait qu'il existe une frontière physique entre les deux utilisateurs qui les empêchent de communiquer directement. En remplacement, ils disposent d'un canal de communication audio symbolisé par la flèche double entre les canaux 'A' de chaque utilisateur et qui passe à travers le monde virtuel. Les transducteurs qui permettent le passage réel – virtuel (en l'occurrence deux paires de micros et haut-parleurs) ne sont pas représentés pour ne pas alourdir le schéma (Cf. § VI.2.7.2).

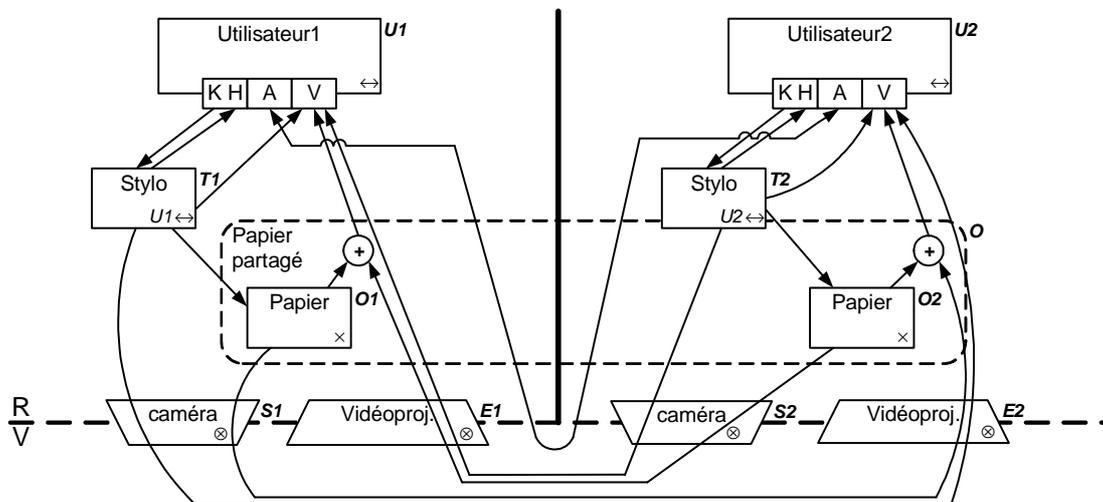


Figure 74 : Modélisation du DoubleDigitalDesk avec IRVO

Chaque utilisateur écrit avec un stylo réel (T1 et T2) sur une feuille de papier réelle (O1 et O2). Tout ce qui est tracé par l'utilisateur 1 est capturé par une caméra fixe (S1) placée au-dessus de la table et affiché sur la feuille de papier de l'utilisateur 2 avec un vidéoprojecteur (E2) placé au-dessus de la table : il résulte de cette opération que l'utilisateur 2 perçoit simultanément ses propres tracés réels augmentés des tracés de l'utilisateur 1, le symbole '+' mettant en évidence cette fusion perceptuelle. Symétriquement, une capture des tracés de U2 est effectuée par S2 puis restituée par E1 ce qui permet la perception augmentée pour l'utilisateur 1. Globalement les deux utilisateurs ont l'illusion d'écrire sur la même feuille de papier augmentée qui est l'objet du domaine de la tâche partagée (O, représenté en traits pointillés).

Grâce à la caméra, le stylo (ou même la main) de chaque utilisateur est visible par l'autre tant qu'il se trouve au-dessus de la feuille : il est ainsi possible de pointer des éléments du tracé qu'ils soient réels ou « virtuels » ce qui permet de supporter les activités de coordination par exemple.

Comme il est expliqué dans [Wellner 1993b], le signal provenant des caméras doit être soigneusement traité pour limiter les effets de boucle vidéo (la caméra ne doit pas filmer ce qui est affiché par le vidéoprojecteur car cela sera réaffiché sur le dispositif distant et ainsi de suite). Rien dans le formalisme d'IRVO n'est prévu pour exprimer ces contingences techniques car ce n'est pas son rôle. Cependant, le formalisme graphique met en évidence cette boucle passant par les caméras et les video-projecteurs ce qui permet d'aider les concepteurs à détecter ce type de problèmes.

### VI.3.1.4 Modélisation du jeu de Mah-Jongg

Ce jeu proposé par [Szalavári 1998] est présenté au Chapitre II, § II.2.9.1. Il existe de nombreuses actions possibles dans ce jeu et nous avons choisi de ne représenter que deux d'entre-elles, la préparation de combinaisons sur la tablette individuelle (la *main*) et le jeu d'une combinaison sur la table de jeu commune.

#### Tâche de préparation de combinaisons

La tâche de préparation de combinaisons est individuelle et peut se faire concurremment avec les autres phases de jeu. La modélisation avec IRVO est donnée en Figure 75.

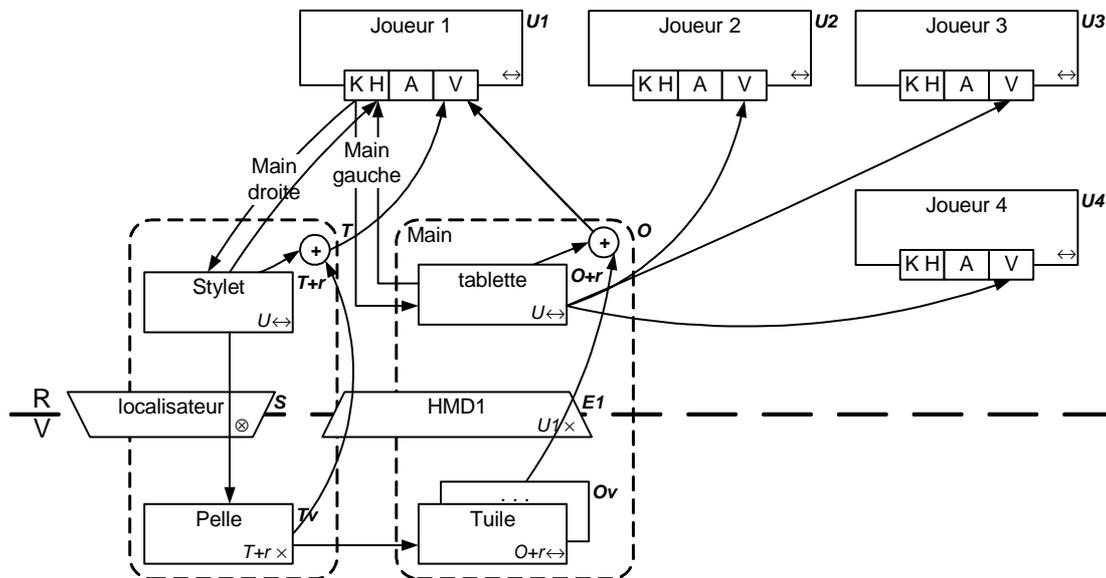


Figure 75 : Modélisation de la tâche « préparer des combinaisons » du jeu de Mah-Jongg avec IRVO.

Le Joueur 1 tient dans sa main gauche (ou plus exactement sa main non-dominante) la tablette réelle (O+r) sur laquelle sont posées les tuiles virtuelles (Ov) le tout constituant sa *main*. Il tient dans sa main droite un stylet (T+r) dont la position en 6 degrés de liberté est suivie par un

localisateur (S), qui lui permet de manipuler l'outil virtuel représenté sous forme d'une petite pelle (Tv). L'utilisateur est muni d'un casque de visualisation semi-transparent (E1) qui lui permet d'avoir une fusion perceptuelle d'une part de l'outil ( $T+r$ ) et d'autre part de l'objet ( $O=Ov+r$ ).

C'est donc bien une application de la Virtualité Augmentée car l'action se situe dans le monde virtuel : avec la pelle le joueur prend et déplace des tuiles pour faire des combinaisons, le stylet et la tablette n'étant que des « augmentations » réelles de ces artefacts pour qu'ils puissent être « empoignable » par l'utilisateur, ce qui justifie bien l'utilisation des notations  $T+r$  et  $O+r$  pour ces artefacts.

Nous avons fait figurer dans ce schéma les 3 autres joueurs uniquement pour montrer qu'ils ne peuvent pas tricher car ils ne perçoivent que la tablette du joueur 1 (et son stylet) et non les tuiles qui ne sont visibles qu'à travers le casque du joueur 1. Ceci est en fait une simplification de l'application qui montre en réalité aux autres joueurs, à travers leur propre casque de visualisation, une représentation des tuiles du joueur 1 débarrassée des symboles : ils ne peuvent donc que voir le nombre de tuiles possédées par ce joueur mais pas le détail de sa *main*.

A nouveau, sur ce schéma, nous n'avons pas fait figurer la localisation du casque de visualisation du joueur par le localisateur S car cela ne fait pas partie de l'interaction. Par contre, la localisation de la tablette aurait pu être représentée car dans une certaine mesure, elle peut aussi être considérée comme un outil qui permet de déplacer les tuiles de la *main* toutes ensemble. Cependant, si on se place du point de vue de l'utilisateur, le véritable outil est bien la pelle, la tablette n'est qu'un support : l'utilisateur pourrait d'ailleurs la poser sur la table sans que cela modifie en rien ses possibilités de jeu. La liaison entre les tuiles et la tablette est malgré tout exprimée par la propriété de mobilité ( $O+r \leftrightarrow$ ) placée dans la Tuile virtuelle ( $Ov$ ).

### Tâche de jeu d'une combinaison

La tâche « poser une combinaison sur la table de jeu commune » est une tâche collaborative même si en pratique un seul joueur agit. La modélisation correspondante est représentée Figure 76.

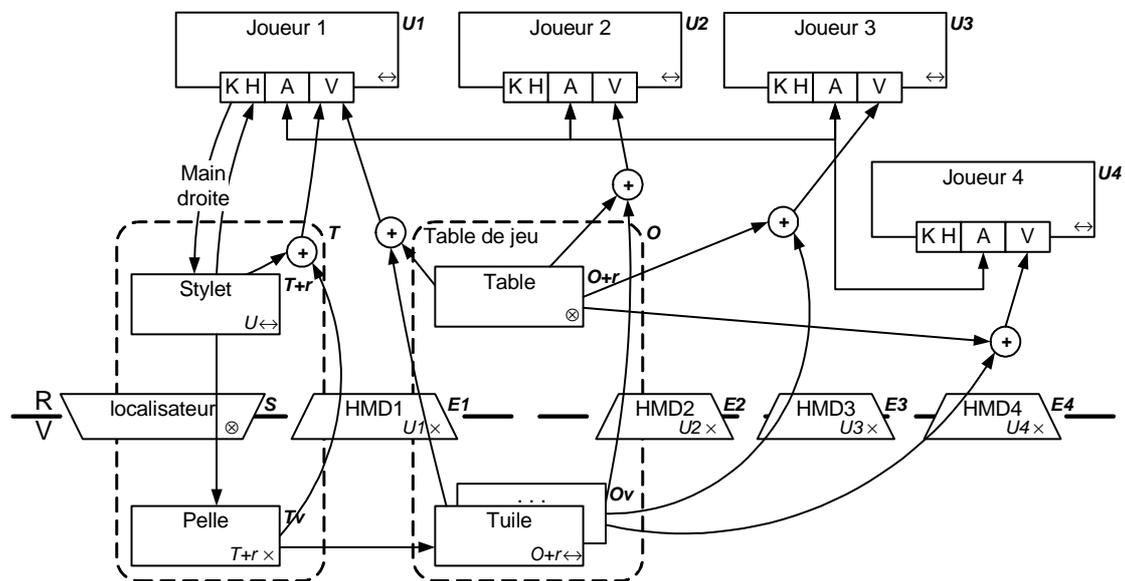


Figure 76: Modélisation de la tâche « poser une combinaison sur la table de jeu commune » du jeu de Mah-Jongg avec IRVO.

Le joueur utilise toujours le stylet ( $T+r$ ) prolongement réel de la pelle ( $Tv$ ), outil virtuel, lui permettant de transporter la combinaison de tuiles depuis sa tablette jusqu'à la table de jeu ; celle-ci est composée d'une table réelle fixe ( $O+r$ ) et des tuiles que dépose le joueur ( $Ov$ ) qui constituent l'objet de cette tâche (sur la table de jeu se trouvent aussi les autres tuiles déjà jouées, la muraille,

etc. qui ne sont pas représentées ici car elles ne font pas partie de la tâche). A la différence de la tâche précédente, les autres joueurs voient la combinaison jouée par le joueur 1 dans leur propre casque de visualisation (E2 à E4) car cette fois-ci la combinaison devient publique (elle est jouée) : ils peuvent donc prendre connaissance des tuiles jouées ce qui va évidemment influencer sur leur propre jeu ultérieur. Le stylet et la pelle du joueur 1 sont aussi observables par les autres joueurs mais cela n'est pas représenté sur le schéma pour ne pas l'alourdir par des relations secondaires. On a également représenté sous forme d'une flèche bidirectionnelle la communication orale possible entre les joueurs pour leur permettre de se coordonner et de faire tous les commentaires appropriés.

### VI.3.1.5 Un autre exemple de modélisation

Jusqu'à maintenant nous avons montré que IRVO était tout à fait approprié pour modéliser l'interaction d'un ou de plusieurs utilisateurs avec un système de Réalité Mixte. Mais en fait, il peut tout aussi bien modéliser des interactions classiques avec des interfaces graphiques : nous avons vu en particulier comment représenter une souris avec IRVO au paragraphe VI.2.6.

Bien sûr cela ne présente pas beaucoup d'intérêt sauf dans le cas d'interactions avancées allant au-delà des interfaces WIMP telle que l'interaction avec des « outils transparents » (*Toolglasses*) proposés par Bier et al. [Bier 1993].

Prenons un exemple tiré de la littérature : il s'agit d'une application permettant l'édition de réseaux de Petri, CPN2000, qui utilise l'interaction bi-manuelle et les *toolglasses* [Beaudouin-Lafon 2000b]. Nous représentons Figure 77 la modélisation avec IRVO des tâches utilisant des *toolglasses*. L'outil à activer ( $T_n$ ) est placé dans une palette d'outil  $T_1$  que l'utilisateur déplace avec sa main gauche à l'aide d'un trackball pour l'amener au-dessus d'une place ou d'une transition du réseau de Petri (c'est l'objet de la tâche,  $O$ ) sur laquelle il veut appliquer cet outil : pour que cela soit faisable, il faut que la palette et les outils soient semi-transparentes d'où le nom de *toolglass*. Pour activer l'outil, l'utilisateur vient avec le pointeur de la souris (qu'il manipule avec sa main droite comme d'habitude) cliquer sur l'outil et donc simultanément sur l'objet qui est derrière.

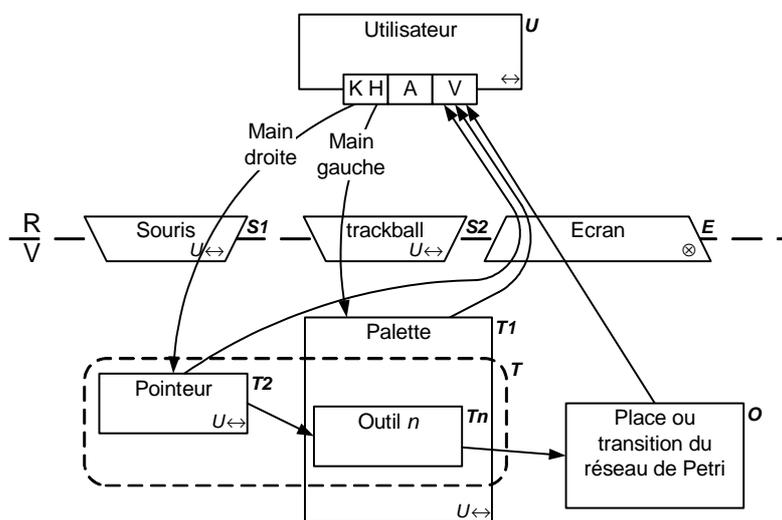


Figure 77 : Modélisation de l'interaction bi-manuelle avec un *Toolglass*.

Nous utilisons le cadre en traits pointillés pour montrer que globalement l'outil est composite (pointeur  $T_2$  et outil  $T_n$ ) mais il ne s'agit pas ici d'un objet mixte au sens de la Réalité Mixte, car le rectangle en traits pointillés ne franchit pas de frontière. Cet exemple montre l'usage que nous pouvons faire d'IRVO pour modéliser des interactions complexes même s'il n'y a aucun objet réel impliqué dans l'interaction.

### VI.3.2 Notre Taxonomie des systèmes de Réalité Mixte

Parmi les taxonomies examinées au Chapitre II, § II.3, nous avons vu qu'aucune ne permettait de distinguer les applications qui sont du domaine de la Réalité Mixte et celles qui sont des applications informatiques ordinaires, ou encore de classer finement les différentes applications de Réalité Mixte entre elles. Nous proposons donc ici une taxonomie qui cherche :

- A affiner la classification entre Réalité Augmentée et Virtualité Augmentée de Dubois (Cf. § II.3.5) en distinguant des sous-catégories,
- A mieux distinguer les applications classiques ou du domaine de la Réalité Virtuelle des applications de Réalité Mixte.

Notre taxonomie se base donc sur la nature des objets présents dans l'application en distinguant les deux types : outils et objets du domaine de la tâche. Pour la construction de la taxonomie nous considérons dans un premier temps : **un seul utilisateur** qui utilise **au plus un outil** sur **un objet du domaine** dans **une tâche** élémentaire donnée.

Nous généralisons par la suite aux cas plus complexes : plusieurs utilisateurs, plusieurs outils, plusieurs objets et/ou plusieurs tâches.

#### VI.3.2.1 Taxonomie basée sur l'interaction d'un utilisateur

Nous avons vu au paragraphe VI.2.4.2 que, dans IRVO, les objets de la tâche peuvent être « à cheval » sur la frontière réel/virtuel. Plus précisément, nous avons 4 cas :

- L'objet est entièrement dans le monde réel ( $O=Or$ ),
- L'objet est entièrement dans le monde virtuel ( $O=Ov$ ),
- L'objet est dans le monde réel et augmenté par des annotations virtuelles ( $O=Or+v$ ),
- L'objet est dans le monde virtuel et augmenté par des « choses » réelles ( $O=Ov+r$ ).

Les mêmes cas peuvent être considérés pour les outils avec les notations correspondantes :  $Tr$ ,  $Tv$ ,  $Tr+v$  et  $Tv+r$ . En combinant toutes les possibilités nous obtenons une matrice 4 x 4 (Figure 78) qui est à la base de notre taxonomie.

Dans cette matrice, nous retrouvons en bas à droite (cas  $Tv/Ov$  et  $Tv+r/Ov$ ) les interfaces traditionnelles du type WIMP ou les interfaces de la Réalité Virtuelle. Même si l'outil est augmenté, il s'agit toujours d'un outil virtuel agissant sur un objet virtuel.

En dehors de ces deux cas nous considérons que tout les autres cas font partie de la réalité mixte :

- Sur la partie gauche de la matrice (huit cases), nous retrouvons le domaine de la **Réalité Augmentée** (RA), tel que défini par Dubois (Cf. § II.3.5), puisque l'objet de la tâche est réel (éventuellement augmenté).
- Sur la partie droite de la matrice (six cases restantes), nous retrouvons le domaine de la **Virtualité Augmentée** (VA) puisque l'objet de la tâche est virtuel (éventuellement augmenté).

Pour les six cas de Virtualité Augmentée, nous considérons que dans les deux cas en haut à droite ( $Tr/Ov$  et  $Tr+v/Ov$ ) nous retrouvons les **interfaces tangibles** : en effet, l'objet du domaine de la tâche est virtuel mais les outils pour agir sont des objets réels (briques, cubes, etc...). Cependant, il n'est pas exclu que des applications de Virtualité Augmentée tombent dans ces catégories sans pour autant être de véritables interfaces tangibles (comme par exemple Charade). Nous confirmons ainsi notre position prise au Chapitre II, § II.1.6 sur le fait que les interfaces tangibles sont un cas particulier de Virtualité Augmentée.

Dans la partie de la matrice recouvrant la Réalité Augmentée, nous avons deux cas particuliers :

- Le premier, le cas situé dans l'angle en haut à gauche ( $Tr/Or$ ),
- Le second, les deux cas situés dans l'angle en bas à gauche ( $Tv/Or$  et  $Tv+r/Or$ )

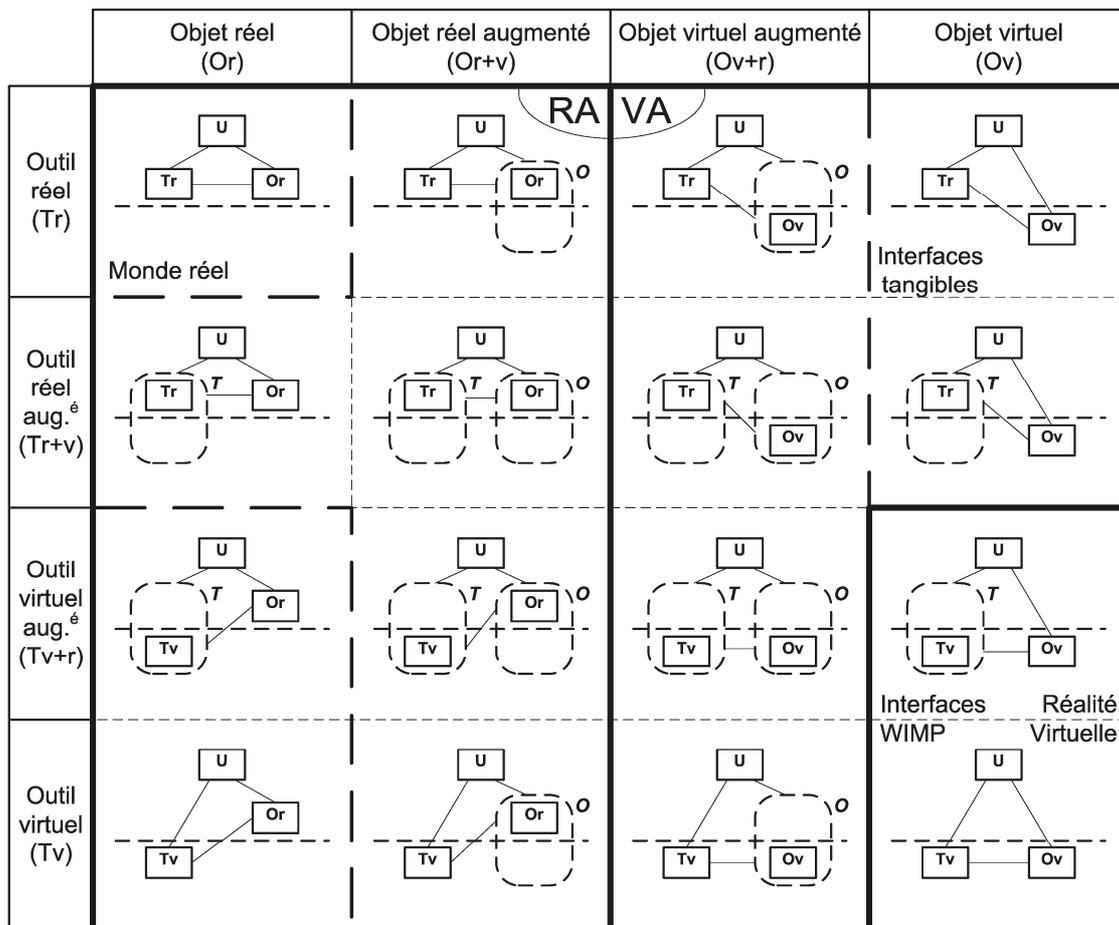


Figure 78 : Notre taxonomie des systèmes de réalité mixte (pour un utilisateur).

Le cas Tr/Or correspond au monde réel « ordinaire », c'est-à-dire sans aucune augmentation. Il semble donc qu'il nous faudrait exclure *a priori* ce cas car, s'il n'y a plus besoin de monde virtuel, il n'y a donc plus besoin d'ordinateurs et nous n'avons plus à nous poser la question de développement d'une interface homme-machine ! Même si cela a une portée limitée, cela montre qu'IRVO permet même de modéliser les actions d'un utilisateur dans le monde réel (taper sur un clou avec un marteau, par exemple).

En réalité, le cas Tr/Or n'est pas aussi trivial qu'il y paraît au premier abord. En effet, si l'on se place dans le cas où les objets réels ne sont pas perçus directement par l'utilisateur mais à travers un médium numérique comme la vidéo d'un objet distant (Trevisan *et al.* appellent ces objets des « objets numériques réels », Cf. § II.1.2.1) alors la modélisation par IRVO reste tout à fait pertinente pour modéliser des interactions de cette nature. Nous verrons justement au paragraphe suivant un exemple d'application collaborative (DoubleDigitalDesk) qui rentre parfaitement dans cette catégorie, mais il faut bien reconnaître qu'il s'agit d'un cas marginal.

Les cas Tv/Or et Tv+r/Or sont assez déroutant. En effet, il s'agit de cas où un objet réel est modifié par l'action d'un outil virtuel (éventuellement augmenté). Evidemment cela nécessite l'usage d'un 'effecteur' *ad hoc* permettant de passer d'une action virtuelle à un effet dans le monde physique (par exemple une machine-outil à commande numérique). Il n'y a aucune raison *a priori* de ne pas considérer que ces cas fassent partie de la Réalité Augmentée et nous les laissons donc à l'intérieur de ce périmètre. Par contre, comme nous le verrons ci-dessous, nous n'avons trouvé aucune application dans la littérature qui rentre dans ces deux cas ce qui fait que nous avons quelques doutes sur l'existence effective de ces catégories comme faisant partie de la Réalité Augmentée ... Cela méritera sans doute de plus amples réflexions.

Enfin, nous pouvons faire un parallèle entre notre taxonomie et la taxonomie des systèmes interactifs mixtes de Nigay *et al.* présentée au Chapitre II, § II.3.6 :

- notre axe horizontal sur la nature de l'objet du domaine de la tâche correspond exactement à l'axe vertical de leur taxonomie sur la nature de la tâche,
- notre axe vertical sur la nature de l'outil correspond à leur axe horizontal sur l'utilisation d'objets physiques ou numériques dans l'interaction.

Ainsi, les 4 cas en bas en gauche de notre matrice correspondent à des cas d'*Interaction Mixte en Réalité Augmentée* et les 4 cas en haut à droite correspondent à des cas d'*Interaction Mixte en Virtualité Augmentée*. Nous retrouvons donc les mêmes cas identifiés par Nigay *et al.* mais en distinguant, de manière plus fine, 4 sous cas pour chaque catégorie identifiée.

### VI.3.2.2 Taxonomie pour les applications collaboratives et/ou complexes

La matrice que nous avons présentée Figure 78 a été construite à partir de l'interaction d'un utilisateur qui utilise **au plus un outil** sur **un objet du domaine** dans **une tâche** élémentaire donnée. Voyons donc ce qui se passe lorsque nous considérons :

- Des applications collaboratives, donc **plusieurs** utilisateurs,
- Des interactions complexes avec **plusieurs** outils utilisés simultanément,
- Des applications complètes, donc comportant **plusieurs** tâches et **plusieurs** objets du domaine.

Dans le cas d'applications collaboratives on peut facilement replacer chaque interaction individuelle avec le système dans une des cases de la matrice. Si les différents utilisateurs n'utilisent pas les mêmes outils ou utilisent des objets du domaine de nature différente (par exemple, un utilisateur travaille sur l'original réel et les autres sur une copie virtuelle), alors les différentes interactions se trouveront réparties sur la matrice dans plusieurs cases.

Le cas des interactions complexes est similaire : chaque outil peut être de l'un des 4 types et donc l'interaction peut chevaucher plusieurs cases dans la même colonne. Le seul cas présenté est celui des *Toolglases* (Cf. § VI.3.1.5) mais les deux outils sont de même nature (virtuels en l'occurrence).

Pour les applications complètes, si l'utilisateur est amené à changer d'objet du domaine en changeant de tâche ou plus fréquemment d'outil, il se peut très bien que certains de ces objets du domaine (respectivement outils) soient réels et d'autres virtuels. Ainsi, en remplaçant chaque interaction d'une tâche élémentaire dans la case correspondante on peut donc se retrouver avec des interactions réparties dans plusieurs cases de la matrice pour l'application complète.

Au final, quelles que soient les causes (plusieurs utilisateurs, plusieurs outils, plusieurs objets du domaine et/ou plusieurs tâches) une application peut donc être amenée à couvrir plusieurs cases de la matrice.

Cependant, les applications collaboratives ne se limitent pas à une simple juxtaposition d'interactions individuelles comme dans une application multi-utilisateurs. En effet, les interactions de chaque personne concourent à une production commune et ceci apparaît dans les diagrammes IRVO par le fait que l'objet de la tâche est commun à tous les utilisateurs :

- c'est un véritable partage de l'objet dans le cas des applications collaboratives locales ou des applications avec WYSIWIS strict.
- c'est un partage de la représentation interne, mais chaque utilisateur possède sa vue propre de l'objet dans les cas du WYSIWIS relâché.

Reprenons la modélisation du DoubleDigitalDesk (Cf. § VI.3.1.3) : tous les outils (des stylos) sont dans le monde réel, ainsi que les objets de la tâche (des feuilles de papier). Chaque utilisateur interagit avec son stylo (Tr) sur sa feuille de papier (Or) donc on se trouve bien dans le cas Tr/Or de la matrice. Mais si l'on considère l'objet véritable qui est partagé par les deux utilisateurs il est

composé de la feuille réelle augmenté de la projection de la vidéo de la feuille distante qui est un objet réel lui aussi. On se trouve donc dans un cas d'augmentation singulier d'augmentation **d'un objet réel par un autre objet réel distant** que nous notons, en suivant nos conventions :  $O=Or+r!$  Ce type d'augmentation ne « brise » pas la frontière réel/virtuel mais la frontière physique entre 2 lieux à distance, ce qui est une manière de mettre en place le concept de Téléprésence [Buxton 1992].

Nous pourrions imaginer un DoubleDigitalDesk « amélioré » qui, en plus, pourrait ajouter des augmentations virtuelles sur les deux objets réels et nous aurions le cas :  $O=Or+r+v!$  On pourrait aussi avoir un « TripleDigitalDesk » (sans doute techniquement difficile à construire) et obtenir  $O=Or+r+r!$

Par contre, ce n'est pas symétrique et le même procédé n'a pas de sens avec les objets virtuels. En effet, par construction un objet virtuel est ubiquitaire (Cf. les remarques sur l'absence de frontières dans le monde virtuel au § VI.2.1.3) et peut être rendu perceptible aux utilisateurs de la même manière quel que soit l'endroit où ils se trouvent : il suffit de rajouter les transducteurs *ad hoc* pour que les utilisateurs puissent percevoir l'objet et éventuellement agir dessus.

### VI.3.2.3 Application de notre taxonomie aux applications de la littérature

Nous reprenons les 44 applications que nous avons recensées dans la littérature au Chapitre II pour leur appliquer notre taxonomie et les classer de manière plus fine que la simple dichotomie RA/VA que nous avons utilisé dans la colonne 'type d'augmentation' dans les tableaux de synthèse au Chapitre II, § II.2.11. La Figure 79 montre le résultat de ce classement. Nous pouvons constater que les cases ne sont pas uniformément couvertes, en particulier la case  $Tr/Or+v$  est saturée par une vingtaine d'applications qui entrent dans cette catégorie.

Conformément à la discussion du paragraphe précédent, certaines applications couvrent plusieurs cases :

- Elles possèdent plusieurs tâches utilisant des objets du domaine et/ou des outils différents comme :
  - DigitalDesk : certaines applications utilisent des objets réels augmentés (PaperPaint) et d'autres des objets virtuels augmentés (Calculator).
  - TROC qui permet soit de cliquer avec la souris (outil virtuel) sur les objets virtuels soit de les attraper avec un cube physique (outil réel augmenté).
  - Caméléon qui permet d'écrire sur les strips avec un stylo réel ou d'établir une communication audio en appuyant sur un bouton virtuel.
- Elles sont collaboratives et les utilisateurs n'ont pas la même interface comme LivePaper qui dans l'application bureautique permet à l'utilisateur local d'écrire sur du papier avec un stylo (outil réel) alors que les utilisateurs distants utilisent une souris (outil virtuel).

Pour InteriorDesign, le cas nous paraît plus délicat car si le but ultime est bien de meubler une pièce réelle, toute l'application ne repose que sur la manipulation d'objets virtuels superposés à la pièce réelle qui ne sert que de « toile de fond » : selon le point de vue, cette application est donc soit du domaine de la Réalité Augmentée, soit du domaine de la Virtualité Augmentée.

Nous pouvons constater que la case  $Tr/Or$  n'est pas vide puisque le DoubleDigitalDesk rentre dans cette catégorie (Cf. discussion du paragraphe précédent).

Par contre, nous n'avons pas trouvé d'application pour remplir les deux cases en bas à gauche ( $Tv/Or$  et  $Tv+r/Or$ ) ce qui nous laisse dubitatif sur la pertinence de ces catégories en tant que catégories de Réalité Augmentée.

Enfin, l'application SeamlessDesign se retrouve bien en dehors de la Virtualité Augmentée car, conformément à notre discussion du Chapitre II, il s'agit en réalité d'une application de Réalité

Virtuelle. Nous avons également fait figurer pour mémoire les *Toolglasses* que nous avons modélisé avec IRVO (Cf. § VI.3.1.5) même s'il s'agit ici d'une interaction entièrement dans le monde virtuel.

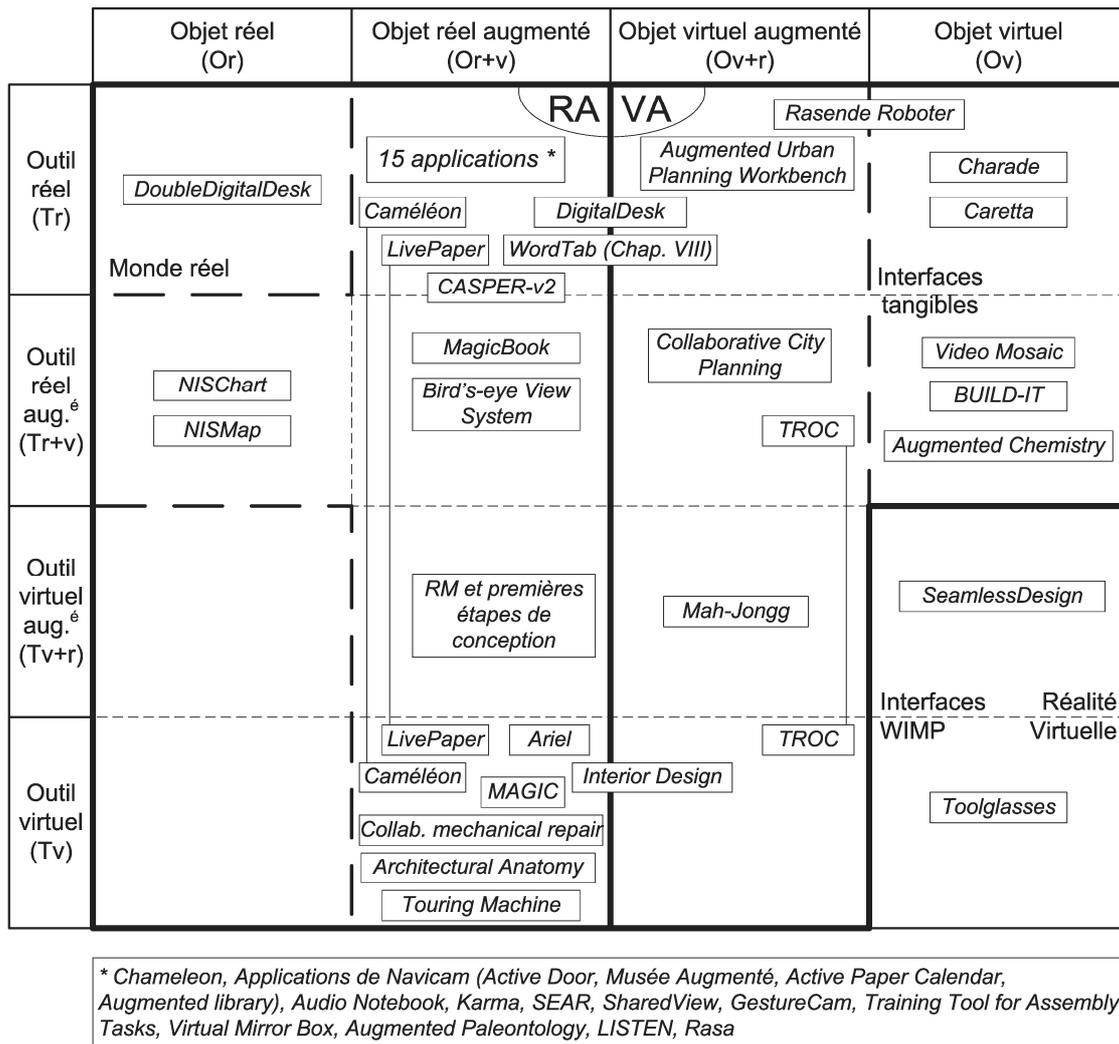


Figure 79 : Les 44 applications recensées au Chapitre II placées dans notre taxonomie ainsi que les *Toolglasses* (§ VI.3.1.5) et l'application *WordTab* du projet Accoréa (présentée au Chapitre VIII).

### VI.3.3 Comparaison d'IRVO avec ASUR et le formalisme de Renevier

Dans cette section, nous comparons le formalisme d'IRVO avec les deux autres formalismes proposés pour modéliser l'interaction dans les systèmes de Réalité Mixte : le formalisme ASUR et le formalisme de Renevier.

Cette comparaison est effectuée sur des exemples typiques pour chacun des deux formalismes.

#### VI.3.3.1 Comparaison IRVO / ASUR 2004

Une comparaison complète des deux formalismes a été faite par David Juras [Juras 2004]. Nous nous limitons ici à une comparaison sur un exemple pour montrer les différences principales qui existent entre les deux formalismes sans chercher l'exhaustivité.

Nous reprenons l'exemple de *CASPER-v2+* que nous avons modélisé au paragraphe VI.3.1.1 avec le formalisme IRVO et qui a été modélisé avec le formalisme ASUR 2004 par [Juras 2004]. Les deux modèles sont présentés à la Figure 80.

Au niveau des dispositifs d'entrée/sortie on constate que dans les deux formalismes ils sont représentés de manière équivalente. Dans cet exemple, la distinction que nous faisons entre



IRVO va plus loin en considérant que les informations liées à l'outil ont valeur d'augmentation virtuelle de l'outil réel (d'où la notation  $T+v$ ) et que les informations liées au patient (cône de la trajectoire idéale) ont valeur d'augmentation de l'objet du domaine réel (d'où la notation  $O+v$ ). La liaison fonctionnelle entre ces augmentations et les objets qu'elles augmentent est renforcée par les cadres en traits pointillés placés autour de ces deux objets.

Enfin, au niveau de la perception visuelle de l'utilisateur, dans le schéma ASUR trois flèches figurent en provenance du HMD, de l'aiguille et du patient, ce qui pourrait laisser croire que le chirurgien doit continuellement porter son attention sur les 3 objets et qu'il y a donc discontinuité perceptuelle. Dans IRVO, le HMD n'apparaît pas comme une entité sur laquelle le chirurgien doit porter son attention. Par contre la perception des objets virtuels est directement représentée sur le schéma par le passage de la relation à travers la frontière R/V et le HMD : la continuité perceptuelle de l'outil et de son augmentation d'une part et de l'objet du domaine et de son augmentation d'autre part sont donc explicitement montrées dans le schéma à l'aide de l'opérateur  $\oplus$  qui montre la fusion perceptuelle opérée par le système. Dans le schéma IRVO, il n'y a donc que deux flèches qui arrivent sur le canal visuel du chirurgien, une pour l'outil et une pour l'objet du domaine.

### VI.3.3.2 Comparaison IRVO / Formalisme de Renevier

Le formalisme de Renevier (Cf. § V.2.1) est conçu initialement pour représenter des scénarios projetés à un niveau abstrait, c'est-à-dire en ne prenant pas en compte les dispositifs d'entrée/sortie mis en œuvre. IRVO permet également de ne pas représenter les transducteurs, et donc peut aussi, dans une certaine mesure, se placer à ce niveau d'abstraction.

Nous comparons les deux formalismes sur un exemple de la littérature proposé dans [Renevier 2003c] : il s'agit du Navicam+ version collaborative du Navicam (Cf. § II.2.8.1). Philippe propose l'étude de plusieurs scénarios et nous reprenons ici un scénario où l'expert est à distance et définit de manière synchrone l'augmentation du tableau pour son ami visiteur qui est lui dans le musée. Les modélisations de cette application avec les deux formalismes sont présentées à la Figure 81.

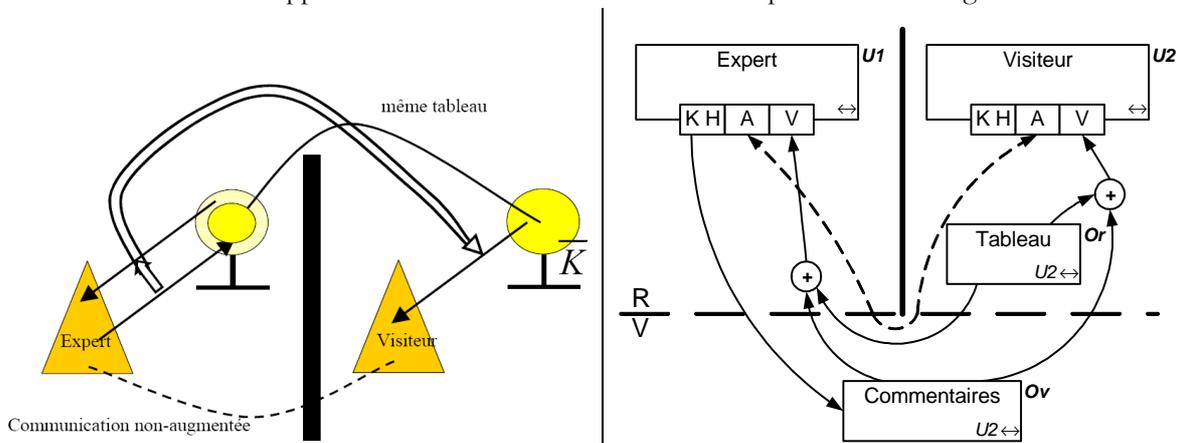


Figure 81 : Comparaison de la modélisation du Navicam+ : à gauche, le formalisme de Renevier (d'après [Renevier 2004c]), à droite, le formalisme IRVO.

Nous voyons que les deux formalismes font apparaître le même objet du domaine, réel, le tableau dans le musée. Dans le schéma de gauche, cet objet est présent pour l'expert sous forme d'un représentant numérique de l'objet réel, alors que dans IRVO, c'est bien le même objet perçu par une relation de perception qui traverse le monde virtuel pour franchir la distance qui sépare les deux lieux.

La création de l'augmentation par l'expert est clairement représentée dans le schéma de gauche par une flèche double, alors que dans le schéma de droite cet aspect dynamique n'est pas visible. Par contre, dans IRVO la perception augmentée du visiteur apparaît explicitement avec l'opérateur  $\oplus$ .

Dans les deux schémas la possibilité de communication par téléphone est symbolisée mais le sens du trait pointillé est différent : dans le schéma de gauche, cela représente un échange non-augmenté alors que à droite cela signifie que la relation est secondaire voire facultative.

Une autre différence tient dans le fait qu'IRVO représente les sens perceptifs et actionnels des utilisateurs explicitement ce que ne fait pas le formalisme de Renevier. Dans cet exemple, la communication entre l'expert et le visiteur est bien orale alors que l'augmentation produite est visuelle. Même dans des stades très amont de la conception du système, il nous semble que ces informations sont essentielles pour comprendre comment l'interaction doit se dérouler dans le futur système.

### VI.3.3.3 Bilan de ces comparaisons

Les comparaisons que nous venons de faire n'ont pas valeur de démonstration mais permettent toutefois de faire ressortir quelques points du formalisme d'IRVO qui ne figurent pas dans les deux autres formalismes. Ces points sont :

- Les sens perceptifs et actionnels des utilisateurs sont explicitement représentés dans les schémas IRVO. Même si cela est très limité, il s'agit d'une information essentielle permettant de voir immédiatement les sens mis en jeu dans l'interaction avec le système.
- Les objets mixtes sont clairement identifiables par la mise en relief de la relation d'augmentation représentée sous forme d'un rectangle en traits pointillés.
- Les perceptions augmentées sont représentées sous la forme d'une fusion (à l'aide de l'opérateur  $\oplus$ ) des perceptions de l'objet réel et de l'objet virtuel.

Il faut constater que les formalismes ASUR et de Renevier permettent de modéliser des applications mobiles ce que IRVO ne prend pas en compte à l'heure actuelle (ou plus exactement ne prend que partiellement en compte).

### VI.3.4 Bilan sur les capacités descriptives d'IRVO

Dans cette section, nous avons utilisé IRVO pour décrire des dispositifs existants afin de les analyser et les comprendre. Nous avons examiné en détail cinq applications tirées de la littérature (quatre de Réalité Mixte et une classique) montrant les apports du formalisme d'IRVO pour la description de ces systèmes.

Nous avons fait ce travail systématiquement pour les 44 applications recensées au Chapitre II. Même si certaines applications nous ont posé quelques problèmes sur la définition précise des outils ou des objets de la tâche (souvent parce que la description qui est faite du dispositif dans les articles laisse la place à plusieurs interprétations), toutes les applications ont pu être modélisées avec succès en utilisant le formalisme d'IRVO.

Nous avons proposé, en partant de la nature des objets présents dans les systèmes et en distinguant les deux types d'objets (outil et objet du domaine), une taxonomie comportant 16 cas. D'une part, cette taxonomie affine des taxonomies existantes :

- Taxonomie de Dubois (Cf. § II.3.5) – Réalité Augmentée/Virtualité Augmentée,
- Taxonomie de Nigay *et al.* (Cf. § II.3.6) – Interaction Mixte en Réalité Augmentée / Interaction Mixte en Virtualité Augmentée.

D'autre part, elle permet de différencier les systèmes de Réalité Mixte des systèmes classiques ou des systèmes de Réalité Virtuelle car ils apparaissent comme des cas particulier de la taxonomie générale.

Ainsi, en réalisant la modélisation avec IRVO d'un système existant, on peut facilement, par la nature des outils et des objets du domaine, le classer finement et le comparer aisément à d'autres systèmes similaires. Cette approche nous a permis de reclasser une application que ses auteurs

appelaient « augmentée » mais qui en pratique est du domaine de la Réalité Virtuelle du point de vue de l'interaction homme-machine.

Nous avons enfin comparé les capacités descriptives d'IRVO avec les deux autres formalismes existants : le formalisme ASUR et le formalisme de Renevier. Cette comparaison ne se veut pas exhaustive mais a été faite sur un exemple significatif pour chacun des modèles. Il apparaît de cette comparaison qu'IRVO a un pouvoir d'expression équivalent à ces deux formalismes, voire supérieur pour les aspects liés à la représentation des augmentations des objets et de la représentation de la perception augmentée de ces objets par les utilisateurs.

Si ces résultats ne nous permettent pas de prouver « formellement » les capacités descriptives d'IRVO ils nous permettent toutefois de nous montrer plutôt confiants quant à son potentiel de représentation de l'interaction dans les systèmes de Réalité Mixte.

## VI.4 Utilisation d'IRVO à des fins d'analyses prédictives

Nous voulons qu'IRVO puisse servir à la conception d'applications nouvelles. Il faut donc, au-delà de capacités descriptives, qu'IRVO permettent de comparer différentes solutions alternatives en phase de conception afin d'éliminer les moins bonnes et de retenir les meilleures. Pour cela, il faut avoir un certain nombre de critères les plus objectifs possibles et éventuellement chiffrés pour évaluer chaque solution et faire des analyses prédictives sur l'utilisabilité du système en cours de développement.

Comme, dans notre cas nous nous intéressons à l'interaction de l'utilisateur avec le futur système, il s'agit avant tout d'évaluer *a priori* la qualité de l'interaction, ce qui nous amène naturellement à considérer des propriétés ergonomiques.

Nous proposons donc un ensemble de règles permettant de vérifier la complétude d'un schéma, le respect du critère d'observabilité, de la continuité et la cohérence. Ces règles ne prétendent pas à l'exhaustivité mais constituent une base permettant d'évaluer la modélisation d'un système en utilisant le formalisme d'IRVO.

### VI.4.1 Vérification de la complétude de l'interaction

Une modélisation d'un système faite avec IRVO peut être évaluée du point de vue de la complétude à partir du schéma. En effet, l'utilisateur est réputé être en interaction avec le système c'est-à-dire qu'il y a action de l'utilisateur sur le système et perception du nouvel état du système après action. Cela conduit donc à proposer la règle suivante :

**Règle 1 :** *Il doit exister sur le schéma une boucle action-perception partant d'un canal de l'utilisateur et arrivant à lui en passant par l'outil (les outils) et l'objet de la tâche :  $U \rightarrow T \rightarrow O \rightarrow U$ . Il ne doit donc pas y avoir d'outils ou d'objets non-reliés dans un schéma IRVO.*

Dans certains cas l'outil n'est pas explicité sur le schéma car l'action de l'utilisateur peut se faire directement sur l'objet de la tâche (Cf. Audio Notebook, § VI.3.1.2), ce qui fait que le schéma ne montre plus qu'une boucle :  $U \rightarrow O \rightarrow U$ .

Il existe un second cas particulier : certaines tâches peuvent être des tâches d'observation pure de l'état du système sans action de l'utilisateur et donc sans outil. Dans ce cas, nous donnons la règle suivante, comme exception à la règle 1 :

**Règle 1 bis :** *Dans les modélisations correspondantes aux tâches où il n'y a pas d'action de l'utilisateur, il n'existe pas de boucle action-perception, il n'y a pas d'outils, et seule est présente la flèche  $O \rightarrow U$ .*

## VI.4.2 Vérification de l'observabilité

Les propriétés d'observabilité d'une interface homme-machine dictent que l'utilisateur doit pouvoir contrôler son action. Il doit donc pouvoir percevoir l'objet de la tâche, ce qui est pris en compte par les règles 1 et 1 bis. Il doit aussi percevoir l'outil en train d'agir sur l'objet par un ou plusieurs de ses sens (par exemple : perception kinesthésique de la réaction de l'objet sur l'outil + perception visuelle). Ceci dicte la deuxième règle.

**Règle 2 :** *Il doit exister sur le schéma au moins une flèche  $T \rightarrow U$ .*

Cas des objets du domaine mixtes

Un objet mixte est composé d'une partie réelle et d'une partie virtuelle. Ces deux parties doivent donc être perçues simultanément par l'utilisateur :

- Soit par le même sens avec une fusion perceptive qui est exprimée par l'opérateur  $\oplus$
- Soit par plusieurs sens (multimodalité)

**Règle 3 :** *La perception d'un objet du domaine mixte doit se faire en percevant chacune des composantes. Il doit donc exister :*

- Soit des relations :  $Or \rightarrow \oplus$ ,  $Ov \rightarrow \oplus$  et  $\oplus \rightarrow U$ ,
- Soit des relations :  $Or \rightarrow U$  et  $Ov \rightarrow U$  qui arrivent sur des canaux différents de l'utilisateur.

Cas des outils mixtes

Pour un outil mixte :

- soit il est augmenté en perception, et la même règle s'applique que pour les objets mixtes,
- soit il est augmenté en action alors, les 2 parties de cet outil doivent avoir des actions simultanées dans les mondes réel et virtuel : la partie réelle de l'outil doit agir sur un objet réel et la partie virtuelle doit agir sur un objet virtuel.

**Règle 4 :** *Si un outil mixte est augmenté en perception, chacune de ses composantes doit être perçues. Il doit donc exister :*

- Soit des relations :  $Tr \rightarrow \oplus$ ,  $Tv \rightarrow \oplus$  et  $\oplus \rightarrow U$ ,
- Soit des relations :  $Tr \rightarrow U$  et  $Tv \rightarrow U$  qui arrivent sur des canaux différents de l'utilisateur.

**Règle 4 bis :** *Si un outil mixte est augmenté en action, Il doit exister des relations  $Tr \rightarrow Or$  et  $Tv \rightarrow Ov$ .*

Les deux augmentations ne sont pas exclusives. On pourrait même affirmer que comme corollaire de la règle 2, la règle 4 devrait toujours être appliquée. Cependant, en cas d'augmentation de l'action de l'outil, si les deux actions dans les mondes réel et virtuel sont cohérentes, alors il est tout à fait imaginable que la perception d'une seule des deux parties de l'outil soit suffisante pour respecter la propriété d'observabilité.

## VI.4.3 Vérification de la cohérence des transducteurs vis-à-vis des relations

On peut vérifier sur un schéma IRVO que les transducteurs sont correctement utilisés.

Tout d'abord, il faut que le transducteur soit compatible avec l'information qui le « traverse ». Pour les relations partant ou aboutissant à un utilisateur, le canal utilisé renseigne sur le type d'information et on peut donc vérifier que le transducteur permet bien la traduction de cette information du monde virtuel au monde réel. On en déduit les 2 règles suivantes.

**Règle 5 :** *Un senseur n'admet que des flèches qui le traversent dans le sens monde réel – monde virtuel. L'information véhiculée (et dont la nature est donnée par le canal de l'utilisateur sur lequel aboutit la flèche) doit être d'un type compatible avec les capacités de traduction analogique/ numérique du senseur.*

**Règle 6 :** *Un effecteur n'admet que des flèches qui le traversent dans le sens monde virtuel – monde réel. L'information véhiculée (et dont la nature est donnée par le canal de l'utilisateur d'où part la flèche) doit être d'un type compatible avec les capacités de traduction numérique / analogique de l'effecteur.*

Par exemple, une flèche qui traverse un effecteur « vidéoprojecteur » doit aller du monde virtuel au monde réel et aboutir sur le canal V d'un utilisateur (éventuellement après fusion avec l'opérateur  $\oplus$ ).

Par contre, pour les flèches entre des outils et des objets, il n'est pas possible de faire cette vérification a priori car la nature de l'information échangée n'est pas explicitée par la notation graphique.

Une deuxième vérification au niveau des transducteurs est aussi possible, elle porte sur le « partage » de ce transducteur par plusieurs relations (Cf. § VI.2.6, 2<sup>o</sup> alinéa). Par exemple un localisateur 3D peut localiser plusieurs objets simultanément par contre un casque de visualisation est individuel et donc toutes les flèches qui le traversent doivent aboutir à l'utilisateur qui est lié au casque (Cf. le jeu de Mah Jongg, § VI.3.1.4). Il est cependant difficile de formuler une règle générale sur cette notion de « partage » car elle n'est pas suffisamment formalisée et dépend beaucoup de la nature exacte des senseurs et des effecteurs, et doit donc être étudiée au cas par cas. Il faut noter que cette notion de partage est également présente dans le modèle ASUR pour tous les objets.

Une troisième vérification peut aussi être faite sur le délai apporté par les transducteurs sur la boucle sensori-motrice. En effet, comme nous l'avons étudié au § VI.2.6.1, le temps de traversée des transducteurs peut être important en fonction des technologies utilisées. Il est donc important de vérifier, en fonction des relations qui traversent les transducteurs et en fonction de l'action accomplie par l'utilisateur, que les délais ne compromettent pas l'interaction envisagée. Il est difficile de définir une règle universelle mais le formalisme IRVO permet de voir aisément sur un schéma quels sont les « chemins critiques » qui peuvent poser problème. Il est certain que le formalisme pourrait être étendu pour représenter explicitement ces contraintes temporelles.

#### VI.4.4 Vérification de la continuité

La propriété ergonomique de continuité est largement étudiée dans la Réalité Mixte. Elle a été définie par [Dubois 2001b] qui distingue la continuité perceptuelle et la continuité cognitive.

La **continuité 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 perceptuel de l'utilisateur et les sens perceptifs mis en jeu.

Sur un schéma IRVO, l'opérateur  $\oplus$  montre explicitement la fusion perceptuelle d'informations en provenance d'objets réels ou virtuels. Ainsi, la présence de l'opérateur  $\oplus$  pour percevoir les différentes composantes d'un objet mixte montre que la continuité perceptuelle pour les différentes composantes des objets mixtes est assurée. Cependant, l'absence de cet opérateur ne signifie pas forcément qu'il y a discontinuité. Ainsi dans le cas de l'*Audio Notebook* (Cf. § VI.3.1.2), l'augmentation est sonore et ne peut donc pas fusionner avec l'information visuelle provenant du carnet réel. En pratique seule la lecture du carnet demande à l'utilisateur de porter son regard dans la direction des pages. Pour l'écoute des séquences sonores, il n'y a pas de direction privilégiée, elle peut donc se faire simultanément à la lecture. Cependant, le haut-parleur est intégré au support du carnet ce qui renforce la continuité perceptuelle car le son semble bien provenir des pages.

La continuité perceptuelle peut être vérifiée statiquement mais ne plus être vraie lorsque les objets sont en mouvements. En effet, les délais apportés par les transducteurs (Cf. ci-dessus) peuvent désynchroniser l'augmentation virtuelle par rapport à l'objet réel au point que l'augmentation semblera « poursuivre » l'objet réel et l'effet de fusion perceptuelle annihilé. Nous proposons donc de considérer deux sous cas :

- La **continuité perceptuelle statique** lorsque ni les objets ni l'utilisateur sont en mouvements,
- La **continuité perceptuelle dynamique** lorsque au moins une des entités est en mouvement.

La **continuité cognitive** traduit la capacité de l'utilisateur à interpréter correctement les données relatives à un concept en considérant les différents modes de représentation mise en œuvre par le système pour représenter ces données.

Emmanuel Dubois propose d'utiliser les caractéristiques de la modalité de l'information véhiculée par les relations comme base de l'analyse de la propriété de continuité cognitive [Dubois 2001b]. Cette propriété nous semble particulièrement difficile à évaluer car cela fait intervenir l'interprétation de l'utilisateur. Cela reste donc pour nous un problème ouvert.

#### VI.4.5 Vérifications liées au travail collaboratif

Les règles énoncées ci-dessus sont valables pour l'interaction entre un utilisateur et le système. Elles restent valables dans un système collaboratif pour chaque interaction individuelle. Par contre, on peut vérifier d'autres propriétés spécifiques au travail collaboratif.

Il est tout d'abord possible de vérifier des propriétés liées à l'honnêteté comme le WYSIWIS (What You See Is What I See). On peut donc donner la règle suivante :

**Règle 7 :** *un système collaboratif est WYSIWIS si les différents utilisateurs perçoivent le même objet de la tâche directement ou indirectement. Il existe donc des flèches  $O \rightarrow U1$ ,  $O \rightarrow U2$ , etc. où  $O$  est le même objet.*

Si l'objet est mixte, chacune des composantes doit être observable pour tous les utilisateurs pour respecter la règle 3 conjointement avec la règle 7.

Par exemple, le DoubleDigitalDesk respecte cette propriété par contre, le jeu de Mah Jongg ne la respecte pas dans la tâche « préparer des combinaisons » puisqu'elle est en fait individuelle et qu'il s'agit justement de préserver l'espace privé du joueur.

Pour le WYSIWIS relâché, la modélisation IRVO doit normalement faire apparaître des objets virtuels différents (Ov1, Ov2, etc.) pour les utilisateurs qui représentent les « vues » individuelles que ces utilisateurs ont de l'objet de la tâche. La synchronisation de ces vues est assurée par le modèle interne M de IRVO (qui incorpore le contrôleur de dialogue et le noyau fonctionnel de l'application).

Renevier reconsidère la notion de WYSIWIS dans les systèmes mixtes collaboratifs et mobiles sous le nom de *cohérence de l'augmentation* qui traduit la cohérence perceptive et sémantique des liens. Les utilisateurs d'un système mixte collaboratif et mobile vérifiant la cohérence de l'augmentation partagent un même modèle cognitif des liens d'augmentation et ils disposent de la même description de ces liens [Renevier 2003c].

#### Communication entre utilisateurs

On peut, dans une application supportant le travail collaboratif à distance, vérifier que la communication interindividuelle des utilisateurs est possible. Pour cela, soit les frontières du monde physique sont partiellement franchissables (cas ③, ④, ⑤, § VI.2.1.2), soit un canal de communication audio et/ou vidéo est prévu en passant par le monde virtuel. Cependant, dans certains collecticiels, la communication peut être assurée indirectement par une « chat box » textuelle. Il est donc difficile de formuler une règle générale car il existe trop de cas pour qu'on puisse les prendre tous en compte et ces vérifications doivent être effectuées au cas par cas.

### VI.4.6 Automatisation des vérifications

Parmi les règles énoncées ci-dessus, certaines pourraient être vérifiées automatiquement par un outil d'édition. Par exemple les règles 1, 1bis, 2, 3, 4 et 4 bis ne posent pas de problème car il s'agit essentiellement de vérifications syntaxiques. Pour les règles 5 et 6, la première partie de la règle est facilement vérifiable, par contre, la deuxième partie nécessiterait une formalisation plus poussée des transducteurs pour pouvoir automatiser la vérification.

### VI.4.7 Bilan sur les capacités prédictives d'IRVO

Ces règles sont un préliminaire et des travaux complémentaires avec des ergonomes sont nécessaires pour compléter et affiner ces règles. De tels travaux ont déjà été accomplis avec le formalisme ASUR qui sur ces aspects d'analyse prédictive est déjà plus avancé [Dubois 2004b]. Une partie des résultats obtenus avec ASUR devrait pouvoir être transposée à IRVO étant donnée la forte proximité des deux formalismes.

Par ailleurs, comme nous le verrons au chapitre VII, nous préconisons d'utiliser IRVO dans une approche MBA et, dans ce cadre, d'autres modèles permettent aussi de vérifier certains critères ergonomiques qui ne sont pas vérifiables avec IRVO utilisé seul. Nous verrons en particulier au § VII.2.3.2, que, grâce à la liaison que nous établissons entre l'arbre de tâche et IRVO, nous pouvons vérifier certaines propriétés liées à la continuité entre tâches.

## VI.5 Evaluation du formalisme d'IRVO

Dans cette section nous essayons d'évaluer le formalisme d'IRVO par des principes qui ont été proposés par différentes équipes pour évaluer les formalismes de spécification des interfaces homme-machine. Nous présentons deux ensembles de critères, les « principes de Namur » et les critères de Brun, puis nous passons IRVO au crible de ces deux grilles d'analyse.

### VI.5.1 Les « Principes de Namur »

Lors de l'atelier d'Eurographics DSV-IS'96, un groupe de chercheurs a établi un ensemble de critères pour l'évaluation des formalismes de spécification des interfaces homme-machine qui est connu sous le nom de « Principes de Namur » [Johnson 1996], [Jambon 1996, p128] :

- **Capacité à abstraire** (*Support for abstraction*) : c'est le pouvoir d'exprimer différents niveaux de détails ou d'abstraction, de manière à s'adapter aux différentes phases de conception. (exemple : décomposer en tâche et sous tâche).
- **Capacité à structurer** (*Support for structuring*) : les différents aspects d'une interface doivent pouvoir se spécifier séparément, pour être réalisés par des équipes différentes, et rassemblés sans perte d'information.
- **Esthétique** (*Visual appeal*) : un formalisme doit être agréable à utiliser visuellement. L'esthétique d'un formalisme participe à sa lisibilité. Par exemple, un formalisme graphique est plus esthétique qu'un formalisme textuel.
- **Apprentissage progressif** (*Incremental acquisition*) : un formalisme doit pouvoir s'apprendre facilement et progressivement comme un langage de programmation.
- **Existence d'outils d'aide à la spécification** (*Tools support*).
- **Capacité à faciliter la détection des erreurs** (*Support for error identification*) : la détection des erreurs lors des premières phases de conception permet des économies car le coût de correction est élevé.
- **Sémantique claire** (*Clear semantics*) : Un formalisme doit éviter les ambiguïtés et les difficultés d'interprétation, c'est particulièrement vrai lorsque le formalisme est utilisé par plusieurs équipes.

- **Etude de cas** (*Case study*) : un formalisme doit comporter une série d'exemples de spécification portant sur des cas d'études classiques.
- **Finalité clairement définie** (*Clear scoping*) : les buts, avantages, et limitations d'un formalisme doivent être exprimés clairement. L'utilisateur potentiel doit pouvoir identifier facilement les formalismes qui conviennent à ses besoins.
- **Adéquation à la tâche du concepteur** (*Task fit*) : la distance entre le formalisme utilisé et les concepts du domaine ne doit pas être trop importante.
- **Utilité sociale** (*Social value*) : l'utilisation des formalismes doit être replacée dans le contexte industriel de la conception d'interfaces en équipe. Dans le contexte industriel, le formalisme doit aider l'équipe de conception à atteindre des objectifs communs.

### VI.5.2 Evaluation d'IRVO selon les principes de Namur

Dans le tableau de la Figure 82 nous présentons les résultats de l'évaluation du formalisme d'IRVO selon les principes de Namur. Elle correspond à la version actuelle qui, en particulier sur le plan des outils, est améliorable. Nous notons : + pour les critères pris en compte en partie, ++ pour les critères bien pris en compte, +++ pour les critères très bien pris en compte.

Principes de Namur	Evaluation IRVO	Commentaires
Capacité à abstraire	+	IRVO se place à un niveau d'abstraction qui est celui de la tâche. Mais, le § VII.2.3 montre les relations entre IRVO et l'arbre de tâches ce qui permet de se placer à d'autres niveaux d'abstraction.
Capacité à structurer	++	IRVO structure principalement l'interaction entre les utilisateurs et les artefacts.
Esthétique	++	IRVO est un formalisme graphique. Cependant, pour des interactions complexes, les schémas peuvent perdre en lisibilité à cause du nombre de relations.
Apprentissage progressif	++	Certains concepts comme les transducteurs ou les contraintes de mobilité peuvent être ignorés dans un premier temps.
Existence d'outils d'aide à la spécification	+	Une première version d'un éditeur de schéma a été développée comme extension du logiciel Visio de Microsoft, mais il est prévu de l'étendre en particulier en automatisant les règles de vérification du § VI.3.2.
Capacité à faciliter la détection des erreurs	++	Un certains nombres de critères peuvent être vérifiés comme la complétude, la cohérence ou la non observabilité (Cf. § VI.3.2). La continuité entre tâche peut aussi être évaluée (Cf. § VII.2.3.2)
Sémantique claire	++	
Etude de cas	+++	Cf. Exemples du § VI.3.1 et taxonomie (§ VI.3.2)
Finalité clairement définie	+++	Cf. § VI.1.1
Adéquation à la tâche du concepteur	++	IRVO propose une représentation directement interprétable des utilisateurs et des artefacts.
Utilité sociale	++	IRVO s'inscrit dans une approche plus globale basée sur les modèles (Cf. § VII.1.3)

Figure 82 : Evaluation du formalisme d'IRVO selon les « principes de Namur »

S'agissant d'une autoévaluation, celle-ci est obligatoirement subjective. En pratique, ce n'est pas la valeur absolue qui nous semble intéressante dans cette évaluation mais plus la valeur relative entre chaque critère, faisant apparaître ainsi les points forts et les points faibles du formalisme d'IRVO.

### VI.5.3 Critères d'évaluation de Brun

Philippe Brun, quant à lui, propose dans sa thèse 12 critères regroupés en 3 catégories [Brun 1998, p28], [Jambon 2001], certains de ces critères étant éventuellement décomposés en sous critères.

**Le pouvoir d'expression** : il caractérise la capacité du formalisme à décrire le plus grand nombre de caractéristiques d'un système. Il se décompose en 6 critères :

- **L'adéquation entre le formalisme et le cycle de développement** ; un formalisme capable de couvrir plusieurs étapes du cycle de développement est plus avantageux pour les concepteurs.
- **L'état de l'interface et le feed-back du système** ; le formalisme doit permettre de prendre en compte les propriétés orientées vers l'utilisateur afin de mieux implémenter des propriétés comme l'honnêteté, l'observabilité et l'insistance
- **Le séquençement des actions et les contraintes temporelles** ; le formalisme doit pouvoir exprimer la séquence entre deux tâches, l'alternative, la composition dans un ordre quelconque, l'itération, l'expression de délais entre tâches.
- **La concurrence des actions** ; cela concerne la capacité à décrire le parallélisme vrai ou l'entrelacement de tâches.
- **La description des interruptions** : il s'agit de caractériser l'interruption d'une tâche par une autre. 7 sous critères sont considérés : l'interruption avec parallélisme, l'interruption avec entrelacement, l'interruption sans reprise, l'interruption avec reprise en un point, l'interruption avec reprise au début, l'interruption avec reprise à la fin, l'interruption avec reprise au point d'interruption.
- **La présentation de l'interface** : ce critère rend compte de la présentation de l'interface en particulier des objets de l'interface à l'écran, de la présentation de l'information, des messages du système, etc. Il rend compte en particulier des problèmes de cohérence, d'observabilité et de facilité de manipulation.

**Les capacités génératrices** : Les capacités génératrices d'un formalisme représentent sa capacité à dériver du code ou des informations importantes à partir des spécifications. Elle dépend de la possession ou non d'une sémantique claire et du modèle mathématique sous-jacent utilisé par le formalisme. Il y a 4 critères :

- **Générer tout ou partie du système final** ; La génération de tout ou partie du système final peut s'effectuer de plusieurs façons à partir des spécifications détaillées : manuellement, ce qui est le plus souvent le cas et de manière automatique ou semi-automatique suivant l'outillage disponible.
- **Effectuer une analyse prédictive** (surcharge cognitive, temps d'exécution d'une tâche donnée, etc.) ; La possibilité de réaliser une analyse prédictive permettra de connaître les performances de l'application orientée vers l'utilisateur ou vers le système. L'analyse prédictive peut être orientée explicitement vers l'utilisateur ou vers le système.
- **Prouver certaines propriétés du système** ; On peut distinguer deux grands types de propriétés relatives aux systèmes interactifs : d'une part les propriétés liées à l'utilisateur du système (aussi appelées propriétés externes), comme l'utilisabilité et les propriétés s'y rapportant (accessibilité du côté de l'utilisateur, ...), qui sont très générales et très difficiles à démontrer et d'autre part les propriétés relatives au système comme la terminaison ou l'accessibilité, qui sont démontrables avec des formalismes possédant un outillage théorique adapté.
- **Dériver des fonctions génériques de l'interface** ; Une fonction générique est une fonction qui s'applique à des ensembles différents de données et que l'on retrouve dans pratiquement tous les systèmes. Les opérations de couper/copier/coller ou faire/défaire sont des exemples typiques de fonctions génériques. Un formalisme prenant en compte ce type de fonction

permettrait de gagner beaucoup de temps de développement. Cette capacité d'un formalisme à générer des fonctions génériques de l'interface est aussi liée à sa sémantique propre.

**L'extensibilité et l'utilisabilité :** L'**extensibilité** d'un formalisme est sa capacité à être adapté par extension à de nouveaux besoins. Un formalisme extensible a plus de chances de survivre dans le monde réel qu'un formalisme fermé car il a plus de capacités à intégrer des avancées technologiques. Une notation peu formelle semble plus facile à étendre qu'un formalisme très strict mathématiquement parlant, car elle est moins contrainte par les notations. Un formalisme possède une bonne **utilisabilité** (corrélée à son temps d'apprentissage) si « la description de cas simples est simple, et la description de cas complexes est possible ». Le temps d'apprentissage et la sémantique sont les deux facteurs les plus importants déterminant l'utilisabilité d'un formalisme.

#### VI.5.4 Evaluation du formalisme d'IRVO selon les critères de Brun

Dans le tableau de la Figure 83 nous présentons les résultats de l'évaluation du formalisme d'IRVO selon les critères de Brun. Nous notons :  $\emptyset$  pour les critères non significatifs, + pour les critères pris en compte en partie, ++ pour les critères bien pris en compte, +++ pour les critères très bien pris en compte.

	Critères de Brun	Evaluation IRVO	Commentaires
Pouvoir d'expression	Adéquation entre le formalisme et le cycle de développement	+	Phase couverte : « spécification du système interactif ». Voir le § VII.1.3 pour l'intégration de IRVO dans un approche plus globale basée sur les modèles.
	L'état de l'interface et le feed-back du système	+++	C'est le but principal d'IRVO !
	Le séquençement des actions et les contraintes temporelles	$\emptyset$	Il n'y a aucune relation temporelle entre tâches car IRVO se positionne au niveau d'une tâche. Voir le § VII.2.3 pour une discussion approfondie sur les relations entre IRVO et l'arbre de tâches.
	La concurrence des actions	+	Plusieurs actions parallèles peuvent être représentées sur un même schéma IRVO mais sans précisions sur leur synchronisation.
	La description des interruptions	$\emptyset$	Non prévu (Cf. remarque sur le séquençement des actions et les contraintes temporelles)
	La présentation de l'interface	++	Les outils et objets de la tâche sont représentés. Cependant les détails de « mise en page » sont en dehors du champ d'IRVO car il se place à un niveau semi-concret.
Capacités génératrices	Générer tout ou partie du système final	+	Le § VII.3.3 explique comment il est possible de dériver manuellement à partir d'IRVO le modèle de dialogue (exprimé avec le formalisme d'AMF-C).
	Effectuer une analyse prédictive	$\emptyset$	Non prévu. Extension possible ?
	Prouver certaines propriétés du système	++	Un certains nombres de critères peuvent être vérifiés comme la multiplicité des périphériques, des outils et des objets, l'observabilité, la cohérence, ou la continuité entre tâche. Il ne s'agit cependant pas de preuves formelles.
	Dériver des fonctions génériques de l'interface	$\emptyset$	En dehors des buts d'IRVO.
	Extensibilité	++	Le formalisme peut être étendu (mais ce n'est pas la direction que nous préconisons).
	Utilisabilité	+++	Les exemples (§ VI.3.1) montrent que le formalisme est simple d'usage et facilement compréhensible pour les cas simples.

Figure 83 : Evaluation d'IRVO selon les critères de Brun

On peut constater que plusieurs critères ne sont pas couverts et de nombreux ne le sont que peu. Ceci est aisément explicable par notre approche pour la conception de ce formalisme. Il s'agit d'aborder les aspects qui étaient mal pris en compte dans le contexte de la Réalité Mixte par les formalismes classiques, mais pas de refaire un nouveau formalisme « universel ». On pourrait bien sûr être tenté de rajouter dans IRVO les concepts qui manquent mais ce n'est pas notre approche :

- Nous voulons garder ce formalisme simple et utilisable et donc minimiser les concepts. Pour cela le rasoir d'Occam<sup>16</sup> restera le principe fondamental qui nous dirige.
- Il est inutile de refaire ce qui existe déjà et qui est bien fait. Notre approche est donc d'articuler IRVO avec d'autres modèles et formalismes en l'intégrant dans une approche basée sur les modèles (MBA – Model-Based Approach), ce que nous verrons en détail au Chapitre VII.

### VI.5.5 Bilan sur l'évaluation d'IRVO

Les évaluations présentées ci-dessus sont intéressantes mais ne sont pas vraiment « formelles » et de plus difficilement applicables, surtout sous forme d'auto-évaluation. En effet, ces critères ont été initialement établis pour pouvoir comparer des formalismes entre eux et non pour l'évaluation individuelle d'un formalisme.

Elles permettent de dégager des tendances en termes de point faibles et de points forts mais ne peuvent prouver les capacités du modèle IRVO notamment les capacités prédictives et génératives. Il est d'ailleurs très difficile de prouver ces capacités pour un modèle car il faudrait pouvoir faire développer le même système à deux équipes différentes, l'une disposant du modèle pour la conception, l'autre pas, et de comparer la qualité des systèmes résultant des conceptions.

## VI.6 Bilan du chapitre

Dans ce chapitre nous avons présenté le modèle IRVO et son formalisme. L'objectif principal d'IRVO est de modéliser l'interaction entre le (ou les) utilisateur(s) et le système de Réalité Mixte en explicitant les outils et objets du domaine (tant réels que virtuels) mis en jeu et leurs relations, en se plaçant au niveau des tâches élémentaires d'interaction.

IRVO insiste tout particulièrement sur la modélisation des outils et objets du domaine : ils peuvent être réels ou virtuels selon leur position par rapport à la frontière entre le monde réel (dans lequel est placé l'utilisateur) et le monde virtuel (dans lequel est le programme informatique). Les objets mixtes apparaissent ainsi comme une abstraction de conception : ils sont composés d'un ou plusieurs objets réels et virtuels et l'augmentation provient soit de la perception « fusionnelle » que les utilisateurs ont des différents composants, soit de leur capacité à agir dans les deux mondes (pour les outils).

Par rapport aux deux autres formalismes existants pour la modélisation des systèmes de Réalité Mixte (ASUR et le formalisme de Renevier), qui représentent eux aussi les utilisateurs et les objets réels et virtuels, IRVO présente quelques spécificités importantes :

- Les objets mixtes sont explicitement représentés par la mise en relief de la relation d'augmentation sous forme d'un rectangle en traits pointillés.
- Les perceptions augmentées sont représentées sous la forme d'une fusion (à l'aide de l'opérateur  $\oplus$ ) des perceptions de l'objet réel et de l'objet virtuel.
- Les sens perceptifs et actionnels des utilisateurs sont explicitement représentés dans les schémas IRVO, ce qui permet d'explicitier les sens qui sont requis pour l'interaction avec le système et permet (dans une certaine mesure) de représenter des interactions multimodales.

---

<sup>16</sup> Voir épigraphie en tête de ce Chapitre.

- Les dispositifs d'entrée/sortie sont considérés comme étant composés de deux parties : la fonction de conversion entre monde réel et monde virtuel (appelée transducteur) et la partie réelle du dispositif qui peut aussi être considérée comme un outil au sens de notre modèle. Cependant, dans la plupart des cas les deux composants peuvent être fusionnés.

IRVO permet de représenter l'interaction de plusieurs utilisateurs avec le système sur un même schéma, ainsi que les communications interindividuelles entre les utilisateurs, qu'elles soient médiatisées ou non. Il permet donc de supporter la conception d'applications collaboratives même si tous les aspects du travail collaboratif ne sont pas couverts par IRVO (en particulier les aspects de coordination). Ces aspects sont également couverts par le formalisme de Renevier mais pas par ASUR.

Toutefois, IRVO ne prend actuellement pas en compte explicitement les aspects liés à la mobilité contrairement aux formalismes ASUR et de Renevier.

La validation de la capacité descriptive du modèle a été notamment faite en l'appliquant aux 44 applications de la littérature que nous avons recensées au Chapitre II. Nous avons plus spécifiquement donné 5 exemples détaillés de modélisation d'applications avec IRVO qui montrent concrètement l'usage qui peut être fait de ce formalisme pour comprendre et analyser l'interaction dans les systèmes de Réalité Mixte.

Au-delà de cet intérêt, IRVO permet aussi de comparer les applications les unes aux autres dans un but de les classer. Nous avons donc construit une taxonomie qui distingue 16 cas de base obtenus par le croisement de la nature de l'objet de la tâche (objet réel, objet réel augmenté, objet virtuel augmenté, objet virtuel) avec la nature de l'outil utilisé pour interagir (outil réel, outil réel augmenté, outil virtuel augmenté, outil virtuel). Elle fait apparaître les deux grandes catégories de la Réalité Mixte qui sont la Réalité Augmentée (l'objet de la tâche est réel) et la Virtualité Augmentée (l'objet de la tâche est virtuel) mais aussi inclut les applications non mixte de l'informatique WIMP ou de la Réalité Virtuelle.

Afin d'explorer les capacités prédictives de IRVO sur l'utilisabilité d'un système en cours de conception, nous avons proposé un ensemble de règles de vérification d'un système de Réalité Mixte modélisé avec IRVO. Ces règles bâties à partir de propriétés ergonomiques classiques (comme l'observabilité) ou plus spécifiques aux systèmes mixtes (comme la continuité) permettent de vérifier la complétude d'une modélisation et certaines propriétés liées à l'utilisabilité du système modélisé. Cependant ces règles constituent un préliminaire et des travaux complémentaires avec des ergonomes sont nécessaires pour compléter et affiner ces règles.

Nous avons essayé d'avoir une approche plus formelle en évaluant le modèle IRVO en appliquant les principes de Namur et les critères de Brun. Il est tout d'abord toujours délicat de s'auto-évaluer, ce qui limite la portée des résultats. De plus, ces critères ont surtout été établis pour pouvoir comparer des formalismes entre eux et non pour une évaluation individuelle d'un formalisme.

Pour pouvoir concevoir des applications de grande envergure, utiliser un formalisme comme IRVO n'est pas suffisant. Il faut inclure le modèle IRVO dans une méthodologie de conception. Le chapitre suivant est donc consacré aux aspects de conception des systèmes de Réalité Mixte où nous verrons que le modèle IRVO prend naturellement place dans une démarche de conception MBA aux côtés des modèles de tâches et de dialogues.



# Chapitre VII Conception de systèmes de Réalité Mixte

*On fait la science avec des faits, comme on fait une maison avec des pierres :  
mais une accumulation de faits n'est pas plus une science qu'un tas de pierres n'est une maison.*  
Henri Poincaré (1854 – 1912).

Dans ce chapitre nous abordons la conception des systèmes de Réalité Mixte. Nous faisons le choix d'une approche basée sur les modèles ou Model-Based Approaches (MBA).

Après une introduction aux approches MBA, nous montrons qu'IRVO peut harmonieusement s'insérer parmi les autres modèles en les complétant. Nous étudions ensuite tout particulièrement les liens entre IRVO et l'un des modèles les plus importants de ces approches, l'arbre de tâches.

Nous proposons ensuite une première version d'une méthodologie de conception des systèmes mixtes et collaboratifs comme une extension de la démarche CoCSys dans laquelle nous nous insérons. Nous montrons les apports d'IRVO dans cette démarche pour la conception des applications de Réalité Mixte. Nous étudions également les conséquences sur l'architecture du logiciel et établissons des liens entre le modèle IRVO et le modèle AMF-C qui modélise le contrôleur de dialogue de l'application dans l'architecture préconisée par CoCSys.

Nous comparons ensuite notre méthodologie avec deux autres méthodologies de conception de systèmes mixtes en faisant ressortir les principales différences avant de conclure ce chapitre.

## VII.1 Approche basée sur les modèles

Parmi les approches de conception proposées pour le développement d'interfaces utilisateurs, les approches basées sur les modèles (MBA Model-Based Approach) bien que anciennes [Szekely 1996] sont toujours d'actualité. En effet, elles présentent plusieurs avantages [Pinheiro da Silva 2000] :

- Elles produisent une description plus abstraite des interfaces utilisateurs que les descriptions fournies par les autres outils de développement d'interfaces,
- Elles facilitent la création de méthodes pour concevoir et implémenter l'interface utilisateur d'une manière systématique car elles offrent les possibilités :
  - de modéliser les interfaces à différents niveaux d'abstraction,
  - de raffiner les modèles de manière incrémentale,
  - de réutiliser les spécifications d'interfaces utilisateurs.
- Elles fournissent l'infrastructure requise pour automatiser les processus de conception et d'implémentation des interfaces utilisateurs.

Cependant, malgré ces avantages, les approches basées sur des modèles ne se sont pas généralisées car on leur reproche la complexité des modèles et des formalismes qui sont difficiles à apprendre et à utiliser [Vanderdonckt 1999], [Pinheiro da Silva 2000].

### VII.1.1 Liste des modèles dans une approche basée sur les modèles

Nous détaillons ici les différents modèles qui composent une approche basée sur les modèles. Les noms utilisés sont les plus communément admis mais il existe une certaine variation dans le vocabulaire employé selon les auteurs. De plus, la limite de chaque modèle est également imprécise

car il existe un recouvrement non négligeable entre ceux-ci. Enfin, cette liste, sans se prétendre exhaustive, est relativement complète bien que, selon leur degré de sophistication, chaque approche MBA utilise un sous-ensemble plus ou moins important de ces modèles.

### Modèle utilisateur

Le *modèle de l'utilisateur* (User Model) représente les différentes caractéristiques des utilisateurs finaux et les rôles qu'ils jouent tant dans ses dimensions dynamiques que statiques. Il peut contenir les habitudes de l'utilisateur, ses compétences, ses expériences, son profil cognitif, ses préférences, des paramètres de personnalisation, etc. Cela permet de personnaliser le format de l'information présentée par l'interface (langue par exemple) ou le contenu de l'interface (par exemple filtrage des résultats d'une requête à une base de données).

### Modèle du domaine

Le *modèle du domaine* (Domain Model) représente les données et les opérations que l'application supporte [Szekely 1996]. Paulo Pinheiro da Silva [Pinheiro da Silva 2000] nomme ce modèle *Modèle d'application* (AM=Application model). Il peut se définir en terme de classes d'objets, d'attributs et d'opérations associés à ces objets, et de relations entre les classes et les objets.

### Modèle de tâche

Le *modèle de tâche* (Task Model) décrit les tâches que l'utilisateur peut réaliser avec l'application. Les modèles de tâches sont typiquement représentés sous forme arborescente décomposant chaque tâche en sous tâches et ainsi de suite jusqu'aux tâches feuilles qui représentent les opérations de l'application [Szekely 1996].

Paulo Pinheiro da Silva [Pinheiro da Silva 2000] propose un *Modèle de Tâche-Dialogue* (TDM=Task-Dialogue model). Il propose de regrouper les modèles de tâches avec les modèles de dialogue (Cf. ci-après) car ils modélisent tous les deux les tâches possibles qu'un utilisateur peut accomplir en interagissant avec l'application mais à des niveaux d'abstraction différents.

### Modèle de présentation

Le *modèle de présentation* définit les éléments visuels qu'une interface utilisateur propose aux utilisateurs. Classiquement, on considère 2 niveaux [Pinheiro da Silva 2000] :

- *Modèle de présentation abstraite* (APM=Abstract presentation model),
- *Modèle de présentation concrète* (CPM=Concrete presentation model).

Le *modèle de présentation abstraite* représente la structure et le contenu de l'interface en terme d'abstractions [Szekely 1996] :

- Les *objets abstraits d'interaction* (AIO=Abstract interaction objects) sont des tâches d'interface de bas niveau comme sélectionner un élément d'une liste ou afficher une unité de présentation,
- Les *éléments d'information* (information elements) représentent les données à afficher, soit constantes comme les étiquettes, soit venant du modèle du domaine,
- Les *unités de présentation* (presentation units) qui sont les abstractions des fenêtres. Elles spécifient une collection de AIO et d'éléments d'information qui doivent être présentés à l'utilisateur comme un ensemble.

Le *modèle de présentation concrète* spécifie le style pour afficher les unités de présentation, et les AIO et les éléments d'information qu'elles contiennent. Il représente l'interface en termes de primitives de la boîte à outils comme des fenêtres, des boutons, des menus et de primitives graphiques comme des lignes, des images, du texte, etc. En plus, la spécification concrète décrit la disposition des éléments dans la fenêtre (layout).

### Modèle de dialogue

Le *modèle de dialogue* (Dialog Model) décrit les aspects dynamiques de l'interface utilisateur, l'aspect statique étant décrit par le Modèle de présentation. Ce modèle définit comment les utilisateurs peuvent interagir avec les objets de présentation et les réactions que l'interface utilisateur communique via ces objets.

Les aspects dynamiques du modèle peuvent être exprimés par différents formalismes, comme par exemple, les grammaires hors-contexte, les diagrammes d'états et les réseaux de Petri. Un formalisme comme les réseaux de Petri a l'avantage d'être rigoureux car fondé sur des bases mathématiques strictes ce qui permet de faire des évaluations formelles du dialogue. Le principal inconvénient est la complexité de la modélisation dès que l'on s'attaque à des applications de grande taille.

### Modèle de plateforme

Le *modèle de plateforme* décrit les caractéristiques des ordinateurs (système d'exploitation, processeur, vitesse, taille mémoire, taille disque, etc.) et surtout des périphériques utilisés : souris (nombre de boutons), écran (taille, résolution, nombre de couleurs), etc.

### Modèle d'environnement

Parfois présent, le *modèle d'environnement* spécifie le contexte physique et culturel de l'interaction. Par exemple, pour des applications mobiles il peut s'agir de la localisation géographique actuelle de l'utilisateur, les contraintes introduites par cette localisation, le temps, etc.

## VII.1.2 Les modèles par rapport à l'architecture du logiciel

Il est possible de positionner les différents modèles énoncés ci-dessus dans le modèle d'architecture des logiciels interactifs Arch (Cf. § IV.2.1.2) comme le montre la Figure 84. On peut interpréter les modèles comme les spécifications pour le développement de la ou les couches correspondantes du modèle Arch ; pour certains modèles la correspondance est assez directe :

- Le modèle de présentation correspond aux deux couches de présentation et interaction abstraites et concrètes.
- Le modèle de dialogue correspond logiquement au contrôleur de dialogue qui gère la dynamique de l'interface homme-machine de l'application.
- Le modèle du domaine correspond directement aux spécifications du noyau fonctionnel de l'application.
- Le modèle de plateforme est plus éclaté car les périphériques physiques se trouvent tant du côté de l'utilisateur pour les entrées/sorties interactives que du côté du noyau fonctionnel pour l'accès aux supports physiques (stockage, réseaux, etc.).

D'autres modèles ont une relation moins directe avec l'architecture Arch :

- Le modèle de l'utilisateur semble *a priori* externe au modèle Arch. Cependant il influence la conception des couches de présentation comme expliqué ci-dessus.
- Le modèle d'environnement a aussi une influence tant sur le noyau fonctionnel que sur les couches présentation mais non directement représentable sur le schéma.
- Le modèle de tâche est quant à lui à un autre niveau d'abstraction comme l'explique Paulo Pinheiro da Silva [Pinheiro da Silva 2000] mais avec une couverture semblable au modèle de dialogue.

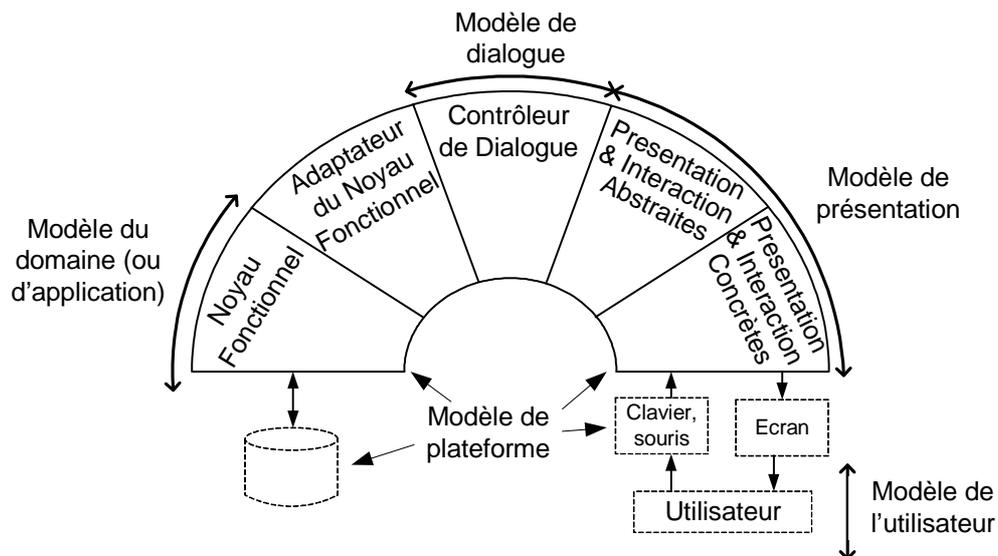


Figure 84 : liens entre une approche basée sur les modèles et le modèle Arch.

### VII.1.3 Positionnement de IRVO dans une approche basée sur les modèles

En reprenant les objectifs d'IRVO nous constatons qu'il couvre les aspects liés à l'interaction entre les utilisateurs et les artefacts manipulés et les relations entre ces différents artefacts (outils, objets de la tâche et transducteurs).

Classiquement les utilisateurs sont modélisés par le **modèle de l'utilisateur**. IRVO représente explicitement l'utilisateur et les canaux de communication qui sont utilisés dans les interactions avec les objets et les autres utilisateurs : il couvre donc une partie de ce modèle.

Parmi les objets du monde numérique (appelés objets et outils virtuels dans IRVO), seuls les objets « appréhendables » sont représentés, c-à-d les objets que l'utilisateur peut percevoir et/ou avec lesquels il peut interagir. Classiquement, ces objets sont les unités de présentation et leur modélisation est couverte par le **modèle de présentation**. De ce point de vue, on peut donc dire que IRVO se rapproche le plus de ce modèle mais en se limitant aux objets de la tâche et sans détailler les autres éléments de l'interface comme les boutons ou les menus de l'application. De plus, il n'aborde pas les détails d'apparence et de « mise en page » des éléments d'interface comme dans un modèle de présentation classique. IRVO se trouve donc dans un niveau intermédiaire entre le modèle de présentation abstraite et le modèle de présentation concrète.

Trévisan et al. (Cf. § VII.4.1) proposent d'étendre le **modèle du domaine**, pour les environnements de Réalité Mixte, aux objets réels directement manipulés par les utilisateurs. Nous n'adhérons pas complètement à cette proposition, car cela dépend en fait du type d'application considéré. Par exemple, dans le cas des interfaces tangibles, les objets réels ne servent qu'à agir sur les objets numériques ; ils ne participent pas au cœur de l'application mais seulement à l'interface et il est donc faux de les ranger du côté du modèle du domaine. Par contre, dans des cas où l'objet de la tâche est réel (comme en chirurgie par exemple), les objets réels doivent être modélisés par le noyau fonctionnel de l'application et donc de ce point de vue alimentent le modèle du domaine.

Enfin, IRVO représente les transducteurs qui, comme nous l'avons montré au § VI.2.6 sont proches de la notion classique de périphérique d'entrée/sortie. Ainsi, IRVO couvre une partie du **modèle de plateforme** puisqu'il participe à la modélisation des périphériques.

Finalement, on peut dire que le modèle IRVO, de part sa centration sur les outils et les objets manipulés par les utilisateurs, se rapproche conceptuellement le plus du **modèle de présentation à un niveau semi-concret**. Il a cependant une couverture plus large et **permet d'établir un pont entre les différents modèles** d'une approche basée sur les modèles :

- Il participe au modèle de l'utilisateur car il modélise explicitement le ou les utilisateurs et leurs interactions (entre eux et avec les artefacts).
- Il représente des objets physiques qui, dans certains cas, alimentent la définition du modèle du domaine (Cf. discussion ci-avant).
- Il représente les transducteurs d'où découlent les périphériques d'entrée/sortie ce qui participe au modèle de plateforme.
- Il établit les relations entre tous ces éléments du point de vue de l'interaction avec l'utilisateur.

## VII.2 Arbre de tâches et modèle IRVO

Dans cette section, nous examinons plus particulièrement le modèle de tâches et plus exactement les arbres de tâches qui sont la représentation la plus communément admise pour ce modèle.

Le modèle de tâches est la clef de voûte des approches basées sur les modèles car elle se trouve en amont de la conception d'autres modèles. En particulier, il est généralement admis [Pinheiro da Silva 2000] que le modèle de dialogue découle en grande partie du modèle de tâches car il s'agit de l'instanciation des tâches à accomplir dans l'interface du logiciel en termes d'actions de l'utilisateur et de réactions du système.

Nous avons vu que IRVO modélise l'interaction au niveau d'un tâche élémentaire. Il est donc important d'étudier les relations entre IRVO et l'arbre de tâches. C'est à partir de ces relations que nous pouvons relier les modélisations IRVO de chaque tâche élémentaire entre elles et construire une modélisation globale de l'application.

Dans un premier temps nous rappelons les principales définitions, puis nous présentons quelques formalismes dont le formalisme CTT que nous avons retenu pour modéliser l'arbre de tâches. Enfin nous montrons les relations entre l'arbre de tâches et les modélisations IRVO d'où nous déduisons des propriétés ergonomiques globales de l'application.

### VII.2.1 Définitions

D'après [Mori 2002], on peut définir une tâche comme une activité qui doit être effectuée pour atteindre un but. Un but est soit une modification désirée d'un état ou une requête pour obtenir de l'information sur l'état courant. Les tâches peuvent aller d'un niveau élevé d'abstraction à un niveau très concret orienté vers des actions simples.

Les modèles de tâches cherchent donc à **analyser et à représenter l'activité** d'un ou plusieurs utilisateurs sous la forme d'une structure de tâches qui décrit ce que l'utilisateur peut ou doit faire.

Ils peuvent être utilisés à des fins descriptives (représentation des tâches supportées par un système existant) ou prescriptives (prescription des tâches sur un système en cours de développement).

Les modèles de tâches sont aussi utilisés à des fins d'évaluation en particulier pour comparer la tâche réelle d'un utilisateur à la tâche prescrite.

### VII.2.2 Présentation de quelques formalismes

Il existe de nombreux formalismes pour la représentation des modèles de tâches. Parmi les plus connus, nous pouvons citer d'après [Mori 2002] :

- GOMS (Goals, Operators, Methods, Selection Rules) : il permet de décrire des séquences d'actions nécessaires pour atteindre un but et de les grouper dans des méthodes qui peuvent être associées à une règle de sélection qui permet de savoir quand leur exécution est recommandée. Cependant GOMS ne considère que des tâches séquentielles, ce qui limite grandement son usage surtout dans le contexte d'interfaces avancées.

- UAN (User Action Notation) possède un ensemble plus riche d'opérateurs temporels. C'est un formalisme textuel avec lequel on peut décrire les relations temporelles entre différentes tâches asynchrones et, au final, les tâches élémentaires sont décrites par des tables qui spécifient : les actions de l'utilisateur, le retour du système, et les modifications de l'état de l'application. Son formalisme textuel a cependant été un frein à sa diffusion.

### VII.2.2.1 Formalisme MAD

**MAD** (Méthode Analytique de Description des tâches) est basé sur une méthodologie issue de l'ergonomie et développé par Dominique Scapin et son équipe [Pierret-Golbreich 1989]. Ce formalisme s'appuie sur une décomposition hiérarchique des tâches reliées par des constructeurs : SEQ (pour exprimer l'exécution séquentielle des tâches), ALT (exprime l'alternative en plusieurs tâches), PAR (exprime l'exécution entrelacée de plusieurs tâches par un même agent), SIM (exprime l'exécution simultanée de plusieurs tâches par des agents distincts). Chaque tâche est caractérisée par :

- Une identification : numéro et nom de la tâche,
- Des éléments de structures : but, état initial, pré-conditions, corps de la tâche, post-conditions, état final,
- Des attributs : FAC (tâche facultative), @ (boucle), PRIOR (tâche prioritaire), INTER (tâche interruptible) et un niveau de priorité.

Le formalisme a évolué sous le nom de **MAD\*** (Modèle Analytique de Description des tâches utilisateur orienté spécification d'interface) [Scapin 2001]. Les ajouts les plus notables concernent l'apport des concepts orientés objets (introduction de classes de tâches, de l'héritage et des instances) ainsi que l'augmentation du nombre de constructeurs (rebaptisés opérateurs dans MAD\*) et leur structuration en quatre catégories : opérateurs de synchronisation (SEQuentiel, PARallèle, SIMultané), opérateurs d'ordonnancement (ET, OU, alternatif), les opérateurs temporels (début, fin, durée) et les opérateurs auxiliaires (élémentaire, inconnu).

Un exemple utilisant le formalisme MAD est donné au Chapitre VIII, § VIII.3.2.

### VII.2.2.2 Formalisme CTT

Pour notre part, notre préférence va actuellement au formalisme CTT (ConcurTaskTrees) [Paternò 2000, Mori 2002], pour plusieurs raisons :

- Ce formalisme a été plus spécifiquement conçu pour la modélisation des interfaces homme-machine dans un but prédictif,
- Il prend en compte explicitement le travail collaboratif,
- Un éditeur graphique, CTTE, est disponible.

Ce formalisme considère 4 catégories de tâches :

- Les tâches de l'utilisateur ; ce sont des tâches cognitives comme choisir une méthode de résolution de problème.
- Les tâches de l'application ; ce sont des tâches où seul le système informatique intervient comme produire le résultat à une requête.
- Les tâches d'interaction ; ce sont des actions de l'utilisateur sur le système avec la possibilité d'un retour immédiat.
- Les tâches abstraites ; ce sont des tâches de plus haut niveau qui se décomposent en sous-tâches pouvant appartenir à l'une des 4 catégories.

Les tâches sont reliées par des opérateurs temporels. Il en existe huit catégories principales qui sont :

- La concurrence (notée  $T1 \parallel T2$ ) : les 2 tâches peuvent être exécutées en parallèle sans aucune contrainte. Il existe une variante (notée  $T1 \parallel\parallel T2$ ) où les 2 tâches doivent se synchroniser pour échanger de l'information.
- Le choix (noté :  $T1 \square T2$ ) : les 2 tâches sont alternatives et le début d'exécution de l'une empêche l'autre de s'exécuter tant que la première n'est pas terminée.
- L'indépendance d'ordre (notée  $T1 \mid = \mid T2$ ) : les 2 tâches doivent être exécutées mais dans un ordre quelconque T1 puis T2 ou T2 puis T1.
- La désactivation (notée  $T1 [ > T2$ ) : la tâche T1 est définitivement stoppée dès que la tâche T2 commence à s'exécuter.
- L'activation (notée  $T1 \gg T2$ ) : lorsque la tâche T1 est terminée la tâche T2 commence à s'exécuter. Une variante (notée  $T1 \square\gg T2$ ), permet de signifier qu'en plus la tâche T1 transmet des informations à la tâche T2.
- La suspension-reprise (notée  $T1 \mid > T2$ ) : la tâche T2 peut suspendre l'exécution de la tâche T1. Lorsqu'elle est terminée, la tâche T1 reprend son exécution là où elle a été interrompue.
- L'itération (notée  $T^*$ ) : la tâche est exécutée en boucle indéfiniment jusqu'à ce qu'elle soit désactivée par une autre tâche. Une variante (notée  $T(n)$ ) permet de limiter la boucle à n itérations.
- Les tâches optionnelles (notées  $[T]$ ) : leur exécution est facultative.

Chaque tâche est décrite par un ensemble d'attributs :

- Des informations générales : un identificateur, son nom, sa catégorie (Cf. ci-dessus), son type, sa fréquence d'utilisation, des pré-conditions éventuelles, et des annotations libres.
- Les objets (éventuels) qui sont manipulés pour accomplir la tâche. Ce sont soit des objets de l'interface utilisateur, soit des objets de l'application (Cf. Modèle du domaine).
- Temps d'exécution : il est possible d'indiquer des temps d'exécution de la tâche (minimum, maximum, moyen) pour permettre des évaluations de performance de tâches.

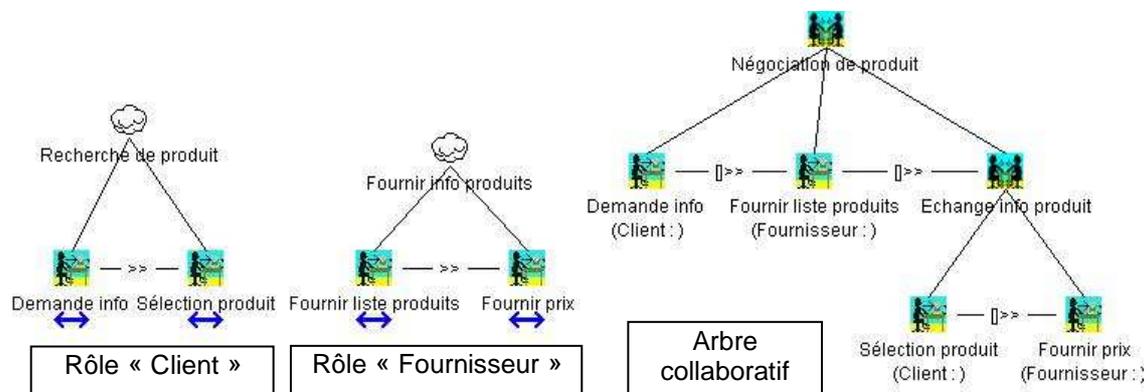


Figure 85 : Exemple de modélisation de tâches collaboratives avec CTT (d'après [Mori 2002]).

CTT prend en compte explicitement les tâches collaboratives. Le modèle de tâches dans ce cas se compose de plusieurs arbres de tâches (Figure 85) :

- Pour chaque rôle identifié, un arbre de tâches individuel est construit. Ces arbres sont similaires aux arbres mono utilisateurs décrits ci-dessus.
- Un arbre de tâches coopératif exprime les relations entre les différentes tâches individuelles. La principale différence est qu'il inclut des tâches coopératives qui impliquent des actions par deux ou plusieurs utilisateurs pour les accomplir. Ces tâches coopératives sont décomposées hiérarchiquement en sous-tâches jusqu'à ce qu'on atteigne des tâches exécutées par un seul utilisateur. Ces tâches individuelles apparaissent aussi dans le modèle de tâches du rôle associé : elles sont appelées tâches de connexion et une flèche double les différencie des autres tâches.

### VII.2.3 Lien entre IRVO et l'arbre de tâche

Le rôle de l'arbre de tâches est d'exprimer le découpage en tâches plus ou moins microscopiques et leur organisation sous forme d'arbre. IRVO permet d'exprimer les objets (outils et objets du domaine de la tâche) participant à la réalisation d'une tâche. On peut donc dire que IRVO apparaît naturellement au niveau des tâches élémentaires (feuilles de l'arbre) : à chaque tâche élémentaire peut être associé un schéma IRVO modélisant l'interaction de l'utilisateur avec les objets dans le contexte de cette tâche (Figure 86, Tâches 11 à 22). Si plusieurs tâches partagent les mêmes artefacts alors il est naturel de remonter jusqu'à la sous-racine de l'arbre de tâches qui partage les mêmes artefacts pour mettre en facteur le schéma IRVO et les sous-tâches correspondantes (Figure 86, Tâche 3).

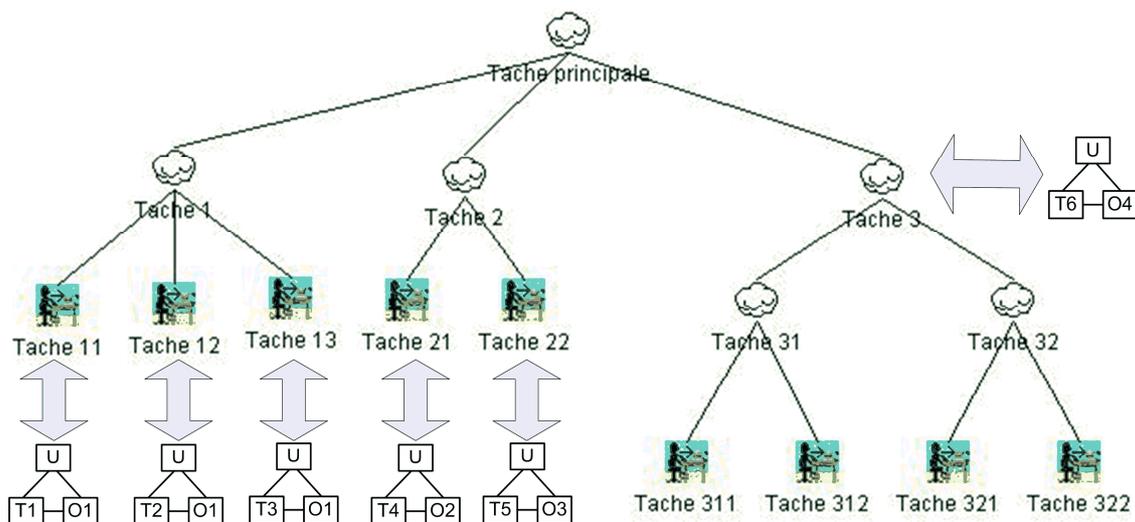


Figure 86 : Liens entre l'arbre de tâches et les schémas IRVO.

#### VII.2.3.1 Fusion de schémas IRVO et relations temporelles

Même dans le cas où les tâches ne partagent pas les mêmes outils et objets du domaine, nous pensons qu'il est utile de fusionner les différents schémas IRVO de plusieurs sous-tâches d'une tâche donnée. Ainsi, le schéma fusionné correspond au schéma IRVO de cette tâche plus globale. Comme il existe plusieurs objets sur ce schéma de fusion, il est nécessaire d'indexer ces objets par le numéro ou le nom de la tâche à laquelle ils participent.

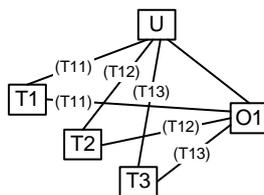


Figure 87 : Exemple de schéma IRVO correspondant à la Tâche 1.

L'intérêt de cette approche est de pouvoir représenter sur un même schéma plusieurs tâches très proches qui utilisent quasiment les mêmes outils et objets du domaine. S'il y a trop d'objets différents, le schéma devient vite très touffu et cette opération perd donc de son intérêt. On remarquera que les différentes entités restant rattachées à une tâche précise, les relations temporelles exprimées par l'arbre de tâches sont donc directement transposables aux artefacts correspondants. Cette propriété explique qu'il ne semble donc pas utile *a priori* d'introduire des relations temporelles directement dans le formalisme d'IRVO puisqu'elles sont déjà exprimées dans l'arbre de tâches.

Cependant, à l'intérieur même d'une tâche il existe une relation temporelle fine dans l'interaction entre l'utilisateur et le système : l'utilisateur agit **puis** il perçoit la réponse. Pour l'instant, cette séquence est implicite dans IRVO : l'action part de l'utilisateur qui agit avec les outils sur les objets et il perçoit le résultat de son action (boucle action – perception au sens de Norman).

Nous avons déjà vu au paragraphe VI.4.3 que les transducteurs pouvaient apporter des délais importants à un point tel que l'interaction peut devenir difficile voire impossible. Dans un but d'analyse prédictive d'un système en cours de développement, il est donc essentiel de pouvoir prendre en compte ces aspects temporels dans le modèle et il semble inévitable qu'à l'avenir il faudra ajouter des opérateurs temporels au formalisme d'IRVO.

Ceci peut aussi être particulièrement pertinent pour des interactions complexes ou interviennent plusieurs utilisateurs et dont les actions peuvent dépendre les uns des autres.

### VII.2.3.2 Schémas de synthèse d'une application

Il semble souhaitable de continuer à remonter dans l'arbre de tâches jusqu'à la racine en fusionnant les différents schémas IRVO ce qui conduit à obtenir un schéma de synthèse montrant tous les outils et objets du domaine participant à l'ensemble des activités. Il est intéressant d'exprimer sur ce schéma de synthèse la « fréquence » d'utilisation des différents constituants : plusieurs représentations sont possibles, notre préférence allant actuellement à la représentation en faisceaux (Figure 88). En analysant ce schéma de synthèse nous pouvons facilement détecter des configurations singulières (cas de T3/O3 dans Figure 88) qui doivent être examinées en vue de déterminer la raison d'être (tâche très spécifique) ou leur re-conception éventuelle dans le respect de la règle de continuité d'interaction [Florins 2004].

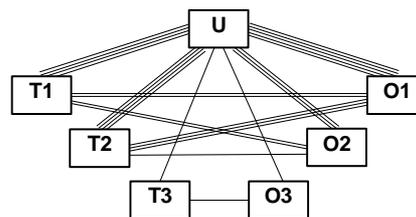


Figure 88 : Exemple de schéma de synthèse d'une application mono-utilisateur.

La même démarche de fusion d'arbres peut aussi être effectuée sur l'arbre de tâches collaboratif (Figure 89). Dans ce cas, on peut aisément voir apparaître les outils qui sont partagés par plusieurs utilisateurs (cas de T1 et T2) et ceux qui sont individuels (cas de T3 utilisé par U3). On pourrait aussi détecter que certains objets ne sont pas visualisés par tous les utilisateurs ce qui peut révéler un problème de conception ou être normal si le rôle de cet utilisateur ne l'autorise pas à accéder à cet objet.

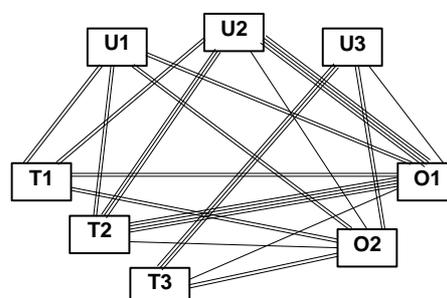


Figure 89 : Exemple de schéma de synthèse d'une application collaborative.

### **VII.2.4 Bilan de la liaison entre IRVO et l'arbre de tâches**

Dans cette section, nous avons détaillé le modèle de tâches (et plus exactement les arbres de tâches qui sont la représentation la plus communément admise pour ce modèle) car c'est l'un des modèles principaux des approches basées sur les modèles. Mais surtout nous avons montré que l'arbre de tâches était le moyen de relier toutes les modélisations IRVO d'une application car chaque schéma IRVO représente l'interaction de l'utilisateur avec le système au niveau d'une tâche élémentaire.

En remontant l'arbre de tâches, il devient donc possible de fusionner des schémas IRVO et, en appliquant le principe jusqu'à la racine d'obtenir un schéma de synthèse de l'application. Ce schéma de synthèse permet de facilement détecter des configurations singulières (d'outils ou d'objets du domaine) qui doivent être examinées en vue de déterminer la raison d'être (tâche très spécifique) ou leur re-conception éventuelle dans le respect de la règle de continuité d'interaction.

## **VII.3 Vers une méthodologie de conception de systèmes de Réalité Mixte et collaboratifs**

Dans cette section nous présentons une méthodologie de conception pour des systèmes de Réalité Mixte et collaboratifs basée sur une approche MBA. Plutôt que de développer une nouvelle méthodologie nous cherchons à nous inscrire dans une méthodologie existante, CoCSys qui est en cours de développement au laboratoire ICTT et qui est initialement adaptée aux systèmes mobiles et collaboratifs.

Dans un premier temps nous décrivons la méthodologie CoCSys. Ensuite nous examinons où et comment le modèle IRVO s'insère dans cette méthodologie au niveau du modèle comportemental. Enfin, nous examinerons comment, dans la phase de développement de l'application, le modèle IRVO permet d'assister la construction de l'architecture du logiciel en établissant des liens avec le modèle d'architecture AMF-C.

### **VII.3.1 Présentation de la méthodologie CoCSys**

CoCSys (acronyme de Cooperative Capillary Systems) est une méthodologie développée par notre laboratoire à l'Ecole Centrale de Lyon [David 2003a, Delotte 2003, David 2004]. Elle est orientée initialement vers la conception de systèmes mobiles et collaboratifs que nous appelons Travail Coopératif Capillaire (TCC).

La méthode s'appuie sur un ensemble de scénarios (Figure 90, item 1) au sens de Carroll [Carroll 1997] et/ou de cas d'utilisation UML qui servent à la construction du modèle de comportement collaboratif (CAB=Collaborative Application Behavior Model).

Le but principal du modèle CAB (Figure 90, item 2) est de décrire explicitement les acteurs, les artefacts, les contextes et la structure de l'arbre des tâches qui caractérisent le comportement de l'application coopérative visée. Les tâches exprimées dans différents scénarios sont analysées puis organisées en arbre de tâches. L'objectif est d'éliminer les redondances et d'élaborer l'arbre des tâches aussi propre que possible. Dans le cas de la présence de scénarios conduisant à la mise en place de dépendances entre tâches et/ou acteurs, ceci se traduit par la mise en place d'un workflow, qui est représenté de façon implicite, pour être évolutif. La construction du modèle CAB dépend de la qualité et de la quantité des scénarios.

L'infrastructure du système coopératif est organisée en un ensemble de 3 couches spécialisées (Figure 90, item 3) :

- Le niveau le plus haut est celui de l'application collaborative elle-même. Son but est de superviser les interactions et de proposer les interfaces de présentation. Il est appelé Modèle CUO (Collaboration User-Oriented Model). Le modèle AMF-C (Cf. § V.3.3) est le modèle proposé pour tenir ce rôle.

- Le niveau intermédiaire est le Modèle CSA (Collaborative System Architecture Model). C'est un niveau générique entre l'application et le système distribué. Il contient les éléments communs de l'activité du groupe et peut être vu comme un « système d'exploitation » dédié aux collectifs. Il supporte le travail collaboratif en contrôlant les sessions, les utilisateurs et les groupes, fournit des outils coopératifs génériques (par exemple, un télépointeur) et est responsable de la gestion de la concurrence.
- Le dernier niveau est en charge essentiellement de la distribution des messages et de la gestion du contrôle du contenu. Il est appelé Modèle DSI (Distributed System Infrastructure Model). C'est un niveau orienté système qui fournit les mécanismes pour la communication et la synchronisation des composants distribués qui ne sont pas adaptés au travail collaboratif mais qui sont utiles à ce niveau technique.

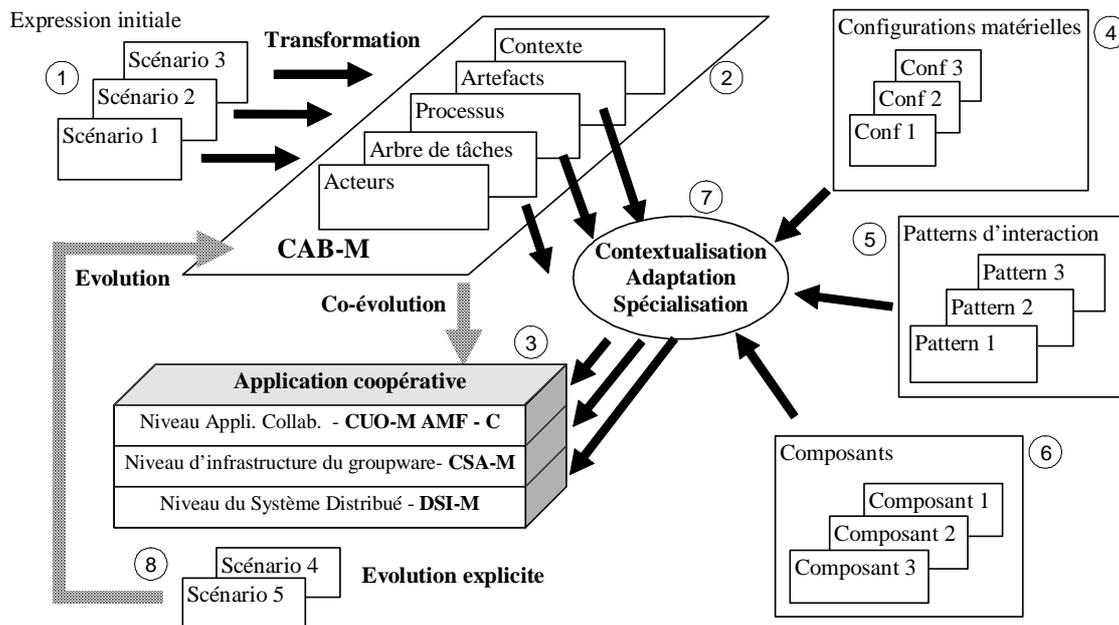


Figure 90 : Principaux constituants de l'environnement de conception d'applications collaboratives mobiles et adaptatives (d'après [Delotte 2003])

Le processus de **contextualisation, d'adaptation et de spécialisation** (Figure 90, item 7) assure le passage entre le modèle comportemental et sa mise en œuvre sur les dispositifs réels de manière la plus automatique possible (en accord avec l'approche MDA [OMG 2004]). Il utilise un ensemble de mécanismes afin de transposer les comportements en services informatiques situés sur les trois niveaux du modèle d'architecture collaborative. En effet, chaque élément du modèle CAB est en relation plus ou moins importante avec les trois niveaux du modèle d'architecture logicielle. C'est naturel pour la couche d'interaction, mais c'est également le cas pour d'autres aspects, comme par exemple la nature de réseau à prendre en compte selon le contexte exprimé et l'infrastructure matérielle visée, qui concerne la couche DSI.

Pour tenir compte de la multiplicité des terminaux et des interfaces des dispositifs mobiles, la transformation du modèle CAB vers l'architecture logicielle prend en compte les plates-formes matérielles disponibles et s'appuie sur des catalogues de patterns d'interaction (et de composants logiciels).

Les configurations matérielles (Figure 90, item 4) décrivent des caractéristiques, liées à celles-ci, qui ont un impact sur la spécialisation à prendre en compte, notamment en matière d'interactions, d'échanges d'information et en capacités de traitement et de stockage. Le processus d'adaptation filtre donc en tenant compte des caractéristiques des configurations disponibles. Par exemple,

L'arbre des tâches est ainsi spécialisé en un arbre des tâches spécifique en éliminant des tâches non pertinentes.

Les patterns d'interaction proposés (Figure 90, item 5) permettent de choisir pour les différents éléments du modèle CAB les interactions à mettre en place en tenant compte de la configuration matérielle disponible.

Les composants (Figure 90, item 6) sont les éléments logiciels mis à la disposition du système et qui peuvent être intégrés lors de l'adaptation. Il s'agit par exemple d'applications spécifiques (progiciels, ...) dont la portabilité doit être gérée en amont.

Enfin, la méthodologie prend en compte l'évolution ultérieure du système. Il peut s'agir d'une adaptation légère de l'application (qui peut éventuellement être faite par l'utilisateur lui-même) ou de la prise en compte de nouveaux comportements qui apparaissent au fil de l'usage. Dans ce cas, cette adaptation passe par la création de nouveaux scénarios (Figure 90, item 8), qui correspondent à la situation à prendre en compte et qui doivent être intégrés au modèle comportemental CAB.

### VII.3.2 Insertion du modèle IRVO dans le modèle comportemental

Nous proposons d'insérer IRVO comme un nouveau modèle dans le modèle CAB (Collaborative Application Behavior) pour compléter les autres modèles déjà présent (Figure 91). Le but de cette introduction est de couvrir plus complètement la modélisation des systèmes de Réalité Mixte en proposant un modèle spécifiquement adapté à la conception de tels systèmes. Plus généralement IRVO peut servir à la conception de toute application qui envisage des interactions « non-standard » comme l'interaction bi-manuelle (Cf. § VI.3.1.5).

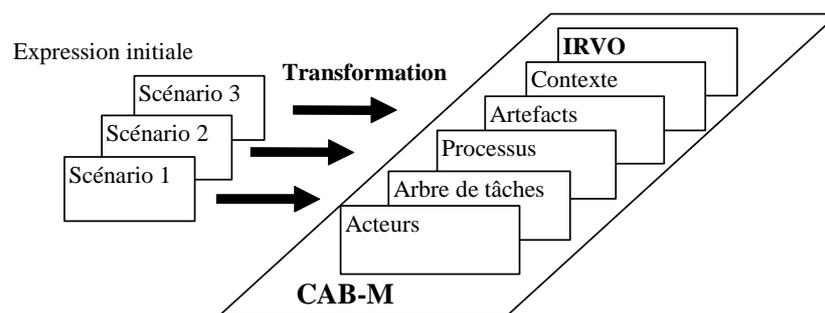


Figure 91 : première phase de la méthodologie CoCSys avec le modèle comportemental étendu.

Nous allons donc examiner les impacts de l'ajout de ce modèle au niveau de la méthodologie CoCSys et décrire les adaptations et ajouts dans les deux principales étapes de la conception qui sont la phase de recueil des scénarios et la phase de construction du modèle comportemental

#### VII.3.2.1 Phase de recueil des scénarios

Comme dans la méthodologie CoCSys initiale il s'agit tout d'abord de recueillir les scénarios qui décrivent des situations d'utilisation et des actions à effectuer. Les scénarios peuvent être utilisés à des fins descriptives pour décrire les situations actuelles de travail ou à des fins prescriptives pour décrire les futures situations d'utilisation du système.

La notion de scénario a particulièrement été étudiée en IHM par Carroll [Carroll 1997]. Les scénarios sont des petites histoires qui décrivent des situations de travail dans des contextes d'utilisation précis. Cette perception analytique des situations de travail peuvent être capturées et exprimées soit par des observateurs, soit par les acteurs eux-mêmes dans une approche de conception participative.

Les scénarios sont concrets car ils sont une représentation d'une solution spécifique dans un contexte donnée, mais ils sont souvent généraux en spécifiant les tâches que l'utilisateur peut

effectuer sans descendre à des niveaux de détails élémentaires. Cependant, nous pensons que pour les situations nouvelles, en particulier les tâches où il est envisagé d'utiliser des interactions utilisant des dispositifs d'interaction avancés comme la Réalité Mixte, les scénarios doivent descendre à un degré de détail suffisant pour anticiper au mieux ces situations qui peuvent être en rupture avec des pratiques acquises avec les interfaces homme-machine traditionnelles.

### Apports d'IRVO à cette étape

Dans la méthodologie CoCSys, même si les scénarios textuels sont privilégiés, il n'est pas exclu de les compléter par d'autres représentations comme des cas d'utilisations (exprimés en UML) ou des diagrammes, dessins, etc.

Dans ce contexte, il est tout à fait envisageable d'utiliser IRVO à cette étape pour représenter des situations qui font intervenir la Réalité Mixte. Dans ces cas il ne s'agit pas de construire une modélisation complète de l'interaction mais juste de schématiser les éléments principaux, typiquement les utilisateurs et les objets du domaine, en omettant les outils et les dispositifs d'entrée/sortie. Un exemple d'un tel schéma a déjà été donné au paragraphe § VI.3.3.2, Figure 81 sur l'exemple du Navicam+ et que nous reproduisons en Figure 92.

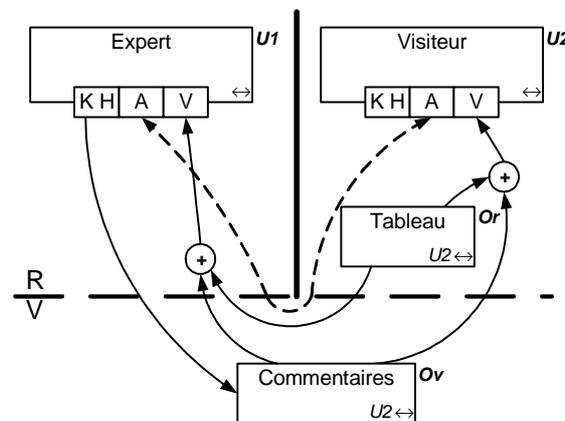


Figure 92 : Exemple d'utilisation d'IRVO pour la représentation de scénarios sur le système Navicam+

Etant donné la relative simplicité du formalisme graphique, ces schémas peuvent être compréhensibles par les utilisateurs finaux ; on peut même imaginer qu'ils participent à la construction de ces schémas avec l'assistance des concepteurs dans un processus de conception participative.

### VII.3.2.2 Elaboration du modèle comportemental (CAB)

Le modèle comportemental de l'application (CAB – Collaborative Application Behavior) est élaboré à partir des scénarios. Il s'agit donc d'extraire, à partir des descriptions textuelles, les éléments principaux comme les acteurs, les artefacts (objets du domaine et outils), les tâches, les processus et les contextes d'usages (plates-formes, périphériques et environnement).

Cette étape est essentiellement manuelle et fait intervenir un degré élevé d'expertise du concepteur qui construit graduellement le modèle CAB à partir des scénarios. Il est important de conserver les liens entre les éléments du modèle comportemental et les scénarios d'origine :

- Pour la validation du modèle : il s'agit de vérifier qu'il n'y a pas de contradiction entre le modèle et les scénarios et que tous les scénarios ont bien été pris en compte.
- Pour l'évolution du système : dans des phases d'évolution ultérieure du système de nouveaux scénarios peuvent être ajoutés et d'anciens qui n'ont plus de raison d'être, enlevés. En gardant ce lien, il est aisé d'identifier les éléments du modèle comportemental qui deviennent inutiles ou qui doivent être modifiés par les nouveaux scénarios.

Les scénarios peuvent présenter des redondances, des incohérences, des imprécisions ou des manques qui vont apparaître pendant la construction du modèle comportemental. Il s'agit de retourner auprès des utilisateurs et des experts du domaine pour corriger, préciser et compléter les scénarios afin d'affiner le modèle par itération successive jusqu'à obtenir un modèle le plus cohérent et complet possible.

Pour assister le concepteur à construire le modèle, à maintenir les liens entre les scénarios et les modèles et à vérifier la cohérence globale du modèle, un outil CBME est en cours de développement [Delotte 2004].

### *Apports d'IRVO à cette étape*

Les scénarios peuvent faire apparaître des tâches qui utilisent la Réalité Mixte. Pour chacune de ces tâches, une modélisation avec le formalisme IRVO est construit afin de formaliser l'interaction des utilisateurs avec le système et expliciter les relations avec les objets du domaine, les outils et les dispositifs d'entrée/sortie. Ces modélisations sont donc reliées :

- A des tâches (en général élémentaires) de l'arbre de tâche de l'application, comme nous l'avons expliqué au paragraphe VII.2.3,
- Au modèle des acteurs, en particulier par les rôles joués par ces acteurs dans l'interaction,
- Au modèle des artefacts, en particulier par les outils manipulés et les objets du domaine,
- Au modèle de contexte, en particulier par les dispositifs d'entrée/sortie qui sont utilisés.

La construction de ces modélisations IRVO peut s'appuyer sur un catalogue de *patterns* d'interaction élaboré au préalable. Ce catalogue n'existe pas encore mais devrait pouvoir être facilement construit à partir des 44 modélisations que nous avons faites à partir des applications de la littérature (Cf. § VI.3.2.3). Si les *patterns* permettent de gagner du temps dans la conception, ils sont en même temps un frein à la génération de nouvelles formes d'interactions originales.

C'est à ce stade que des analyses prédictives des propriétés ergonomiques peuvent être réalisées en s'appuyant sur l'arbre de tâches et sur les modélisations IRVO produites en suivant les règles de vérification énoncées au § VI.4 et complétées au § VII.2.3.2. En fonction de ces analyses, différentes modélisations alternatives peuvent être produites, et par le mécanisme itératif énoncé ci-dessus, les scénarios peuvent être modifiés pour tenir compte de problèmes sérieux d'utilisabilité qui apparaissent suite à ces vérifications de propriétés ergonomiques.

D'ailleurs, il n'est pas évident que la Réalité Mixte soit une solution envisagée *a priori* par les scénarios. On peut très bien imaginer que, pour une tâche donnée, on envisage des interactions avec un ordinateur classique ou avec un système mixte. Dans ce cas, IRVO qui permet de représenter avec le même formalisme les deux situations alternatives, facilite la comparaison des propriétés ergonomiques des deux versions mixtes et non-mixtes.

### **VII.3.2.3 Bilan**

Finalement, On voit que le modèle IRVO peut harmonieusement s'intégrer dans le modèle comportemental de la méthodologie CoCSys comme un « outil » supplémentaire à la disposition du concepteur. Le formalisme d'IRVO est utilisable à deux niveaux :

- Comme point de départ de la conception en formalisant des cas typiques d'interaction avec le futur dispositif. De ce point de vue il peut être vu comme un formalisme graphique complétant des scénarios textuels.
- Comme modèle résultant de l'analyse : après décomposition de l'application en arbre de tâches, pour chaque tâche utilisant la réalité mixte, un schéma IRVO associé est construit pour modéliser finement l'interaction de l'utilisateur avec les outils et l'objet de la tâche.

La construction des modélisations peut être assistée par un catalogue de *patterns* d'interaction présentant des cas types d'interaction en Réalité Mixte. Les différentes modélisations produites sont

évaluées du point des propriétés ergonomiques ce qui permet de comparer différentes solutions alternatives afin de retenir les meilleurs.

Cependant, de plus amples travaux sont nécessaires pour véritablement proposer une méthodologie complète et utilisable.

### **VII.3.3 Conséquences de l'introduction de IRVO sur le processus de contextualisation, adaptation et spécialisation**

Nous avons vu aux paragraphes IV.2.3.3 et V.3.5 consacrés à deux évolutions du modèle d'architecture PAC-Amodeux pour les systèmes mixtes, que la prise en compte de la Réalité Mixte avait une incidence au niveau de l'architecture des logiciels interactifs car il s'agit de traiter de nouvelles données provenant des dispositifs d'entrée/sortie utilisés dans ces systèmes en plus des dispositifs classiques comme le clavier, la souris et l'affichage graphique

Il est donc naturel que l'architecture de l'application collaborative dans CoCSys soit aussi affectée en particulier au niveau de l'application collaborative (modèle CUO) pour la prise en compte de ces nouveaux moyens d'interaction. Nous avons vu ci-dessus (§ VII.3.1) que le rôle du modèle CUO dans CoCSys était tenu par le modèle AMF-C (Cf. § V.3.4).

L'introduction de la Réalité Mixte dans CoCSys a probablement d'autres incidences dans l'architecture du modèle CUO aux niveaux de l'infrastructure du collecticiel (CSA) et du système distribué (DSI). Ces aspects ne sont pas traités ici et devront faire l'objet de recherches ultérieures.

Les autres aspects de la phase de « Contextualisation adaptation spécialisation » de la méthodologie CoCSys restent valables, comme par exemple l'utilisation de patterns d'interaction qui dans ce cas incluent des patterns adaptés à la Réalité Mixte. La encore, de plus amples investigations sont nécessaires pour réellement mesurer l'impact de l'introduction de la Réalité Mixte sur l'ensemble de cette phase.

Nous nous limitons donc dans la suite de cette section à examiner quelles modifications sont à apporter à AMF-C pour prendre en compte la Réalité Mixte. Comme le modèle CAB a été étendu avec IRVO, c'est en examinant le lien entre ces deux modèles que nous pouvons en déduire les modifications à apporter au modèle AMF-C.

#### **VII.3.3.1 Liens conceptuels entre IRVO et AMF-C**

Nous avons vu que le modèle AMF-C (Cf. § V.3.4) est un modèle de dialogue adapté aux collecticiels qui couvre fonctionnellement la couche contrôleur de dialogue mais aussi les deux couches adjacentes du modèle Arch, l'adaptateur du noyau fonctionnel et la présentation et interaction abstraites. D'ailleurs AMF-C est non seulement un formalisme pour spécifier le modèle de dialogue mais est aussi le contrôleur de dialogue lui-même car, grâce au « moteur AMF », le formalisme graphique d'AMF utilisé pour modéliser la facette « contrôle », peut être vu comme un véritable langage de programmation visuelle qui pilote le dialogue au sein même de l'application finale (Cf. § IV.2.3.2).

Pour sa part IRVO de par son rapprochement avec le modèle de présentation (Cf. § VII.1.3) est nettement du côté des couches présentations et de l'interface concrète.

La Figure 93 résume ce rapprochement conceptuel en se basant sur le modèle d'architecture Arch. Dans ce schéma, l'entité 'M' de IRVO apparaît clairement comme étant la « boîte noire » des 4 éléments de gauche du modèle Arch, cette boîte noire étant elle-même détaillée par le modèle AMF-C. C'est donc l'entité 'M' qui va nous permettre de faire le lien entre le modèle d'interaction IRVO et le modèle d'architecture AMF-C.

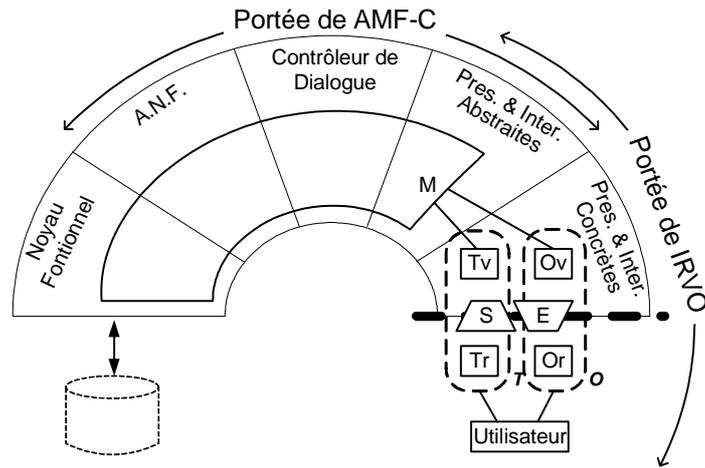


Figure 93 : Relations conceptuelles entre les modèle Arch, AMF-C et IRVO.

### VII.3.3.2 Correspondances entre agents AMF-C et objets du domaine IRVO

Les agents AMF modélisent les objets manipulables par l'utilisateur, la facette présentation permettant à l'utilisateur d'interagir avec cet agent. Il est donc immédiat d'associer un objet virtuel IRVO à la facette présentation de l'agent AMF-C. Pour les objets mixtes, ceci conduit à ne modéliser avec AMF que la partie virtuelle de l'objet, la partie réelle n'étant pas représentée avec AMF. Cependant, cette séparation est absurde car c'est bien l'objet mixte complet qui doit être considéré comme un tout cohérent. C'est la raison pour laquelle nous proposons de rajouter une **nouvelle facette « Réel »** aux agents AMF qui met en correspondance la partie réelle des objets mixtes avec l'agent AMF modélisant l'agent. Cette facette est donc sémantiquement assimilable à une facette de présentation mais n'est pas visible en tant que telle par l'utilisateur.

Concrètement, nous associons un agent AMF-C à tout objet mixte IRVO (Figure 94) :

- A la partie virtuelle de l'objet (Ov) nous associons la facette « Présentation » classique de l'agent AMF,
- A la partie réelle de l'objet (Or) nous associons la nouvelle facette « Réel » de l'agent AMF.

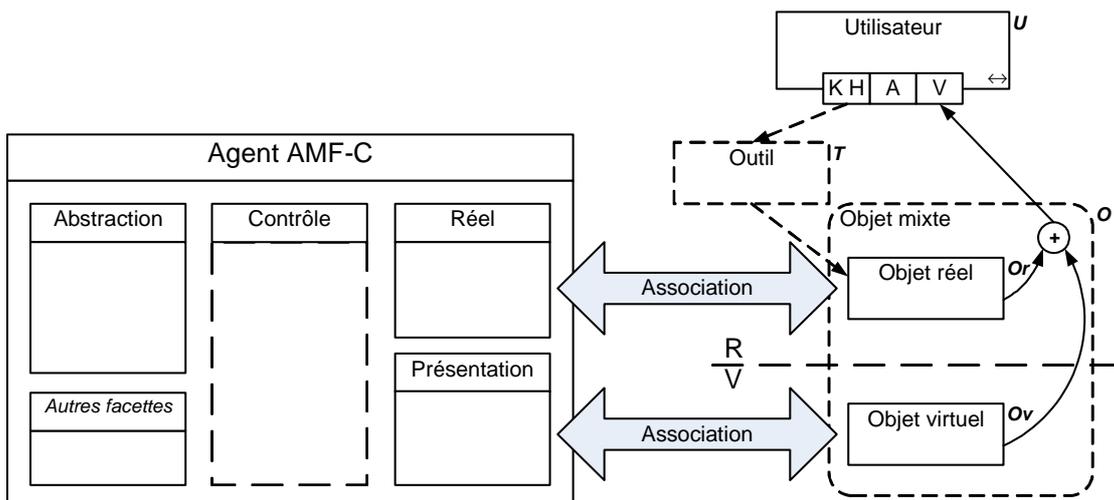


Figure 94 : Correspondance entre les agents AMF-C et les objets IRVO

Conformément à la Figure 93, on voit donc que les agents AMF-C apparaissent comme une modélisation fine de ce que nous avons appelé 'Modèle interne' dans IRVO. En effet, cette entité est responsable de la modélisation des objets (et des outils) au sein du système informatique (d'où le

nom de ‘modèle interne’) mais elle est aussi responsable de la dynamique de l’application et ceci est clairement assuré par AMF-C qui modélise et implémente le contrôleur de dialogue.

Finalement, les objets réels non augmentés peuvent ainsi être modélisés par le système par des agents AMF qui comportent une facette « Réel » mais pas de facette « Présentation ».

Pour que la facette « Abstraction » puisse refléter l’état actuel de l’objet réel il faut bien sûr que des « senseurs » *ad hoc* permettent d’aller chercher cette information dans le monde réel et la numériser : la facette « Réel » **apparaît ainsi comme la passerelle permettant de faire cette correspondance entre le monde réel et le monde virtuel en utilisant les services des senseurs.**

Bien sûr, l’inverse doit aussi être possible. Si la facette « Abstraction » de l’agent a été modifiée (par un autre agent par exemple), le nouvel état doit pouvoir être reflété sur l’objet réel que l’agent AMF modélise. Ce sont dans ce cas les « effecteurs » qui permettent de transformer cette information numérique en une action dans le monde réel et de nouveau la facette « Réel » fait office de passerelle entre les deux mondes en utilisant les services des effecteurs.

Pour les objets mixtes, comportant une partie réelle et une partie virtuelle, l’agent AMF permet donc bien d’assurer la cohérence entre les deux représentations grâce aux facettes « Réel » et « Présentation ». Dans les cas les plus simples, une communication directe (via la facette « Contrôle ») entre ces 2 facettes permet d’assurer la cohérence des deux parties. Dans d’autres cas, le noyau fonctionnel doit être informé du nouvel état et dans ce cas la communication entre les deux facettes se fait indirectement par la facette « Abstraction ».

### VII.3.3.3 Correspondances entre agents AMF-C et outils IRVO

A l’heure actuelle AMF ne considère pas les périphériques d’entrées/sorties explicitement ni la notion d’outil. Cela s’explique par le fait qu’il se positionne au niveau du contrôleur de dialogue et ne considère l’interface qu’au niveau abstrait. Les facettes « Présentation » des agents AMF reçoivent des événements provenant de l’interface concrète (représentés sous forme d’éclairs) qui déclenchent des démons associés aux ports. Une première approche est donc de considérer que les outils modélisés par IRVO doivent se comporter de la même manière en générant des événements à destination des agents AMF, même si ces événements sont « plus riches » que ceux générés par une souris ou le clavier. Cependant, nous pensons que pour les outils complexes, et plus particulièrement les outils mixtes, cette approche conduirait à un manque dans la formalisation de la conception. C’est la raison pour laquelle nous proposons **de modéliser les outils IRVO sous forme d’agents AMF** au même titre que les objets du domaine. De la même manière que pour les objets du domaine, la facette « Présentation » est associée à la partie virtuelle de l’outil (Tv) et la facette « Réel » est associée à la partie réelle de l’outil (Tr).

Malgré tout, dans les cas d’outils simples, et en particulier pour les outils virtuels, il pourrait être fastidieux de modéliser systématiquement les outils sous forme d’agents AMF. Dans ce cas, il est tout à fait admissible de ne pas le faire, ce qui correspond à une modélisation AMF classique où les outils n’apparaissent pas explicitement.

Modéliser les outils avec AMF-C a cependant des conséquences sur la manière même de modéliser une application. En effet, les facettes « Présentation » des objets ne reçoivent plus directement des messages provenant des actions de l’utilisateur avec la souris mais les actions des outils sont modélisées comme une communication entre l’agent AMF représentant l’outil et l’agent représentant l’objet de la tâche, cette communication se faisant par le mécanisme d’export des ports d’AMF. L’exemple présenté ci-après montre clairement ce mécanisme de communication inter-agents.

Ceci ouvre un certain nombre de questions qui devront être examinées mais qui sortent du cadre de ce mémoire : Peut-on mélanger les deux modes dans une même modélisation (certaines actions

sont déclenchées via des ports de la facette « Présentation » et d'autres le sont par le mécanisme de communication inter-agents) ? Cela ne pourrait-il pas conduire à des incohérences dans le modèle ? Plus fondamentalement, est-ce que cette introduction des outils ne distord pas trop le modèle AMF-C lui-même ?

### VII.3.3.4 Exemple de couplage de modèles IRVO et AMF-C

A titre d'exemple nous reprenons la modélisation de CASPER-v2+ (version améliorée de CASPER-v2) présentée dans les exemples du Chapitre VI, § VI.3.1.1. Dans la Figure 72 (reprise à droite de la Figure 95) nous voyons que nous avons deux objets augmentés :

- Un outil augmenté, composé de l'aiguille de ponction réelle (qui est cachée aux yeux du chirurgien car dans le corps du patient) et de son support et d'une aiguille virtuelle visualisée grâce au casque de visualisation semi transparent.
- Un objet du domaine de la tâche augmenté, composé du liquide péri-cardien à ponctionner (et qui est non visible) et d'une trajectoire virtuelle idéale représentée sous forme d'un cône de sécurité.

D'après ce que nous avons exposé ci-dessus dans la modélisation AMF-C nous avons donc deux agents (Figure 95 à gauche) :

- Un agent « Support-Aiguille », dont la facette « Réel » est associée au support (et l'aiguille réelle) et la facette « Présentation » à l'aiguille virtuelle,
- Un agent « Ponction », dont la facette « Réel » est associée au liquide péri-cardien à ponctionner et la facette « Présentation » à la trajectoire idéale virtuelle.

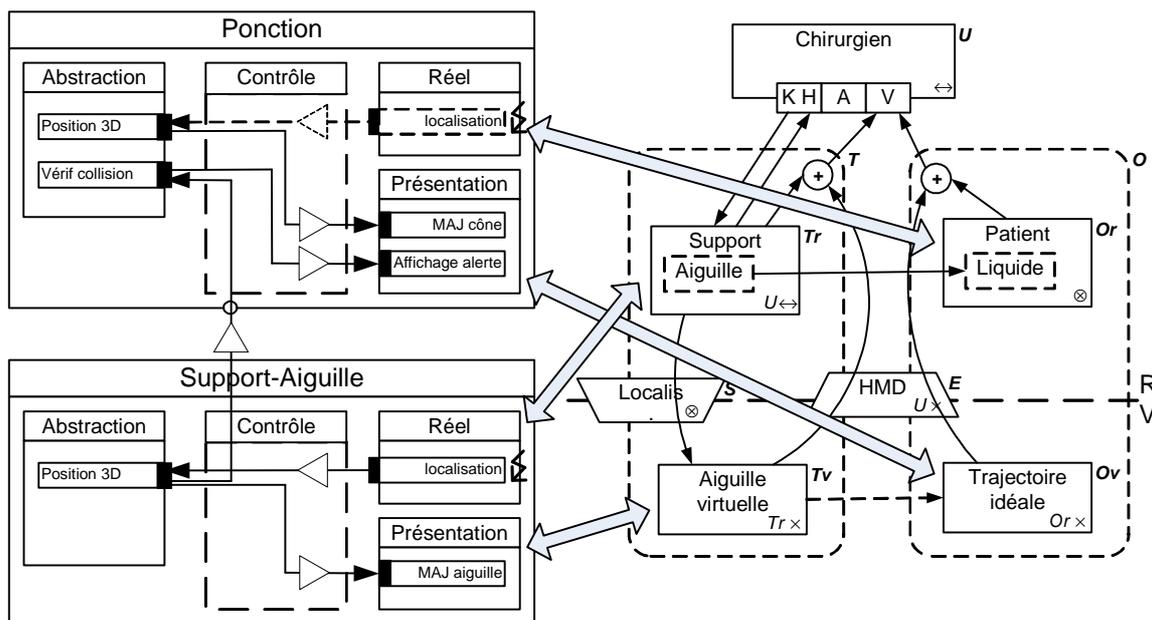


Figure 95 : Exemple de correspondance agent AMF / artefact IRVO

On remarquera que la localisation spatiale (position et orientation) de l'objet réel « Support » de IRVO se traduit par un événement externe (symbole éclair) sur un port « localisation » de la facette « Réel » de l'agent « Support-Aiguille » dans AMF-C. Ceci permet de mettre à jour la position de la seringue dans le référentiel de la salle d'opération et qui est gérée par la facette « Abstraction » de l'agent. Cette position permet, d'une part, de mettre à jour la facette « Présentation » (associée à l'objet virtuel « Aiguille virtuelle » de IRVO) pour afficher l'aiguille virtuelle dans la bonne position et, d'autre part, de vérifier que la position calculée ne rentre pas en collision avec le cône de sécurité qui est géré par l'agent « Ponction ». Ceci est mis en oeuvre en utilisant le mécanisme d'exportation de port de AMF-C (représenté par un petit cercle sur la bordure de l'agent « Ponction »).

Pour l'agent « Ponction », la position du cône dans le référentiel de la salle d'opération est fixe car le patient est supposé immobile (!) sur la table d'opération. Cette position doit simplement être initialisée avant l'opération dans une phase de calibration, soit par une saisie manuelle au clavier (non représentée), soit en utilisant le localisateur ce qui est représenté par l'activation du port « localisation » de la facette « Réel » de l'agent « Ponction ».

Par ailleurs, il est important de remarquer que les facettes « Présentation » des deux agents sont supposées gérer un affichage 3D de l'aiguille et de la trajectoire. La projection 3D-2D sur l'écran du casque de visualisation semi-transparent du chirurgien est supposée être faite au niveau de la présentation concrète et n'est donc pas représentée sur le modèle AMF-C. En effet, pour mener à bien cette opération, le système doit localiser la position du casque de visualisation dans le référentiel de la salle d'opération (en utilisant à nouveau le dispositif de localisation) et utiliser cette information pour projeter la scène 3D sur la surface 2D de l'écran en temps réel. Cette opération de bas niveau est aisément prise en charge par toute « toolkit » graphique et n'est donc pas du niveau d'abstraction traité par le modèle AMF-C. D'un point de vue de la modélisation IRVO, on peut considérer que cette projection 3D-2D est une fonction d'adaptation réalisée par le transducteur avant la transformation de l'information numérique en signal physique.

### VII.3.3.5 Intérêt de ce couplage AMF-C / IRVO

Ce rapprochement entre ces deux modèles a plusieurs intérêts. Tout d'abord, sur le plan théorique, il montre la continuité entre les deux modèles :

- IRVO se positionne comme une modélisation de l'interaction du point de vue de l'utilisateur (en le prenant explicitement en compte) et se positionne clairement du côté de l'interface à un niveau semi-concret,
- AMF-C est un modèle du dialogue permettant de formaliser le contrôleur de dialogue qui est la clef de voûte du modèle Arch et représente l'explicitation de l'entité 'M' du modèle IRVO.

Ainsi, toute la partie centrale et droite du modèle Arch est couverte par ces deux modèles, chacun selon une approche différente mais complémentaire.

Sur un plan plus pragmatique, le fait d'avoir montré la continuité entre IRVO et AMF-C permet de se reposer sur tous les travaux existants (et à venir) autour du modèle AMF-C liés au travail collaboratif (Cf. § V.3.3) :

- Agents fragmentés et répliqués,
- Administrateurs *ad hoc* pour les facettes « Contrôle » qui permettent la sélection/libération d'objet et la validation des actions des utilisateurs.

On peut passer ainsi d'un niveau de conception de l'interface avec IRVO à un modèle complet et précis du dialogue avec AMF-C qui ensuite peut gérer le dialogue dans l'application elle-même dans une approche MDA (Cf. § IV.2.3.2 sur le « moteur AMF »).

### VII.3.3.6 Bilan sur les modifications de l'architecture du logiciel

Nous avons vu que la prise en compte de la Réalité Mixte avait une incidence sur l'architecture du logiciel interactif. Nous nous sommes limités dans cette partie aux conséquences sur le modèle CUO de l'architecture des applications collaboratives dans la méthodologie CoCSys. Nous avons donc examiné les liens entre le modèle IRVO (qui est inséré dans le modèle comportemental) et le modèle AMF-C qui tient le rôle de modèle CUO dans l'architecture.

Nous proposons de modéliser les objets du domaine mixtes d'IRVO (c'est-à-dire des objets composés d'une partie réelle et d'une partie virtuelle) sous la forme d'un agent AMF-C :

- A la partie virtuelle de l'objet (Ov) nous associons la facette « Présentation » classique de l'agent AMF,
- A la partie réelle de l'objet (Or) nous associons la nouvelle facette « Réel » de l'agent AMF.

Pour les objets du domaine, il est bien nécessaire de créer un agent AMF-C puisque la facette « abstraction » est en liaison avec le Noyau Fonctionnel de l'application qui gère les objets du domaine. Les outils ne sont initialement pas modélisés par AMF-C et il doit toujours être possible de construire l'architecture du logiciel de cette manière. Nous pensons cependant que, pour certains systèmes, il devient nécessaire de modéliser aussi les outils avec un agent AMF-C. Dans ce cas, il s'agit d'un agent intermédiaire dont la facette « abstraction » n'est pas connectée au Noyau Fonctionnel de l'application.

Cette proposition doit être considérée comme un travail préliminaire qui demande de nombreuses vérifications, tant sur la validité théorique vis-à-vis du modèle AMF-C que d'un point de vue pragmatique sur l'intérêt d'un tel ajout dans le modèle.

### VII.3.4 Bilan sur la méthodologie de conception

Nous avons, dans cette section, décrit la méthodologie CoCSys destinée à la conception d'applications collaboratives et mobiles qui est en cours de développement au laboratoire ICTT. Nous avons ensuite examiné comment le modèle IRVO s'insère dans cette méthodologie au niveau du modèle comportemental et quelles étaient les conséquences sur la méthodologie tant au niveau de l'élaboration des scénarios que de la construction du modèle comportemental de l'application. Nous avons plus particulièrement étudié les apports des modélisations IRVO dans ces phases de conception d'applications de systèmes mixtes.

Enfin, nous avons examiné comment, dans la phase de développement de l'application, le modèle IRVO permet d'assister la construction de l'architecture du logiciel en établissant des liens avec le modèle d'architecture AMF-C. Il s'agit ici de travaux préliminaires qui demandent encore à être affinés et surtout confirmés.

## VII.4 Comparaison avec d'autres méthodologies de conception pour les systèmes mixtes

Nous examinons dans cette section, deux autres méthodologies de conception des systèmes mixtes et les comparons avec notre proposition.

### VII.4.1 Approche de Trevisan et al.

Daniela Trevisan *et al.* [Trevisan 2003] proposent une approche basée sur les modèles pour la conception des systèmes de Réalité Mixte. Les auteurs proposent de prendre en considération 6 modèles :

- Un *modèle de l'utilisateur*, classique. Il est cependant étendu par des aspects spécifiques aux systèmes de RM comme la localisation spatiale de l'utilisateur pour les applications qui en ont besoin (par exemple pour actualiser l'affichage d'un casque de visualisation)
- Un *modèle de la tâche*, classique lui aussi.
- Un *modèle du domaine*, qui décrit les objets qu'un utilisateur peut manipuler. Il s'agit soit d'objets virtuels manipulés à travers l'interface (et représentés par le modèle de présentation), soit d'objets réels manipulés directement par l'utilisateur.
- Un *modèle de présentation*, qui au-delà de la structuration classique en unités de présentation, doit considérer l'intégration physique et spatiale des informations affichées avec les objets réels. L'intégration physique décrit comment les différents médias visuels sont distribués sur les périphériques de sortie (affichés sur un seul écran ou sur plusieurs écrans). L'intégration spatiale concerne l'arrangement spatial et les caractéristiques topologiques des objets. En particulier il s'agit d'aligner visuellement les objets virtuels et les objets physiques dans le cas des objets augmentés.

- Un *modèle de dialogue*, classique, qui décrit les aspects dynamiques de l'interface. Il définit comment les utilisateurs interagissent avec les objets de présentation (boutons, menus, etc.), avec les médias d'interaction (entrée vocale, écran tactile, etc.) et les réactions que l'interface communique via ces objets.
- Un *modèle d'application et de plateforme*, qui décrit les aspects système et matériel des dispositifs. En particulier, au niveau système, il peut être nécessaire de synchroniser des événements de différents périphériques selon la tâche accomplie et aussi d'intégrer différentes sources de médias dans une ou plusieurs unités de présentation en mixant les informations réelles et virtuelles correctement. Pour la plateforme, au-delà des informations classiques, il faut bien sûr définir les périphériques spécifiques utilisés : gants de données, casques de visualisation, localisateurs, etc.

La structuration du modèle de présentation est développée dans [Trevisan 2004] où les auteurs considèrent *l'Espace d'interaction mixte* qui est défini comme étant l'environnement complet d'interaction requis pour mener à bien une tâche interactive particulière. Il comprend 4 aspects :

- *L'intégration spatiale des objets mixtes* ; il s'agit des relations topologiques entre les objets (disjoints, recouvrements, juxtaposition, etc.), de la direction de ces relations (gauche, droit, au-dessus, au-dessous à gauche, etc.) et des distances entre les objets.
- *L'intégration temporelle* ; il s'agit de synchroniser temporellement les différents médias entre eux. Les relations sont basées sur les relations temporelles d'Allen [Allen 1983].
- *Le contexte d'insertion des périphériques* en considérant 4 niveaux : la zone centrale (0-45 cm de l'utilisateur), la zone personnelle (0,46-1,2 m de l'utilisateur), la zone sociale (1,3-3,6 m de l'utilisateur), la zone publique (plus de 3,6 m de l'utilisateur).
- *Le focus de l'interaction*. 5 cas sont distingués : focus sur un seul objet du monde réel, focus sur un seul objet du monde virtuel, focus sur plusieurs objets du monde réel, focus sur plusieurs objets du monde virtuel, focus sur des objets des mondes réels et virtuels.

#### Comparaison avec notre méthodologie

Dans l'état actuel de la proposition, les deux approches sont similaires puisqu'elles sont toutes les deux basées sur des approches MBA même si le choix des modèles diffère légèrement. Dans la méthodologie CoCSys, la construction des différents modèles a été particulièrement étudiée en se basant sur une approche basée sur des scénarios, alors que dans la proposition de Trevisan *et al.* rien n'est précisé sur la manière de procéder à cette construction. CoCSys va également plus loin du côté du développement du logiciel en formalisant le processus de contextualisation, adaptation, et spécialisation.

### VII.4.2 Approche de Renevier

Philippe Renevier [Renevier 2004c] propose une démarche de conception des systèmes mixtes collaboratifs et mobiles basée sur des scénarios :

- Une *phase d'analyse de la tâche réelle* : dans cette phase des scénarios sont élaborés par une explicitation hors site par les futurs utilisateurs des étapes de leurs activités qui leur semblent pertinentes.
- Une *phase d'analyse de l'activité* : elle est réalisée par une observation instrumentée de l'activité des futurs utilisateurs commentée par eux-mêmes sur site.
- Une *phase d'identification des besoins* réalisée à partir des scénarios obtenus aux phases précédentes.
- Une *phase de spécifications externes* de l'application : Elle s'appuie sur l'élaboration de scénarios projetés qui décrivent l'activité future avec le système à développer. Cette phase est itérative pour affiner les spécifications externes.
- Une *phase de développement* du logiciel.

- Une *phase d'évaluation expérimentale* : les outils sont testés auprès des futurs utilisateurs tant sur les aspects des fonctionnalités (utilité) que sur l'interface (utilisabilité).

Dans cette démarche les scénarios projetés constituent un outil central de la conception de l'interaction. Ils servent de support à une évaluation prédictive du futur système interactif. L'auteur distingue deux niveaux d'abstraction pour les scénarios projetés :

- Les *scénarios projetés abstraits* : ils décrivent les fonctions de l'interface du système interactif sans les détails des modalités d'interaction employées. Seuls les objets, aussi bien physiques que numériques, indispensables à l'activité sont présents dans ces scénarios. L'auteur propose, plutôt que d'utiliser des scénarios narratifs, de les décrire avec un formalisme graphique qu'il propose et que nous avons présenté au paragraphe V.2.1.
- Les *scénarios projetés concrets* : ils décrivent l'interaction telle qu'elle va se dérouler. Il s'agit de la réalisation des scénarios abstraits en choisissant les modalités d'interaction. L'auteur propose de décrire ces scénarios concrets en utilisant la notation ASUR d'Emmanuel Dubois (Cf. § IV.1.7.1).

### Comparaison avec notre méthodologie

Cette méthodologie présente de nombreuses similitudes avec celle que nous proposons.

Les scénarios projetés abstraits et concrets sont présents dans la méthodologie CoCSys sous forme de scénarios. Ils permettent de manière itérative de définir les spécifications externes qui sont représentées dans le modèle CAB dans notre démarche. Par contre le modèle CAB va plus loin puisqu'il formalise complètement les spécifications sous forme de modèles qui permettront de produire le logiciel suivant une approche MDA.

Nous avons vu (Cf. § VI.3.3.1) que le modèle IRVO est fonctionnellement équivalent au modèle ASUR et permet donc la représentation des scénarios projetés concrets. Mais IRVO permet aussi dans une certaine mesure la représentation des scénarios projetés abstraits comme nous l'avons montré dans ce chapitre au paragraphe VII.3.2.1 et par la comparaison avec le formalisme de Renevier au paragraphe VI.3.3.2. Cependant, IRVO n'ayant pas été conçu dans cette optique il est forcément moins souple et moins expressif que le formalisme *ad hoc* de Renevier.

Finalement, IRVO permet le passage des scénarios abstraits aux scénarios concrets sans changer de formalisme, mais en affinant et en complétant des modélisations au fur et à mesure de la conception alors que dans la proposition de Renevier il y a une conversion à réaliser entre son formalisme et ASUR ce qui conduit nécessairement à une rupture dans le processus de conception.

## VII.5 Bilan du chapitre

Dans ce chapitre nous avons examiné en détail l'articulation entre le modèle IRVO et l'activité de conception des systèmes de Réalité Mixte. En partant des approches MBA nous avons montré que IRVO s'inscrit pleinement dans ces démarches sans les remettre en cause mais en les complétant.

Nous avons montré en particulier l'articulation entre IRVO et l'arbre de tâches : chaque schéma IRVO est associé à une tâche élémentaire ou de plus haut niveau dans l'arbre. L'arbre de tâches apparaît ainsi comme le lien entre les différents schémas IRVO d'un système et en particulier exprime les relations temporelles qui existent entre ces schémas. En remontant l'arbre de tâches il est également possible de bâtir un schéma de synthèse de l'application permettant de détecter d'éventuels problèmes de discontinuités entre tâches.

Nous avons ensuite proposé une première ébauche d'une méthodologie de conception de systèmes mixtes et collaboratifs en nous basant sur une méthodologie existante, CoCSys, basée sur une approche MBA. Nous avons plus particulièrement étudié l'insertion du modèle dans le modèle comportemental et avons montré l'apport d'IRVO à deux niveaux :

- Comme point de départ de la conception en formalisant des cas typiques d'interaction avec le futur dispositif. De ce point de vue il peut être vu comme un formalisme graphique complétant des scénarios textuels.
- Comme modèle résultant de l'analyse : après décomposition de l'application en arbre de tâches, pour chaque tâche utilisant la réalité mixte, un schéma IRVO associé est construit pour modéliser finement l'interaction de l'utilisateur avec les outils et l'objet de la tâche.

Nous avons ensuite examiné quelles étaient les conséquences sur la méthodologie tant au niveau de l'élaboration des scénarios que de la construction du modèle comportemental de l'application. Nous avons plus particulièrement étudié les apports des modélisations IRVO dans ces phases de conception d'applications de systèmes mixtes. Cette conception pourrait s'appuyer sur un ensemble de patterns d'interaction qui peuvent être facilement établis à partir des exemples de la littérature que nous avons modélisés.

Enfin, nous avons examiné comment, dans la phase de développement de l'application, le modèle IRVO permet d'aider la construction de l'architecture du logiciel en établissant des liens avec le modèle d'architecture AMF-C.

La méthodologie que nous proposons ici est encore dans sa première version et de plus amples travaux sont encore nécessaires pour affiner cette proposition et la rendre pleinement opérationnelle.



# Chapitre VIII    **Projet ACCORéA — Aide à la Conception Coopérative supportée par la Réalité Augmentée — et prototypes associés**

*They did not know it was impossible, so they did it!*  
Mark Twain (1835 – 1910).

Dans ce chapitre nous abordons les développements que nous avons réalisés pour mettre en pratique nos travaux conceptuels exposés dans les chapitres précédents. L'essentiel de ces travaux s'inscrit dans le projet ACCORéA commencé en 2001 et des sous-projets associés comme TableGate (1 et 2) que nous allons présenter plus en détail.

Dans une première section nous décrivons le projet ACCoRéA et ses objectifs tels qu'ils étaient posés initialement et la démarche adoptée. Ensuite nous présentons les résultats de travaux préliminaires sur la rédaction collaborative et l'écriture d'articles scientifiques.

Dans la section VIII.4, nous exposons le projet Tablegate qui nous a permis de développer un prototype d'éditeur collaboratif utilisant une table interactive mettant en œuvre une partie de l'activité de la rédaction collaborative. La conception de ce logiciel a été l'occasion de tester partiellement la méthodologie de développement que nous avons proposée au chapitre précédent ainsi que le modèle IRVO.

Enfin, nous présentons succinctement le projet en cours Tablegate 2 avant de faire le bilan du chapitre.

## **VIII.1 Description du projet ACCORéA**

Le projet ACCORéA (Aide à la Conception Coopérative supportée par la Réalité Augmentée) cherche à étudier comment la Réalité Mixte peut aider les individus dans des **activités de conception** tant dans ses **dimensions individuelles que collectives** et que les acteurs soient dans un **même lieu ou à distance**. L'approche de ce travail vise à la création d'un continuum aussi fluide que possible entre la réalité et la virtualité tant au niveau de l'objet en cours de conception que des outils permettant d'agir sur ces objets. L'objectif de nos travaux n'est pas seulement un dispositif technique (un ensemble de logiciels animant une plate-forme matérielle), mais également une organisation du travail, dans ses dimensions individuelles, collectives et temporelles et de façon plus précise l'identification des activités, de tâches, d'actions, d'outils et de représentations les plus appropriés pour un travail intellectuel de conception.

Ce projet a été initié par l'auteur de ce mémoire au printemps 2001 pour une durée de 3 ans et a obtenu un financement de 50 000 F sous forme de BQR (Bonus Qualité Recherche) de l'Ecole Centrale de Lyon. L'équipe du projet était composé d'une équipe pluridisciplinaire de 5 chercheurs permanents du laboratoire ICTT (Interaction Collaborative, Téléactivités, Téléformation) de l'Ecole Centrale de Lyon : 3 chercheurs en informatique et 2 en psychologie. Ont également participé à ce projet, une étudiante du DEA de « communication homme/machine et ingénierie éducative » de

l'Université du Maine, Magali Beldame, dans le cadre de son stage de DEA [Beldame 2001], ainsi que 5 étudiants de la promo 2005 de l'École Centrale de Lyon dans le cadre d'un projet d'Etude d'une durée de 15 mois (octobre 2002 - janvier 2004), TableGate [Bianchi 2004]. Une nouvelle équipe de 5 étudiants de la promo 2006 a pris le relais en octobre 2003 sous le nom de TableGate II.

Dans le projet initial trois contextes applicatifs étaient envisagés :

- La rédaction collaborative de documents scientifiques (articles de conférences ou de revues, rapports)
- La conception de sites web
- La conception de produits industriels

Par rapport aux objectifs initiaux le projet s'est principalement focalisé sur les activités de conception et de rédaction de documents scientifiques et ce sont ces démarches et résultats que nous présentons dans cette section.

## VIII.2 Démarche

Dans le cadre de la rédaction d'articles scientifiques, une part importante de cette activité est aujourd'hui informatisée. Cette activité intellectuelle de conception (individuelle ou collective) nécessite la manipulation répétée d'un grand volume d'information et les écrans d'ordinateurs sont encore aujourd'hui, à quelques exceptions près, trop petits pour permettre de voir d'un seul coup d'œil cette masse d'information et de la manipuler, ordonner, extraire,... Régulièrement, le document est imprimé pour permettre sa relecture, sa ré-organisation, ou bien encore pour y travailler avec des co-auteurs, etc. Mais, une fois ce travail effectué sur le papier, il faut réintroduire sur la version électronique de manière manuelle toutes les modifications ce qui est consommateur de temps et source d'erreurs. Ainsi, il paraît intéressant de mettre au point un outil permettant de basculer alternativement et de manière souple du réel au virtuel tant au niveau de l'objet de l'action que des outils permettant d'agir.

L'idée est donc de mettre en place un environnement de travail permettant de manipuler des documents réels et virtuels avec des outils réels et virtuels. L'objet de la tâche est donc ici un texte qui est un concept abstrait pouvant être représenté sous différentes formes, document physique ou document numérique ; L'objet peut donc changer de forme à plusieurs reprises lors des activités menées par les différents utilisateurs.

Le point de départ peut être un ensemble de documents réels qui peuvent dans un premier temps être accrochés sur un mur réel observable par un moyen électronique supportant des opérations de désignation, sélection, déplacement, extraction (...) et de constitution d'un support électronique du document en résultant. Un autre point de départ peut être le travail en commun autour d'une table interactive, permettant d'élaborer ensemble le plan du document. Pour la suite de l'activité de conception (individuelle ou collective) il est très souvent nécessaire de re-matérialiser le document. Cette re-matérialisation conduit à replacer sur le mur ou sur la table la nouvelle version papier et pouvoir ainsi la retravailler soit de façon physique (avec les outils physiques) soit de façon virtuelle avec des outils informatiques.

Le travail de conception se définit comme un ensemble de tâches la caractérisant. Ces tâches doivent être identifiées et décomposées à leur tour en tâches élémentaires, activités, actions, etc., afin d'étudier comment les réaliser et les instrumenter dans le contexte de la Réalité Mixte, c'est-à-dire dans le continuum réalité - virtualité à double sens. Leur instrumentation pourrait se réaliser grâce à plusieurs outils (table interactive, C-pen, eye-tracking, object-tracking,...) en tentant de définir leurs implications possibles dans le contexte de réalité augmentée.

Dans le projet ACCORéA nous distinguons plus particulièrement :

- Des objets virtuels : des documents informatisés, organisés sous forme hiérarchique (chapitres, paragraphes, phrases, ...) et visualisables sous différentes formes (continu, par pages, en mode plan, ...)
- Des objets réels : des feuilles sur lesquelles le document est imprimé et pouvant être découpées, assemblées, annotées, ...
- Des outils virtuels : les outils classiques du traitement de texte (écriture, couper/copier/coller, glisser/déposer, ...)
- Des outils réels : les stylos, des surligneurs, la main qui déplace et classe des feuilles, la paire de ciseaux, le tube de colle, ...

L'objectif est d'imaginer, de mettre en œuvre et de valider des tâches, activités et actions et de pouvoir non seulement utiliser les outils virtuels sur les objets virtuels et les outils réels sur les objets réels mais de pouvoir également utiliser les outils virtuels sur des objets réels ou des outils réels sur des documents virtuels. Par exemple, nous voudrions pouvoir faire par désignation directe sur la feuille de papier des opérations de copier/coller qui modifieraient directement le document virtuel comme sur le DigitalDesk [Wellner 1993b].

### VIII.3 Travaux préliminaires

Nous allons présenter dans un premier temps les travaux conceptuels qui ont été effectués sur la rédaction collaborative et le processus d'écriture d'un article scientifique par Magali Beldame et qui ont servi de base au projet Tablegate présenté par la suite.

#### VIII.3.1 Rédaction collaborative

Magali Beldame a étudié en détail le processus de conception [Beldame 2001]. Elle a plus particulièrement étudié l'activité de conception d'articles scientifiques et notamment la rédaction collaborative d'articles et nous reprenons ici les grandes lignes de ses travaux.

Elle définit l'écriture « comme un acte graphique permettant à un individu de faire part à un autre individu de manière différée dans le temps et dans l'espace, des événements, des ou ses actions, ses pensées, ses sentiments, ce qu'il a pu vivre ou observer à un moment donné ». L'écriture est au service de la pensée : elle est une trace qui permet de disposer ce dont a besoin la pensée pour poser les éléments de sa réflexion.

La psychologie cognitive a étudié les mécanismes et les stratégies impliqués dans la production écrite par un individu. Il existe deux principaux types de modèles pour décrire ces processus :

- Le **modèle récursif** qui s'inspire du système de traitement de l'information de Hayes et Flower (1980) : ce modèle est basé sur trois processus animés par un principe de récursivité constante :
  - La *planification cognitive* qui correspond à la récupération des informations stockées en mémoire à long terme de l'individu ainsi que la prise d'information en situation, la sélection et l'organisation des informations.
  - La *production*, processus par lequel l'individu fait les choix linguistiques qui traduisent le mieux les choix sémantiques effectués lors de la phase de planification et fait les actions de génération de texte.
  - La *révision* qui regroupe la relecture d'un texte déjà produit et les modifications.
- Le **modèle basé sur l'interaction sociale** exposé par Nystrand en 1989 qui s'inscrit dans une approche constructiviste. La production écrite apparaît comme une négociation entre le scripteur et le lecteur pour la signification du texte. De ce point de vue, la nature de l'écriture est nécessairement interactive et sociale. Ce modèle comporte trois opérations essentielles :
  - L'initialisation du discours écrit,
  - La maintenance et la régulation du discours écrit,

- L'élaboration du rédacteur.

Il existe différentes manières d'écrire de façon collaborative. Par exemple, Ronald Baecker et al. ont proposé la classification suivante [Baecker 1993] :

- La stratégie de l'écriture unique (*single writer*) : un individu écrit un texte en se basant sur les discussions avec les autres membres du groupe.
- La stratégie du scribe (*scribe*) : les individus travaillent ensemble et l'un deux retranscrit les pensées et décisions du groupe.
- La stratégie de l'écriture séparée (*separate writers*) : un document est divisé en parties et chaque membre du groupe est chargé de l'écriture d'une partie.
- La stratégie de l'écriture jointe (*joint writing*) : plusieurs membres du groupe rédigent le texte ensemble et le contenu est négocié par l'ensemble des membres du groupe.

### VIII.3.2 Modélisation – arbre de tâches

Magali Beldame a modélisé la tâche de rédaction d'article scientifique avec le formalisme MAD (§ VII.2.2). Nous reproduisons ce modèle à la Figure 96.

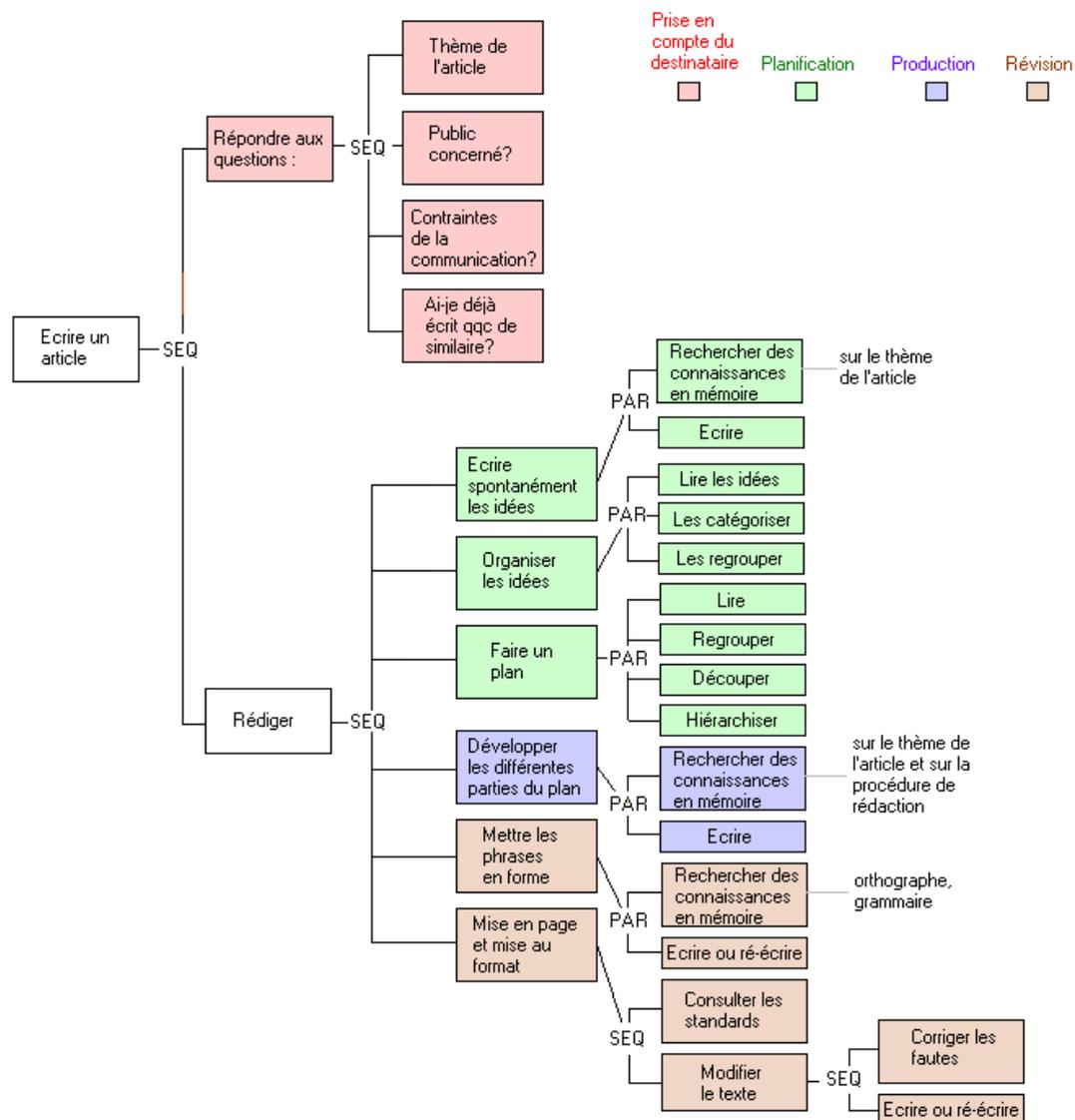


Figure 96 : Modélisation de la tâche de rédaction d'article scientifique (d'après [Beldame 2001]).

Nous rappelons la signification des différents constructeurs utilisés par MAD :

- SEQ : signifie que les sous-tâches sont séquentielles,
- SIM : signifie que les sous-tâches sont simultanées,
- ALT : signifie que les sous-tâches sont des alternatives (exclusivement)
- PAR: signifie que les sous-tâches sont entrelacées.

Ces travaux préliminaires nous ont permis de formaliser l'activité de rédaction d'écriture d'articles scientifiques. Il s'agit maintenant de les mettre en œuvre de manière pratique dans un prototype. Parmi l'ensemble des tâches considérées, nous avons choisi de nous focaliser sur une tâche qui permette le travail collaboratif synchrone autour du même texte et utilisant la Réalité Mixte. La mise en œuvre de cette tâche a fait l'objet du prototype WordTab.

## VIII.4 Projet Tablegate et prototype Wordtab

A partir des études préliminaires présentées ci-dessus nous avons lancé un projet d'étude, TableGate, avec 5 étudiants de l'Ecole Centrale de Lyon de la promo 2005 d'une durée de 15 mois d'octobre 2002 à janvier 2004 [Bianchi 2004]. Le but de ce projet était de réaliser un prototype permettant l'édition collaborative d'un document en cours de rédaction en utilisant notre Table Interactive. Le principe est de pouvoir poser un document papier (une dizaine de feuilles A4 imprimées) et sur lequel plusieurs personnes peuvent travailler simultanément. Les utilisateurs interagissent sur les documents papier en sélectionnant, déplaçant, effaçant ou copiant des paragraphes du texte pour le réorganiser en fonction des objectifs de production.

### VIII.4.1 Objectifs de WordTab

Le but de WordTab est de couvrir une partie du processus d'écriture collaborative essentiellement dans la phase de révision en proposant aux auteurs un outil de réorganisation de la structure du document en modifiant l'ordre des paragraphes ou en supprimant des parties. Il permet une édition macroscopique sur l'ensemble du document, le plus petit grain de travail étant le paragraphe. Il ne permet donc pas l'édition fine du texte comme le propose tout traitement de texte : ajouter/supprimer des lettres, des mots, des phrases.

Nous partons du constat que, pour ce genre de tâche de réorganisation, il faut pouvoir avoir une vue globale de la structure du document, mais en même temps, pour décider si un paragraphe doit être déplacé à tel ou tel endroit, il faut être capable de pouvoir en visualiser aisément le contenu. Or, les écrans informatiques permettent tout au plus des affichages en 2000 x 1600 ce qui permet de visualiser 2 pages simultanément à l'écran. Si l'on souhaite en visualiser plus (par exemple une dizaine), le texte ne devient plus lisible car la résolution est insuffisante. On est donc conduit à faire sans cesse des zooms avant (pour lire le texte) et des zooms arrière pour visualiser la structure.

Le papier n'a pas cet inconvénient : il est facile et peu coûteux d'imprimer le document puis d'étaler les pages sur une table. La structure **et** le texte sont perceptibles simultanément et éventuellement par plusieurs personnes. Bien sûr, toutes les annotations manuscrites qui peuvent être faites doivent être ressaisies à la main par la suite sur l'ordinateur et il serait intéressant d'avoir un moyen de refaire la synchronisation plus facilement. C'est ce problème qui a donc motivé le développement de WordTab : utiliser le papier pour son confort de visualisation mais garder la synchronisation avec la version électronique pour ne pas avoir à refaire le travail.

De nombreux auteurs reconnaissent l'intérêt du papier à l'heure du « tout numérique ». Par exemple, Pierre Wellner considère que le papier et le stylo fournissent un moyen portable, tangible et ubiquitaire de capture de l'information [Wellner 1993b].

Stifelman dit que les documents papiers ont de nombreux avantages sur les documents numériques. Une feuille de papier peut facilement être détachée d'un carnet, fourrée dans une poche pour un accès facile ou donnée à la main à un ami. Les utilisateurs se rappellent souvent comment accéder à l'information désirée par l'emplacement physique de leurs notes [Stifelman 1996].

Pour Harper il existe même 2 raisons significatives à la persistance du papier sur les lieux de travail en particulier en tant que support au travail collaboratif [Harper 1995] :

- Le papier a des propriétés physiques qui font qu'il est particulièrement adapté pour supporter certains aspects importants du travail collaboratif. Ces propriétés particulières ne sont pas fournies de manières satisfaisantes par d'autres médias ou outils collaboratifs.
- Les pratiques organisationnelles de travail dans un grand nombre de domaines ont évolué de pair avec l'usage du papier. Le papier a aidé à affiner des pratiques de travail et les pratiques de travail ont été conçues autour de l'usage du papier.

## **VIII.4.2 Conception de WordTab**

Pour présenter la conception de WordTab, nous suivons la méthodologie CoCSys étendue telle que nous l'avons exposée au chapitre précédent.

### **VIII.4.2.1 Scénario d'utilisation**

La première phase est le recueil des scénarios. Dans ce cas simple, outre la présentation générale présentée au § VIII.3.2, nous nous focalisons sur l'étape de révision du document. Pour cette étape, nous avons un seul scénario qui décrit complètement l'usage de la future application.

Nous considérons une équipe de travail qui rédige un document de manière collaborative. Il peut s'agir par exemple de chercheurs qui rédigent un article scientifique pour une revue. On suppose que les rédacteurs ont déjà initié le travail par exemple en faisant un premier plan pour se répartir le travail et qu'ils ont construit une première version contenant toutes les contributions des auteurs dans un document. Le contenu est donc présent dans ses grandes lignes mais il s'agit maintenant de le réorganiser pour établir le plan définitif avant les dernières corrections et finitions. Les rédacteurs impriment le document et étalent les feuilles sur la table ce qui leur permet d'avoir une vue globale de la structure, tout en pouvant relire certaines parties du texte si nécessaire (par exemple pour prendre connaissance de parties qu'ils n'ont pas écrites). Les rédacteurs collaborent, supprimant des paragraphes, déplaçant d'autres paragraphes directement en pointant sur le texte les parties qu'ils veulent déplacer. Après chaque action, le vidéoprojecteur affiche directement sur les feuilles des zones de couleurs et des flèches qui sont la trace des actions. Ils peuvent revenir en arrière en annulant des modifications ou en changeant la destination de certains paragraphes. Lorsqu'ils sont satisfaits du résultat, ou lorsque les modifications sont trop nombreuses et que cela devient illisible, alors ils sauvegardent leur travail ce qui a pour effet de générer une nouvelle version du document contenant toutes leurs modifications. Ce nouveau document peut de nouveau être imprimé et le processus reprendre comme ci-dessus. Après cette réorganisation, les rédacteurs utilisent ensuite leur traitement de texte habituel pour corriger les fautes et rajouter des phrases ou des paragraphes de transition. Ils peuvent, si nécessaire, décider de refaire une séance sur la table pour affiner encore la structure jusqu'à ce qu'ils convergent vers une version définitive.

### **VIII.4.2.2 Construction du modèle comportemental**

A partir du scénario présenté ci-dessus nous pouvons construire les différents modèles qui composent le modèle comportemental.

#### *Acteurs*

Il n'y a ici qu'un seul type d'acteur qui a soit le rôle de rédacteur du document, soit le rôle de lecteur du document. Un acteur peut changer de rôle à tout moment, la coordination étant assurée directement par les acteurs par un protocole social.

### Processus

Il n'y a pas de processus de type workflow dans ce système car les rédacteurs ont un rôle symétrique et la coordination est assurée par la communication directe des acteurs.

### Artefacts

L'objet du domaine est un document informatique. Il est représenté sous forme de feuilles de papiers imprimées qui sont les pages qui composent le document.

### Contexte

Le dispositif d'interaction est une table interactive (qui permet de capturer les actions des utilisateurs) complété par un vidéoprojecteur pour réaliser les augmentations.

### Arbre de tâches

L'arbre de tâches représenté avec le formalisme CTT résume les différentes actions possibles avec le logiciel (Figure 97).

Dans la tâche principale de travail (tâche « Travailler »), une première sous tâche, est de placer les feuilles sur la table et de calibrer le dispositif (tâche « Placer feuilles »). La deuxième sous tâche, qui peut être répétée autant de fois que nécessaire, est la tâche « Editer » : il faut sélectionner un paragraphe puis choisir une action parmi la copie, le déplacement et la suppression de ce paragraphe. On peut à tout moment imprimer la nouvelle version du document ce qui fait recommencer une nouvelle séquence de travail : « Placer les feuilles » et « Editer ».

Le travail peut être interrompu par la tâche « Sauver et quitter » qui enregistre la nouvelle version du document et sort du logiciel.

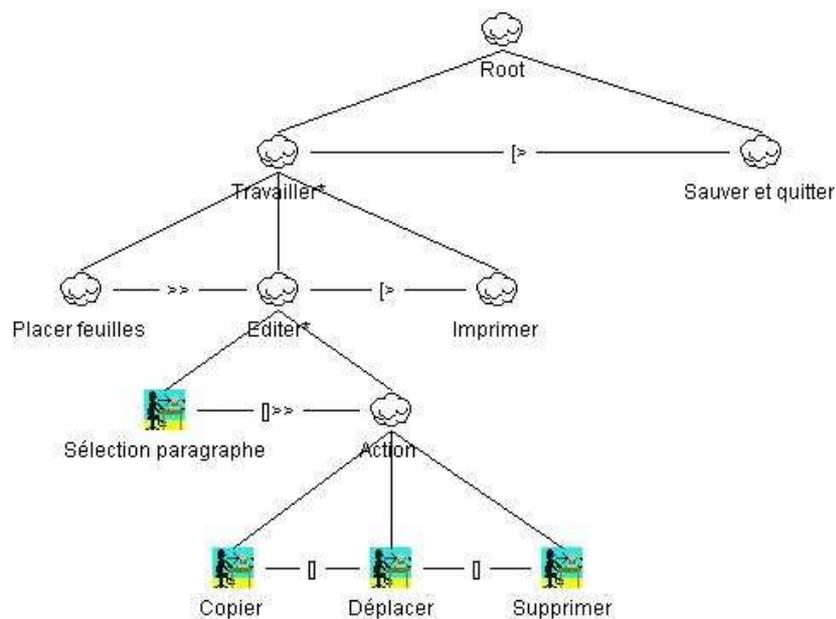


Figure 97 : Modélisation des tâches principales de WordTab avec CTT.

### Modèle IRVO

Pour la tâche principale d'édition (tâche « Editer »), nous donnons le schéma IRVO correspondant (Figure 98). C'est la seule tâche qui utilise la Réalité Mixte, les autres étant des tâches non augmentées. Nous considérons qu'il n'y a qu'un seul rédacteur et plusieurs lecteurs à un instant

donné. Cependant, les rôles peuvent tourner entre les différents acteurs à chaque édition et même à chaque interaction : on peut imaginer qu'une personne sélectionne le paragraphe et qu'une autre choisisse le point de déplacement (et ce type de comportement a effectivement été observé dans les expérimentations, Cf. § VIII.4.4).

Le rédacteur utilise un stylet (voire directement son doigt) qui a valeur d'outil (I) qu'il perçoit de manière tactile et visuelle. Avec ce stylet il pointe directement sur les feuilles réelles (Objet réel de la tâche = Or) pour faire toutes les actions (sélectionner, copier, déplacer, supprimer). Cette action est détectée par la table interactive (Senseur = S) et l'écho de l'action est retourné sous la forme de modifications des pages (augmentation virtuelle = O+v) qui sont superposées par l'intermédiaire du vidéoprojecteur (Effecteur = E) directement sur les feuilles physiques par projection. Cette augmentation est non seulement perçue par le rédacteur mais aussi par les relecteurs grâce à la grande taille de la table et parce que l'augmentation est réalisée par projection. Ceci est représenté dans le schéma par le partage de l'opérateur de fusion '+' par les différents utilisateurs. Finalement l'objet de la tâche, tel qu'il est perçu par les utilisateurs, est bien l'ensemble des feuilles réelles augmentées par des annotations virtuelles ce qui produit bien un « objet augmenté » ( $O = Or+v$ ).

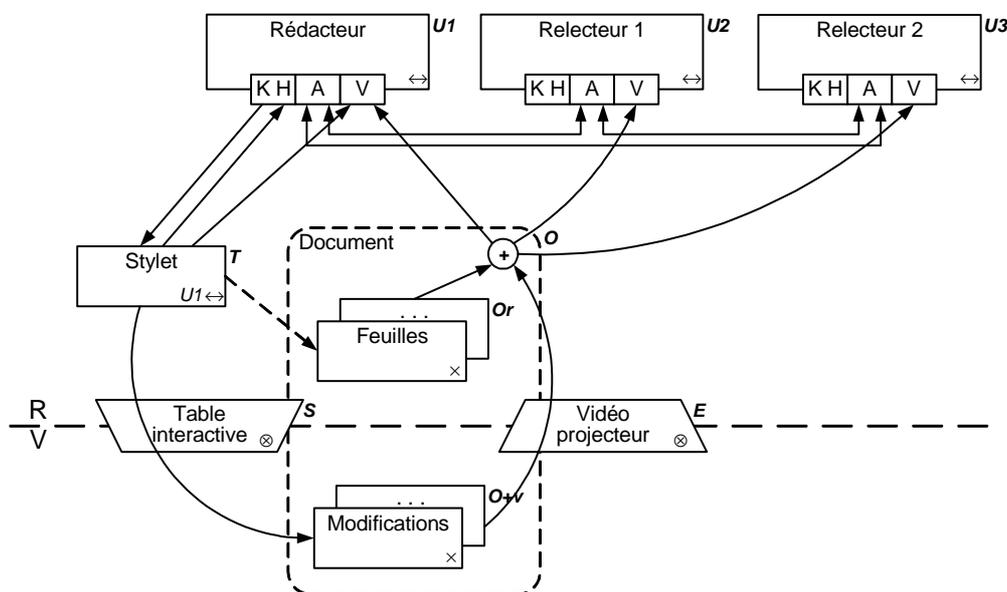


Figure 98 : Activité de co-édition en local avec le logiciel WordTab représentée avec IRVO.

Enfin, sur le schéma IRVO est représentée sous forme de flèches à double pointes la communication orale entre les différents acteurs ce qui permet aussi la coordination nécessaire pour réaliser le travail.

En dehors de l'action d'édition des paragraphes que nous venons de détailler, les autres fonctions de WordTab sont toutes mise en œuvre sous forme classique de boutons et menus déroulants. L'interaction se fait toujours avec la table et le stylet qui se comporte dans ce cas comme une simple souris pour accéder aux différentes fonctions.

### VIII.4.3 Réalisations

#### VIII.4.3.1 Dispositif technique - Table interactive

La Table Interactive que nous avons mis en œuvre est directement inspiré du DigitalDesk de Pierre Wellner (Cf. II.2.1.2). Elle est présentée Figure 99 sous forme de schéma et, en Figure 100, sous forme de photographie du dispositif tel qu'il était installé au laboratoire (avant son démontage pour cause de changement de locaux).

Notre dispositif ne cherche pas à rivaliser avec les nombreuses tables et tableaux augmentés de la littérature (Cf § III.3.1.4). Nous avons choisi d'utiliser des dispositifs du commerce faciles à assembler, simples et fiables d'utilisation.

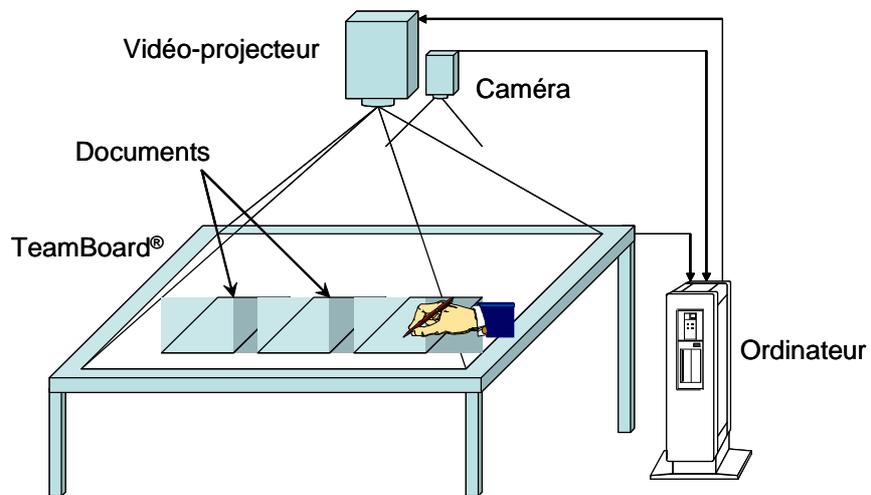


Figure 99 : Schéma de notre dispositif de Table Interactive.



Figure 100 : Photographie de notre Table Interactive (ancienne installation).

Le cœur du dispositif est un tableau blanc interactif de marque TeamBoard<sup>17</sup>. Tel qu'il est vendu, ce tableau blanc interactif permet d'écrire avec des feutres conventionnels comme sur tout tableau blanc. Cette écriture (ou plus généralement tout dessin) est capturée par un logiciel *ad hoc* fourni et les images ainsi obtenues peuvent être archivées sur un ordinateur, et éventuellement modifiées, imprimées ou transmises à d'autres personnes. Dans un autre mode de fonctionnement, il est possible de projeter l'image de l'ordinateur sur le tableau blanc avec un vidéoprojecteur et l'ensemble se comporte comme un écran tactile de grande dimension, chaque pression sur le tableau avec un stylet ou simplement le doigt simulant un clic souris.

Nous avons placé le tableau blanc de manière horizontale pour le transformer en table et en fixant le vidéoprojecteur au-dessus. En pratique, la table faisant près de 1,2 m de profondeur par 1,5 m de large, nous avons gardé une inclinaison d'environ 20 % de la table (à la manière d'une table à dessin) pour que les personnes puissent accéder à la partie arrière de la table sans avoir à trop se pencher en avant. Cette inclinaison n'est pas gênante pour poser des feuilles de papier car les frottements les empêchent de glisser. A l'usage, on s'est même aperçu que, en frottant les feuilles, l'électricité statique permettait de véritablement « coller » les feuilles sur la table.

Le DigitalDesk utilise une caméra pour détecter les actions des utilisateurs. Nous n'avons pas retenu cette solution car nous avons jugé la résolution et la précision insuffisantes. En effet, une image vidéo numérisée permet au mieux une image de 640 x 480 pixels alors que le TeamBoard a une résolution de 2048 x 2048. Comme nous utilisons une résolution d'affichage de 1024 x 768, on voit tout de suite l'intérêt de cette configuration. Toutefois, nous avons également installé dans notre dispositif une caméra au-dessus de la table dans le but de détecter automatiquement les feuilles de papier plutôt que de le faire dans une phase de calibration manuelle. La détection est réalisée en entourant les feuilles d'un cadre noir épais et en utilisant un algorithme de reconnaissance simple (Cf. ci-dessous).

#### VIII.4.3.2 Logiciel WordTab

Au démarrage de WordTab, la seule action possible est d'ouvrir le document de travail (au format Word ou tout autre format reconnu par ce logiciel) en sélectionnant un fichier existant dans le système de fichiers. Il n'est pas possible de partir d'un texte vide puisque WordTab ne permet pas de créer de nouveaux textes mais seulement de déplacer, de dupliquer ou de supprimer des paragraphes existants. Il faut également disposer d'une version imprimée du même document.

Ensuite, le logiciel fonctionne selon deux modes principaux, le basculement se faisant par deux boutons placés en bas vers le centre (le bouton du mode actif reste enfoncé) :

- Un mode de positionnement des feuilles réelles sur la table,
- Un mode d'édition du texte.

##### *Mode de positionnement des feuilles*

L'interface correspondante à ce mode est présentée Figure 101. Cette figure est une recopie d'écran : en réalité, cette image est projetée sur la table et l'intérieur des cadres noirs contient la feuille de papier physique avec son texte imprimé.

Dans ce mode, on pose les feuilles physiques sur la table et on apprend au logiciel le numéro de la page et la position de celle-ci sur la table. Pour pouvoir plaquer facilement les feuilles sans risquer des actions indésirées, le bouton « cadenas » permet de verrouiller la table, c'est-à-dire que plus aucune action n'est possible sauf déverrouiller la table en appuyant de nouveau sur le cadenas. La grille qui s'affiche en fond d'écran permet de mieux aligner les feuilles horizontalement.

---

<sup>17</sup> <http://www.teamboard.com/>

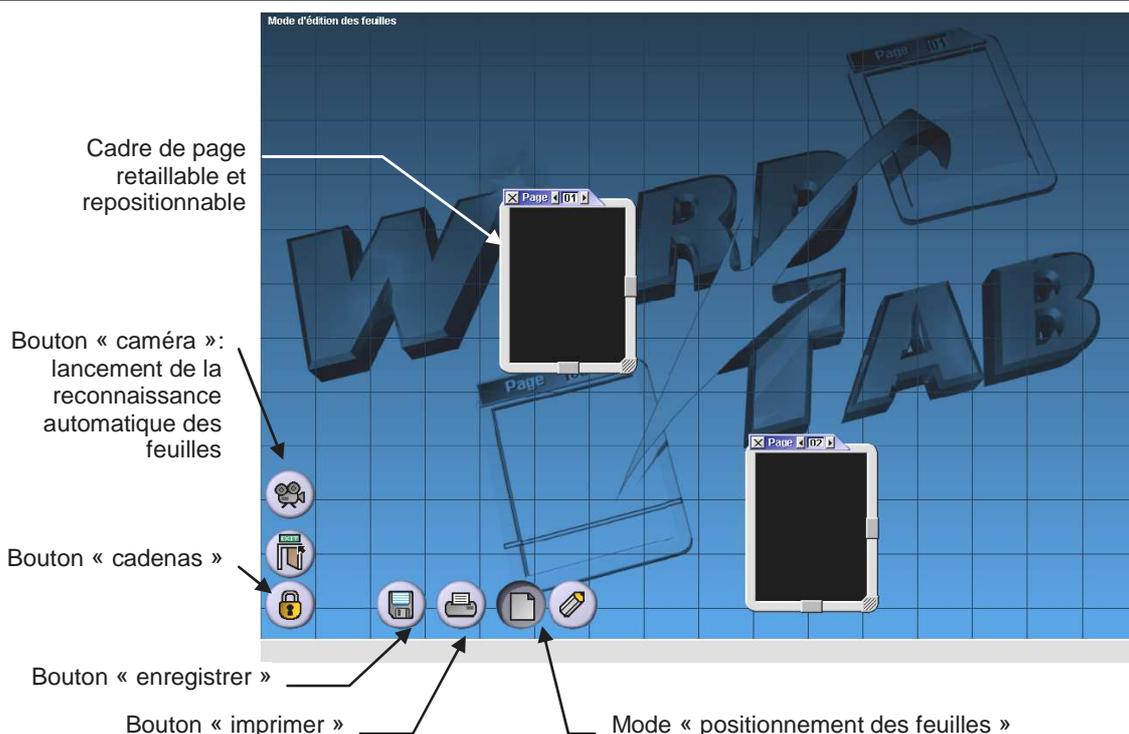


Figure 101 : Interface de WordTab en mode « positionnement des feuilles ».

Une fois les feuilles posées, il suffit de cliquer sur le point haut-gauche de chaque feuille pour faire apparaître un cadre à fond noir qu'il est possible de redimensionner pour qu'il s'ajuste le plus exactement possible à la feuille physique. Il est important d'apporter le plus grand soin à cette phase de calibration car sinon les annotations projetées ne tomberont pas exactement sur les paragraphes. Notons qu'une ligne de texte imprimée en corps 12 correspond à environ 3 pixels de haut pour le vidéoprojecteur !

Par défaut, les pages sont numérotées 1, 2, 3, ... dans l'ordre de création des cadres, mais cela peut être modifié avec les boutons d'incrément/décément placés sur le bandeau.

Dans une version récente du logiciel, il est également possible de faire le positionnement des pages automatiquement (en utilisant la caméra de la Table Interactive) en sélectionnant le bouton « caméra ». Il faut imprimer au préalable les pages avec un cadre noir épais pour que le logiciel puisse les reconnaître. Par défaut, il les numérote dans le sens conventionnel de lecture de gauche à droite et de haut en bas. Malheureusement, le résultat est loin d'être parfait car la résolution de la caméra est insuffisante et l'algorithme de reconnaissance pas assez précis : il faut donc retoucher à la main la position des cadres après la phase de reconnaissance. Malgré ces défauts, cette fonctionnalité gagne du temps par rapport au positionnement entièrement manuel.

### Mode édition

L'interface correspondante à ce mode est présentée Figure 102. Comme nous l'avons déjà précisé il faut imaginer que les cadres noirs des pages contiennent normalement la feuille de papier physique avec son texte imprimé.

Les 5 icônes à bas à droite représentent les 5 actions possibles :

- Le bouton « nouvelle sélection » permet de créer une nouvelle sélection de paragraphe à laquelle est affectée une nouvelle couleur. Il suffit de pointer sur le paragraphe à sélectionner. Ensuite, tous les pointages sur des paragraphes ajoutent ces paragraphes à la sélection courante qu'ils soient contigus ou non. L'écho des sélections est réalisé en projetant la couleur sous forme d'un bloc recouvrant la totalité du paragraphe. Un pointage sur un

paragraphe déjà sélectionné, le désélectionne. Il est également possible de sélectionner plusieurs paragraphes contigus en faisant glisser le stylet sur les paragraphes sans relâcher la pression sur la table.

- Le bouton « aucune action » permet d'affecter à une sélection... aucune action ! C'est surtout intéressant pour annuler une action sur une sélection existante sans supprimer la sélection elle-même.
- Le bouton « supprimer » permet de supprimer la sélection courante. L'action est symbolisée par la projection de hachures sur tous les paragraphes supprimés.
- Le bouton « déplacer » permet de déplacer tous les paragraphes faisant partie de la sélection vers un point d'insertion qui est désigné par pointage. Ce point d'insertion ne peut être qu'entre deux paragraphes. L'écho de cette action est donné, d'une part par une trame fine sur les zones de la sélection, et d'autre part par des flèches partant de ces zones et pointant vers le point d'insertion.
- Le bouton « dupliquer » permet de copier tous les paragraphes faisant partie de la sélection vers un point d'insertion qui est désigné par pointage. L'écho de l'action est similaire au déplacement sauf pour la trame qui est remplacée par un damier.

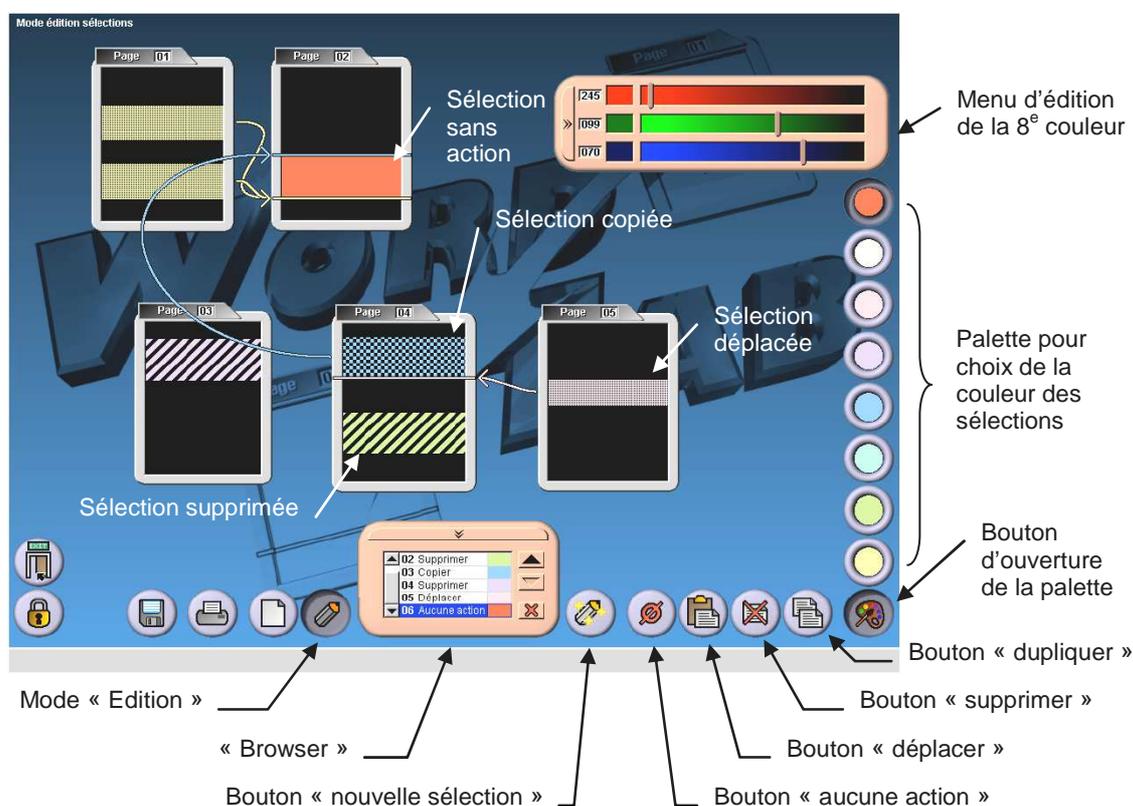


Figure 102 : Interface de WordTab en mode édition

Les 4 actions (aucune action, supprimer, déplacer et dupliquer) sont également accessibles par un menu déroulant (pop-up menu) accessible en pointant n'importe où sur le fond de la table.

Deux outils complémentaires sont disponibles :

- Un menu déroulant, appelé « browser » placé en bas au centre. En l'ouvrant, il donne la liste complète des sélections (numérotées dans l'ordre), de l'action associée et de la couleur de la zone. Il est ainsi possible de rendre active une ancienne sélection, de changer l'ordre des actions en remontant ou descendant les sélections dans la liste et de supprimer des sélections (ce qui a pour effet d'annuler l'action et la sélection pas de supprimer le texte !)

- Une palette de pots de peinture qui s'ouvre par le bouton « palette » placé en bas à droite et qui permet de changer la couleur par défaut des sélections. La huitième couleur (la plus en haut) est même entièrement éditable au format RVB.

#### *Fin du travail*

Nous avons vu que les sélections et les actions associées restent éditables à tout moment : il est possible d'enlever ou supprimer des paragraphes dans une sélection, d'annuler des actions, de changer l'ordre des actions, etc.

Les actions ne sont réellement accomplies qu'au moment de la sauvegarde qui est réalisée en pointant sur le bouton « enregistrer ». A ce moment là, toutes les actions de copie, déplacement, suppression sont accomplies sur la version électronique du document qui se retrouve donc « synchronisé ».

Le bouton « imprimer » voisin, fait la même chose mais en plus imprime directement le résultat. Ainsi, il suffit d'enlever l'ancienne version du document se trouvant sur la table (il est désormais obsolète) et de la remplacer par les feuilles fraîchement imprimées et de continuer sur une version propre.

#### *Implémentation*

WordTab est basé sur le logiciel Word de Microsoft. Il permet donc de manipuler tout format de texte accepté par ce logiciel mais il comporte certaines limitations : par exemple il ne peut traiter que du logiciel sur une seule colonne et les dessins ne sont manipulables que s'ils sont insérés dans le texte.

WordTab est programmé en C++ sous Windows et s'appuie sur la bibliothèque SDL<sup>18</sup> (Simple Direct Media Layer) pour la mise en œuvre de l'interface. L'interface avec Word de Microsoft se fait par l'intermédiaire d'un module programmé en VisualBasic. Le module de reconnaissance des pages par caméra est basé sur la bibliothèque OpenCV<sup>19</sup> fournie par Intel.

### **VIII.4.4 Expérimentation de WordTab**

Une procédure de tests a été mise en place avec un double but :

- Détecter les derniers bogues du logiciel par des utilisateurs externes à l'équipe de développement,
- Essayer d'évaluer l'utilisabilité du dispositif de Table Interactive et du logiciel WordTab dans une situation qui se veut proche d'une situation réelle de travail collaboratif.

#### **VIII.4.4.1 Protocole expérimental**

Les tests, d'une durée d'une quarantaine de minutes, se déroulent par équipe de trois sujets et en présence d'un membre du projet qui présente le dispositif dans une première partie puis qui a un rôle d'observateur pendant les tests proprement dit. Un test se compose de 3 grandes phases :

##### *Une phase de présentation et prise en main du dispositif*

- Une présentation orale du dispositif : projecteur, table tactile, introduction de la notion de réalité augmentée ; suivie d'une présentation rapide du principe de fonctionnement du logiciel : impression d'un texte Word présent dans l'ordinateur, logiciel de traitement de texte où l'unité de travail est le paragraphe.

---

<sup>18</sup> <http://www.libsdl.org/index.php>

<sup>19</sup> <http://www.intel.com/research/mrl/research/opencv/>

- Une démonstration du travail avec le logiciel : après avoir expliqué les étapes à suivre lors de l'utilisation du logiciel (placement des feuilles, sélection des zones à modifier, opérations, sauvegarde, réimpression), le membre du projet présent utilise un texte jouet pour faire une démonstration de ce travail. Ce texte jouet contient des paragraphes dans le désordre et des parties superflues ; le contenu est explicite pour que l'ordre d'origine soit évident : les paragraphes comportent soit des nombres de un à douze (écrits en toutes lettres), soit un mot ressemblant à l'un de ces chiffres, répété sur quelques lignes.
- Après cette démonstration, les sujets sont invités à tester les différentes opérations (placement, sélection, opérations) et fonctionnalités par eux-mêmes sur le même texte : ils peuvent alors poser toutes sortes de questions techniques au membre du projet. C'est la phase pendant laquelle ils apprennent à manipuler le logiciel, d'une durée de 10 à 15 minutes.

### *La phase de test proprement dite*

Les sujets doivent effectuer un travail seul, et sont observés. Pour se mettre dans une situation de travail collaboratif, un texte de 4 pages a été conçu où alternent des morceaux d'un sketch de Bigard découpé en paragraphes et mélangés, et des parties « parasites », des textes sur la réalité augmentée ou des passages d'une nouvelle de Maupassant.

- On donne à lire à chaque sujet une page et demie de ce texte en leur précisant dans une consigne écrite le type et le thème du texte d'origine, ainsi que le travail que l'on attend d'eux : retrouver le texte initial. Ils se trouvent ainsi dans une situation proche de l'édition commune d'un rapport écrit à plusieurs : chacun a écrit une partie et n'a pas lu celle des autres, mais en connaît tout de même le thème.
- Après ce travail préliminaire, tous les sujets se retrouvent sur le dispositif où ils doivent disposer les feuilles sur la table interactive (étape de calibration) puis utiliser le logiciel pour remettre le texte dans l'ordre tout en éliminant les parties superflues.

Lors de cette étape, le membre du projet est présent pour répondre à toute question technique sur la manipulation, il note également ses remarques dans une fiche qui permet de noter un certain nombre d'observations sur le déroulement du test :

- Caractéristiques de l'équipe pouvant avoir une influence sur leur travail : leur expérience en informatique, leur expérience du travail collaboratif et le fait qu'ils se connaissent ou non.
- Le temps qui leur a été nécessaire pour effectuer le travail demandé, et si le texte s'est bien retrouvé dans l'ordre voulu à l'issue de la manipulation.
- Des remarques sur l'utilisation du dispositif de la table interactive (positionnement autour de la table, mobilité, gêne/accès aux fonctions, manipulation ou annotations sur les feuilles de papier réelles)
- Des remarques sur l'utilisation du logiciel WordTab (manipulation aisée des différentes fonctions, erreurs fréquentes).
- Du point de vue du travail collaboratif :
  - La méthodologie employée pour le travail (actions effectuées en premier, en dernier)
  - Les stratégies de partage des tâches (immédiat, maintenu, source de conflits ou non),
  - La communication dans l'équipe,
  - les accidents de manipulation dus à des opérations simultanées et les éventuels énervements ou autres problèmes survenus au sein du groupe.

La Figure 103 montre une équipe en train de travailler avec WordTab.

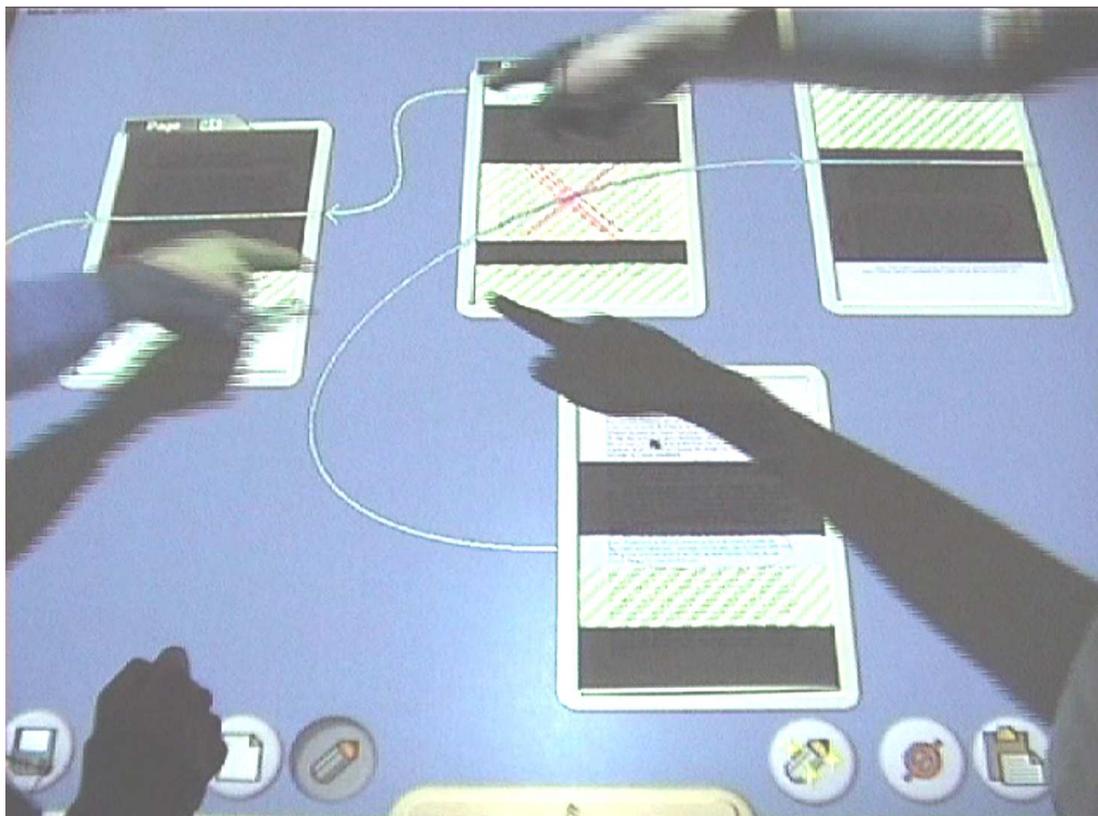


Figure 103 : Exemple d'utilisation de WordTab avec 3 utilisateurs.

#### *Une phase d'évaluation*

Les sujets ont à remplir un questionnaire après le test : celui-ci se compose de huit questions, sous forme de QCM ou de questions ouvertes. Les premières questions portent sur l'utilisation du logiciel, le temps estimé d'apprentissage et le temps d'amortissement. Les questions ouvertes demandent au sujet l'intérêt qu'il a vu au logiciel, les utilisations pour lesquelles on pourrait l'adapter, le temps qu'il estime avoir gagné en utilisant ce logiciel et tout autre commentaire ou suggestion.

Par ailleurs, un test complémentaire a été réalisé : quelques sujets n'effectuent que la phase 2, munis uniquement d'un ordinateur classique : le texte est le même, les utilisateurs ont également une connaissance partielle du texte. Pour ces tests, une observation systématique n'est pas prévue, car cette situation de travail nous est bien connue, et étudier le travail sur ordinateur n'est pas l'objectif de ces tests. On note en revanche le temps mis pour remettre le texte dans l'ordre, que l'on compare au temps mis par les volontaires passant sur le dispositif TableGate.

#### **VIII.4.4.2 Déroulement des tests**

Les sujets étaient des étudiants volontaires de l'Ecole Centrale. Les équipes ont été formées en prenant en compte leur expérience préalable en informatique : des experts, des personnes ayant une expérience correcte des traitements de texte, des équipes ayant déjà rédigé des rapports ensemble, des équipes dont les membres se connaissent, d'autres non. Cinq équipes ont participé aux tests pour un total de 13 sujets (2 équipes ne comportaient que 2 personnes en raison de problèmes d'organisation).

### VIII.4.4.3 Résultats issus des observations

#### *A propos du temps*

Au niveau du temps, il n'y a pas de différence significative : les équipes ont mis entre 2 et 3 minutes pour placer les feuilles, entre 9 et 13 minutes pour effectuer la remise en ordre du texte. Toutes ont réussies à éliminer les paragraphes intrus et à remettre le texte dans l'ordre avec cependant quelques erreurs dues à la compréhension et non à la manipulation. Il est à noter que l'équipe ayant mis le plus de temps est celle qui était la plus expérimentée à la fois en informatique et en travail collaboratif, et que l'équipe la plus rapide était composée d'élèves qui n'avaient jamais travaillé en commun sur un texte, ne se connaissaient pas avant le test, et semblaient assez timides. Puisqu'il ne s'agit que de quelques personnes, on se gardera d'en déduire que l'utilisation de WordTab est la plus facile pour les personnes ne maîtrisant pas les traitements de textes habituels, mais on en retiendra déjà que ce logiciel est assez simple d'utilisation pour que beaucoup de personnes, plus ou moins expérimentées, puissent s'en servir.

#### *A propos de l'utilisation de l'interface*

Les sujets n'ont pas eu de problèmes majeurs pour accéder aux fonctions du logiciel, s'aidant éventuellement les uns les autres pour trouver tel ou tel bouton. Aucun problème opératoire n'a été bloquant. Pour l'accès aux fonctions, il y a beaucoup de variations selon les sujets : certains ont utilisé quasi-exclusivement les icônes, d'autres ont préféré le menu déroulant. Tous ont compris le rôle du *browser*, mais tous, sauf un, ont fait appel au membre observateur de TableGate pour des précisions quant à son utilisation lorsqu'ils ont voulu s'en servir.

Les sélections dites « avancées », c'est-à-dire le fait d'effectuer la même opération (par exemple la suppression) sur plusieurs paragraphes à la fois, ont été utilisées au moins une fois par équipe. Les équipes ont également fait appel au membre observateur dans le cas où plusieurs flèches se sont retrouvées à pointer vers le même endroit : il est vrai qu'il faut alors se rappeler quelle opération a été faite en premier (passer par le *browser* est également possible), ce qui est très différent des situations habituelles où les modifications sont faites en temps réel, permettant une vérification immédiate du travail effectué. Cependant, le fait de ne pas avoir sous les yeux le texte final, à l'issue du travail, n'a pas gêné les sujets : aucune équipe n'a oublié de partie de texte à cause de cela et les seules erreurs dans les textes modifiés ont été dues à des erreurs de compréhension.

Pour le travail sur le texte, dans chaque équipe, chaque personne a manipulé le logiciel. Le partage des tâches s'est organisé de façon naturelle dès le début du travail et n'a pas causé de problèmes, sauf pour quelques sujets qui, voulant aller vite, ont appuyé en même temps sur la table.

Les sujets ont peu profité des possibilités offertes par les feuilles réelles de papier : aucun sujet n'a soulevé une feuille ou pris une feuille pour pouvoir la lire, aucun sujet n'a écrit sur une feuille alors que les consignes qui leur ont été distribuées précisaient qu'ils pouvaient faire ce qu'ils voulaient avec leur copie du texte, et que des stylos leur étaient fournis. Les avantages de la réalité mixte n'ont pas été utilisés et les réponses au questionnaire ont montré que les sujets n'en avaient pas bien compris le principe : pourquoi on utilisait du papier plutôt qu'une projection directe du texte sur la table, et les avantages que l'on pouvait en tirer.

#### *A propos de la méthodologie de travail*

Sur la méthodologie de travail, nous avons constaté une grande uniformité de comportement. En particulier nous avons observé un comportement qu'on pourrait qualifier de « propriétaire » : dans la plupart des cas, chaque sujet plaçait lui-même sur la table les feuilles qui lui avaient été attribuées pour lecture, effectuait la calibration puis restait devant. Etant donné le manque de place autour de la table, on aurait pu imaginer que le sujet qui a fini « sa » mise en place des feuilles se décale sur le côté ou se recule pour ne pas gêner le suivant. Malgré ce comportement, ils n'ont pas eu de

difficultés pour lire le texte, opérant des translations d'un côté à l'autre de la table au cours du travail mais sans ne jamais effectuer de permutation avec l'un des voisins.

Afin de prendre connaissance des parties qu'ils ne connaissaient pas, les sujets ont souvent commencé le travail en survolant rapidement des yeux les parties inconnues, sauf pour une équipe où les membres ont commencé par lire presque intégralement les parties des autres : ceci leur a pris du temps mais au final, ils ont été les plus rapides à effectuer le travail. Il en serait sans doute été autrement avec un texte encore plus long et moins typé. Comparativement, les équipes qui ont utilisé un ordinateur classique ont adopté la même stratégie : avant de commencer les modifications, elles ont fait défiler à l'écran le texte entier. Pour cette étape, TableGate est plus avantageux puisque, le texte étant visible en totalité, les sujets peuvent lire simultanément : ainsi chacun peut découvrir à son rythme les parties inconnues sans avoir à relire celles qu'il connaît déjà.

La méthode de travail généralement employée pour réorganiser le texte est de supprimer en premier lieu les paragraphes superflus, puis de rechercher dans l'ordre le premier paragraphe, puis le deuxième... suivant une démarche complètement linéaire. Au cours du travail, les sujets se font de très brefs résumés de ce qu'ils ont lu, mais sont tout de même amenés à survoler des paragraphes n'appartenant pas au texte qu'ils ont eu à lire, ce qui prend du temps, mais sans les lire entièrement. En ce sens, le transfert d'informations pertinentes dans l'équipe est un peu faible, en revanche la communication est importante. Un bon équilibre et un partage des tâches est trouvé dès le départ, les accidents de manipulation sont peu nombreux, il n'y a pas eu de conflits dans l'équipe. Le travail s'est déroulé dans une bonne ambiance.

Une seule équipe n'a pas commencé la réorganisation en supprimant les paragraphes parasites, mais a quand même effectué un travail entièrement linéaire (comme toutes les autres), faisant toutes les suppressions en dernier. Il est intéressant de noter cet attachement à une démarche linéaire alors que WordTab permet justement d'autres démarches peut-être plus efficaces : ayant tout le texte sous les yeux, on peut tout à fait choisir de regrouper d'abord des paragraphes qui semblent se suivre, ou vouloir commencer par la fin. Lors de la présentation du logiciel, le membre de TableGate prenait pourtant le soin de montrer que c'était possible, mais cet aspect du logiciel n'a pas été assimilé par les sujets. On a même vu quelques équipes reconnaître très vite le paragraphe final (ce qui était assez simple puisque le texte retenu, un texte comique, comprenait une chute facilement identifiable), mais choisir de ne le déplacer vers la fin du texte qu'en dernier, après toutes les autres modifications. Une équipe a mis à profit dans cette situation la fonction "aucune action", qui permet de surligner une partie de texte et de n'effectuer la modification que plus tard ; les autres équipes n'ont jamais utilisé cette fonction.

#### VIII.4.4.4 Résultats issus des questionnaires

##### *A propos de l'utilisation*

Les sujets ont exprimé dans les questionnaires que les menus déroulants étaient moins intuitifs et moins faciles d'utilisation que le *browser*, alors que les observations semblaient montrer qu'ils étaient beaucoup moins à l'aise avec le *browser*. Le plus gros reproche fait aux menus déroulants est le fait qu'il faille rester appuyé avant de choisir l'action à effectuer, et certains ont proposé que le menu déroulant reste ouvert après une pression sur la table et jusqu'à la pression suivante. Cette difficulté avec le menu déroulant provient sans doute d'une expérience trop courte d'utilisation du dispositif. Plus de la moitié des sujets ont répondu qu'ils étaient « à l'aise pour travailler de façon autonome » sur le dispositif après les 10 minutes de présentation et d'essais du logiciel : l'expérience a prouvé le contraire puisque tous ont demandé des renseignements techniques par la suite ! Il faut donc sans doute compter plutôt 15 à 20 minutes pour se familiariser avec le dispositif.

Seul un sujet a estimé ne pas avoir gagné de temps par rapport à une équipe qui aurait travaillé avec un ordinateur classique, les autres ont estimé avoir gagné entre 5 et 10 minutes. Environ la moitié pense que le temps de familiarisation au dispositif est amorti au bout d'une seule utilisation : à la

suite de ces tests, nous sommes un peu moins optimistes qu'eux. Tous pensent que WordTab facilite le travail à plusieurs. En ce qui concerne le travail sur les textes longs, les réponses sont plus mitigées : certains (plus particulièrement ceux qui sont les plus expérimentés en informatique) regrettent que les modifications ne soient pas faites en temps réel. En revanche, un sujet a remarqué que garder le texte dans l'ordre jusqu'à la fin permettait de « se rappeler où on a lu les paragraphes », ce qui peut être un grand avantage lors du travail collaboratif, et auquel nous n'avions pas pensé.

#### *A propos des commentaires et suggestions sur WordTab*

Certains des commentaires ou suggestions ont montré une compréhension incomplète des objectifs et avantages du dispositif, ou étaient impossibles techniquement : désir de voir les modifications en temps réel, modifications du texte à petite échelle (orthographe par exemple), suppression des feuilles de papier... D'autres propositions étaient des choses auxquelles nous avions pensé, même implémentées puis remplacées par des fonctions plus performantes : il s'agissait de numéroter les flèches pour se souvenir de l'ordre des actions, d'accéder à une fonction en traçant un symbole sur la table, d'écrire le nom des fonctions à côté des icônes correspondantes... Certaines remarques étaient liées à des bogues qui ont été corrigés depuis (bogue lors du déplacement en début ou en fin de texte) ou à des problèmes ergonomiques (problème de luminosité des projections lors des suppressions). Ensuite, suite à une remarque d'un sujet, une amélioration a été introduite : un quadrillage s'affiche sur la table pendant la phase de calibration afin de faciliter le placement des feuilles.

Les questions sur les utilisations possibles du dispositif n'ont rien apporté : les sujets n'y ont vu que l'édition de texte et des applications graphiques. Il aurait peut-être fallu poser la question d'une autre manière comme : « Ce dispositif vous semble-t-il utile ? Dans quelles types d'applications ? », en insistant que la question portait sur le dispositif de table interactive et non seulement sur le logiciel WordTab.

#### **VIII.4.4.5 Bilan des tests**

Il est apparu que WordTab était utilisable par un utilisateur assez rapidement au bout d'une quinzaine de minutes de formation. Son fonctionnement est intuitif pour des utilisateurs qui sont déjà des utilisateurs « dégourdis ». Cependant les concepts qu'il utilise sont assez nouveaux et ne sont en général pas compris par les utilisateurs qui voient dans l'usage du papier un retour en arrière.

Les équipes qui ont fait le test sur un ordinateur classique ont mis les mêmes temps en moyenne, mais avec un écart-type plus important. On peut penser qu'il y aurait eu une différence de temps en faveur de WordTab si le texte retenu avait été plus long, ou si les sujets avaient eu plus de temps pour se familiariser avec le dispositif. Les textes utilisés étaient aussi très typés et non représentatifs d'un article scientifique nécessitant des discussions et des réflexions plus approfondies entre les membres des équipes.

Il est certain que les sujets que nous avons choisis (étudiants d'une Ecole d'ingénieur) ne sont pas représentatifs de la population potentiellement utilisatrice du dispositif. Il serait nécessaire de refaire d'autres tests avec des personnes plus âgées, moins rompues à l'usage des interfaces informatiques graphiques mais qui, par leurs expériences de travail, verraient peut-être plus l'intérêt d'utiliser le papier comme moyen d'interaction.

Il est par contre intéressant de constater que dans les équipes tous les sujets ont manipulé le dispositif à tour de rôle, en général chaque sujet manipulant « ses » paragraphes (c'est-à-dire les paragraphes qu'il avait lus), ce qui nous encourage dans le fait que ce dispositif facilite le travail à plusieurs. A l'inverse, les équipes qui ont travaillé sur un ordinateur classique ont, en général, laissé une personne faire les manipulations, les autres se contentant de donner des instructions.

## VIII.5 Projet TableGate II

Faisant suite à TableGate nous avons lancé un second projet d'étude, TableGate II, toujours basé sur l'usage de la Table Interactive mais abordant cette fois-ci les aspects de travail collaboratif à distance. Ce projet a débuté en octobre 2003 pour une durée de 15 mois et implique 5 étudiants de l'Ecole Centrale de Lyon de la promo 2006.

Nous avons laissé le libre choix du thème aux étudiants. Après un long débat, ils ont finalement décidé de développer un prototype pour assister la gestion des interventions des pompiers sur des sinistres.

Le scénario d'étude est le suivant :

Un appel est reçu sur le standard des pompiers. Le pompier d'astreinte prend la nature et les coordonnées du sinistre et les saisit sur un ordinateur pour l'enregistrer dans une base de données. Sur un plan de la ville, posé sur la Table Interactive, un symbole est projeté sur l'emplacement correspondant au sinistre. Alternativement, il peut pointer directement sur la carte l'emplacement ce qui renseigne directement la base de données. Les emplacements des différentes casernes les plus proches du sinistre sont signalées par projection sur le plan avec indication des ressources disponibles sous forme symbolique (nombre et type de camions). Pour choisir un camion spécifique (et l'équipe d'intervention associée) le pompier prend un cube qui représente ce camion pour le déplacer et le poser vers l'emplacement du sinistre. Cette action déclenche l'alerte dans la caserne correspondante et transmet directement les informations aux pompiers par exemple sur une imprimante et/ou sur le PDA de l'officier. Une communication vocale directe peut aussi être établie directement entre le pompier d'astreinte et l'équipe qui part en intervention. Ensuite, lorsque le camion quitte la caserne, il est suivi par GPS et son trajet est affiché au fur et à mesure sur la carte. Pour les sinistres importants, le pompier peut décider d'envoyer plusieurs camions, de la même caserne ou de plusieurs casernes. Il lui suffit de répéter les manipulations ci-dessus. A tout moment il visualise l'état des ressources en action sur le terrain ce qui lui permet, en cas de nouvel appel, de savoir immédiatement quelles ressources sont encore disponibles.

Ce scénario a été conçu *a priori* par les étudiants. En faisant l'interview d'un responsable du SDIS (Service Départemental d'Incendie et de Secours) du Rhône, ils ont pu le confronter avec la réalité. Les pompiers utilisent à l'heure actuelle un système qui leur affiche directement sur l'écran la carte locale du sinistre en fonction des coordonnées entrées et qui leur propose les moyens disponibles dans les casernes les plus proches sur une liste. Le pompier n'a donc plus qu'à valider la proposition s'il en est d'accord. A l'heure actuelle, ils ne voient pas l'intérêt d'avoir une carte à l'échelle de la ville ou de l'agglomération car elle serait surchargée sans apporter d'aide véritable. Par contre, ils voient tout l'intérêt de ce dispositif dans le cas de sinistres très importants où un PC (Poste de Commandement) délocalisé est mis en place. A l'heure actuelle, l'officier qui a le commandement de ce PC n'a que des moyens de liaison radio avec les équipes et doit mémoriser leur emplacement dans sa tête ! S'il pouvait avoir une Table Interactive dans un véhicule qui lui affiche en permanence la position des équipes, voire des pompiers, l'efficacité de ses ordres pourrait être notablement améliorée.

### Modélisation

Le travail étant en cours, nous ne présentons ici qu'un seul modèle correspondant à la tâche de déplacement d'un camion depuis la caserne vers un sinistre sur la Table Interactive (Figure 104).

Le pompier d'astreinte prend le camion qui est représenté par un bloc physique et est visuellement augmenté par la projection d'une image par-dessus qui renseigne sur le type de camion. Il déplace ce camion sur la carte physique qui est elle-même augmentée par les annotations virtuelles :

emplacement de la caserne, emplacement du sinistre. Cette action est suivie par la caméra qui permet d'une part de projeter l'image du camion au bon endroit et d'autre part (via la structure interne du programme symbolisé par le 'modèle interne' M) de déclencher une alerte et l'affichage de la fiche d'intervention pour le pompier distant.

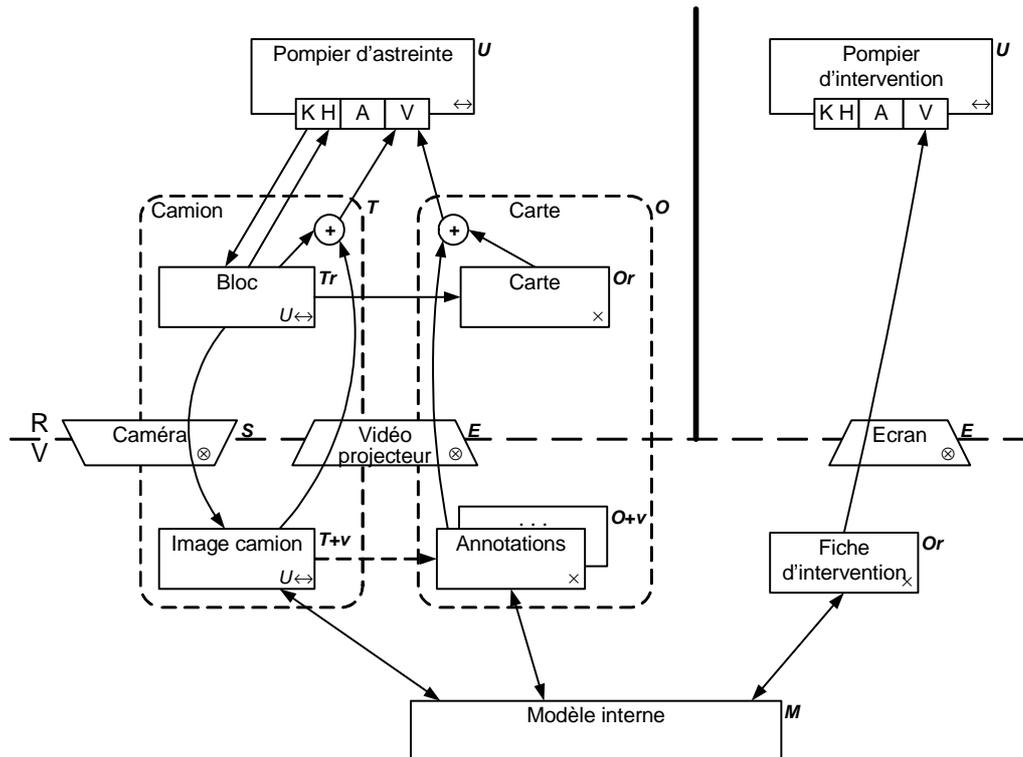


Figure 104 : Modélisation avec IRVO de la tâche « déplacement d'un camion » dans TableGate II.

### Dispositif

Ce projet utilise également notre Table Interactive (c'était une contrainte du sujet !) mais, pour pouvoir assurer le suivi des petits blocs sur la table, les étudiants ont choisi d'utiliser une caméra et l'ARToolkit [Kato 2000] qui permet d'identifier et de localiser facilement des objets en leur ajoutant une petite étiquette comportant un symbole *ad hoc* (Cf. § III.3.1.1).

## VIII.6 Bilan du chapitre

Nous avons présenté dans ce chapitre les réalisations concrètes que nous avons menées dans le cadre du projet ACCORéA au laboratoire ICTT. Une première étude a permis de modéliser l'activité de rédaction d'écriture d'articles scientifiques à plusieurs : parmi les différentes tâches exhibées, nous avons choisi d'en traiter une en détail en la mettant en œuvre dans un prototype, WordTab.

Ce prototype nous a donc permis de confronter la méthodologie de conception que nous avons proposée au Chapitre VII et le modèle IRVO à la conception d'une application nouvelle. Cependant, la tâche mise en œuvre n'était pas suffisamment riche pour pouvoir véritablement servir de test.

L'expérimentation menée a permis de faire ressortir l'intérêt de l'approche pour un travail de réorganisation d'un document à plusieurs même si le document et les sujets n'étaient pas en adéquation totale avec les buts visés par l'application.

Un nouveau projet débute en octobre 2004 au laboratoire. Il s'appelle IMERA (Interfaces Mobiles pour l'Environnement Réel Augmenté) et a pour objectif d'étudier des Interfaces Homme-Machine

innovantes (instrumentées, bi-manuelles, proactives, adaptatives) appropriées pour le travail en situation de mobilité, coopératif et contextualisé dans un environnement réel augmenté.

Ce projet est donc situé dans le domaine de l'informatique ubiquitaire et de l'*Ambient Intelligence*. Ce sera donc une opportunité de tester IRVO dans ce domaine et en particulier de voir comment il peut prendre en compte la mobilité des acteurs et des données. De même, la prise en compte de l'environnement, ou plus globalement du contexte de l'interaction, peut être délicate et nous amener à faire évoluer le modèle IRVO.

Au-delà des laboratoires de recherche, ce projet implique plusieurs industriels ce qui devrait nous permettre d'expérimenter IRVO et la méthodologie associée pour la conception d'applications de plus grande envergure et dans un cadre industriel.



# Chapitre IX Conclusion

*Un beau livre, c'est celui qui sème à foison des points d'interrogation.*  
Jean Cocteau (1889 – 1963).

*La perfection est atteinte non quand il ne reste rien à ajouter, mais quand il ne reste rien à enlever.*  
Antoine de Saint-Exupéry (1900 – 1944).

## IX.1 Bilan de nos contributions

### *Etats de l'art*

Nous avons dressé un état de l'art (Chapitre II) des applications de Réalité Mixte et Augmentée dans un but de recensement et de comparaison des différents domaines d'application et des différentes techniques mises en œuvre. Au-delà de « l'effet catalogue » cette revue nous a été particulièrement utile pour la validation de notre modèle. Nous avons également fait un tour d'horizon des dispositifs d'entrée/sortie utilisés dans les systèmes de Réalité Mixte (Chapitre III) qui nous a permis en particulier d'affiner le concept de transducteur de notre modèle.

### *Analyse critique des modèles existants*

Nous avons passé en revue plusieurs modèles de l'Interaction Homme-Machine (Chapitre IV) et du Travail Collaboratif Assisté par Ordinateur (Chapitre V). Les modèles apparaissent incomplets par rapport à la modélisation des systèmes de Réalité Mixte car ils sont conçus pour la modélisation du logiciel et ne prennent pas (ou peu) en compte les artefacts réels. Seuls deux modèles sont spécifiquement dédiés aux systèmes mixtes :

- Le modèle ASUR [Dubois 2001b], pour la modélisation de l'interaction dans les systèmes mixtes et mobiles, et qui a été récemment étendu [Dubois 2004b].
- Le formalisme de Renevier [Renevier 2004c], pour représenter des scénarios abstraits pour la conception des systèmes mixtes mobiles et collaboratifs.

Les modèles classiques d'architecture logicielle apparaissent eux aussi inadaptés à la prise en compte de la Réalité Mixte et des dispositifs d'entrée/sortie. Cette constatation a conduit à deux extensions du modèle PAC-Amodeus :

- L'une par [Dubois 2001b] pour les systèmes mixtes,
- L'autre par [Renevier 2004c] pour les systèmes mixtes collaboratifs et mobiles.

### *Modèle IRVO*

Notre contribution principale est le modèle IRVO (Chapitre VI). L'objectif principal d'IRVO est de modéliser l'interaction entre le (ou les) utilisateur(s) et le système de Réalité Mixte en explicitant les outils et objets du domaine (tant réels que virtuels) mis en jeu et leurs relations.

IRVO explicite donc l'interaction entre l'utilisateur et les artefacts qu'il manipule. Cette interaction, qui se décompose en action de l'utilisateur sur les outils et les objets et la perception de ces outils et objets, est représentée par un échange d'information entre les entités U (utilisateur), T (outil) et O (objet) du modèle. Pour préciser la nature de cet échange, l'utilisateur est modélisé par ses 'canaux' de communication qui représentent ses sens de perception et ses moyens d'actions.

IRVO insiste tout particulièrement sur la modélisation des outils et objets du domaine : ils peuvent être réels ou virtuels selon leur position par rapport à la frontière entre le monde réel (dans lequel est placé l'utilisateur) et le monde virtuel (dans lequel est le programme informatique). Les objets mixtes apparaissent ainsi comme une abstraction de conception : ils sont composés d'un ou plusieurs objets réels et virtuels et l'augmentation provient soit de la perception « fusionnelle » que les utilisateurs ont des différents composants, soit de leur capacité à agir dans les deux mondes (pour les outils).

Par rapport aux deux autres formalismes existants pour la modélisation des systèmes de Réalité Mixte (ASUR et le formalisme de Renevier) qui représentent eux aussi les utilisateurs, les objets réels et virtuels, IRVO présente quelques spécificités importantes :

- Les objets mixtes sont explicitement représentés par la mise en relief sur le schéma de la relation d'augmentation.
- Les perceptions augmentées sont explicitement représentées sur le schéma sous la forme d'une fusion des perceptions de l'objet réel et de l'objet virtuel.
- Les sens perceptifs et actionnels des utilisateurs sont explicitement représentés dans les schémas IRVO, ce qui permet d'explicitier les sens qui sont requis pour l'interaction avec le système et permet (dans une certaine mesure) de représenter des interactions multimodales.
- Les dispositifs d'entrée/sortie sont considérés comme étant composés de deux parties : la fonction de conversion entre monde réel et monde virtuel (appelée transducteur) et la partie réelle du dispositif qui peut aussi être considérée comme un outil au sens de notre modèle. Cependant, dans la plupart des cas les deux composants peuvent être fusionnés.

IRVO permet de représenter l'interaction de plusieurs utilisateurs avec le système sur un même schéma, ainsi que les communications interindividuelles entre les utilisateurs, qu'elles soient médiatisées ou non. Il permet donc de supporter la conception d'applications collaboratives même si tous les aspects du travail collaboratif ne sont pas couverts par IRVO (en particulier les aspects de coordination). Ces aspects sont également couverts par le formalisme de Renevier mais pas par ASUR.

Enfin, pour franchir la frontière entre le monde réel et le monde virtuel, IRVO introduit la notion de transducteur qui recouvre les senseurs (du réel vers le virtuel) et les effecteurs (du virtuel vers le réel). Ces entités modélisent la fonction de conversion de l'information du monde réel en données numériques permettant au système informatique de « percevoir » les utilisateurs et les objets du monde réel et « d'agir » sur ces mêmes utilisateurs et objets. Dans notre modèle, un dispositif d'entrée/sortie est composé d'un transducteur et d'une partie réelle qui peut aussi être considérée comme un outil. Ainsi, il nous semble important de séparer les deux notions dans notre modèle car, du point de vue de l'utilisateur, le retour haptique que procure le dispositif peut être important dans le contrôle que l'utilisateur a du dispositif. Ceci constitue une différence fondamentale entre les transducteurs du modèle IRVO et les adaptateurs du modèle ASUR qui sont la représentation complète des dispositifs d'entrée/sortie.

#### *Evaluation des capacités descriptives du modèle IRVO*

Nous avons appliqué le modèle IRVO aux applications de la littérature que nous avons recensées au Chapitre II. Nous l'avons également comparé avec ASUR et le formalisme de Renevier sur des exemples. Ceci nous a permis de montrer le pouvoir descriptif d'IRVO : c'est un outil qui permet de comprendre, de comparer et d'analyser les applications de Réalité Mixte existantes.

#### *Taxonomie des systèmes mixtes*

Notre taxonomie (Chapitre VI) se base sur la nature des objets présents dans l'application en distinguant les deux types : outil et objets du domaine de la tâche. Cette taxonomie distingue 16 cas de base obtenus par le croisement de la nature de l'objet de la tâche (objet réel, objet réel augmenté,

objet virtuel augmenté, objet virtuel) avec la nature de l'outil utilisé pour interagir (outil réel, outil réel augmenté, outil virtuel augmenté, outil virtuel). Elle permet :

- D'affiner les deux grandes catégories de la Réalité Mixte qui sont la Réalité Augmentée (l'objet de la tâche est réel) et la Virtualité Augmentée (l'objet de la tâche est virtuel) [Dubois 1999]
- De pouvoir distinguer la Réalité Mixte des systèmes non mixtes de l'informatique classique WIMP ou de la Réalité Virtuelle.

Nous l'avons appliquée aux applications recensées au Chapitre II, ce qui nous a ainsi permis de reclasser une application que ses auteurs appelaient « augmentée » mais qui est en pratique du domaine de la Réalité Virtuelle.

### *Conception de systèmes mixtes*

Le modèle IRVO est aussi un outil adapté à la conception d'applications nouvelles. Il peut être utilisé de deux manières :

- Comme point de départ de la conception en formalisant des cas typiques d'interaction avec le futur dispositif. De ce point de vue il peut être vu comme un formalisme graphique complétant des scénarios textuels.
- Comme modèle résultant de l'analyse : après décomposition de l'application en arbre de tâches, pour chaque tâche utilisant la réalité mixte, un schéma IRVO associé est construit pour modéliser finement l'interaction de l'utilisateur avec les outils et l'objet du domaine de la tâche.

Les modélisations ainsi construites peuvent être évaluées par un ensemble de règles basées sur des propriétés ergonomiques liées à l'utilisabilité ce qui permet de faire une analyse prédictive du système en cours d'élaboration. Il est ainsi possible de comparer différentes solutions alternatives au niveau des dispositifs d'interaction envisagés (outils et transducteurs) et d'écarter les solutions inadaptées à la tâche.

Ces modélisations peuvent ensuite être réunies sur un schéma global, ce qui permet de faire ressortir d'éventuels problèmes de conception au niveau de l'application globale et en particulier de vérifier des règles de continuité de l'action entre tâches.

### *Méthodologie de conception*

Nous proposons d'utiliser IRVO dans une démarche MBA (Model-Based Approach) en partant de la méthodologie CoCSys qui permet la conception de systèmes coopératifs capillaires (Chapitre VII). IRVO s'intègre ainsi comme un modèle complémentaire du modèle de comportement : il s'articule aisément avec les autres modèles, en particulier avec l'arbre de tâches et certains aspects du modèle de l'utilisateur. Nous avons montré les apports d'IRVO, tant au niveau du recueil des scénarios où il peut servir de support à la discussion entre les différents acteurs (concepteurs, futurs utilisateurs, etc.), qu'au niveau du modèle comportemental pour affiner la conception (Cf. ci-avant).

Pour faire le lien avec le logiciel et son architecture nous avons montré les relations qui existent entre les modèles IRVO et AMF-C et le modèle d'architecture Arch. En effet IRVO se positionne du côté de l'interface utilisateur (modèle de présentation) et AMF-C se positionne au niveau du modèle de dialogue et nous faisons correspondre aux objets IRVO des agents AMF-C ce qui permet d'explicitier l'entité 'M' de IRVO et d'aller vers la formalisation du contrôleur de dialogue de l'application.

### *Prototype*

Nous avons développé un prototype d'éditeur de texte collaboratif augmenté, WordTab, basé sur l'utilisation d'une Table Interactive (Chapitre VIII). La conception de ce prototype nous a permis

de faire un premier test de la méthodologie de conception que nous avons proposée. Cependant, l'application n'avait pas une envergure suffisante pour permettre une évaluation significative de la méthodologie.

## IX.2 Perspectives

Les travaux accomplis jusqu'à présent nous ont permis de construire un ensemble cohérent constitué d'un modèle de l'interaction et d'une première version d'une méthodologie de conception associée. Il reste cependant de nombreux prolongements qui se situent sur plusieurs plans.

### *Validation d'IRVO*

Malgré les différentes évaluations que nous avons proposées, le modèle IRVO n'est pas validé. Il est bien sûr très difficile de valider formellement un tel formalisme, et ce n'est que par son utilité dans la conception sur le long terme qu'il prouvera ses capacités.

Au niveau de ses capacités descriptives, il nous faut poursuivre la revue de la littérature que nous avons mené pour modéliser systématiquement les applications de Réalité Mixte ce qui permettra de faire ressortir certaines limites.

Au niveau de ses capacités génératives, nous comptons sur le projet IMERA (Cf. § VIII.6) qui débute pour mettre à l'épreuve IRVO sur des applications de taille conséquente, ce que le projet ACCORÉA ne nous a pas permis de faire.

### *Editeur de schémas IRVO*

A l'heure actuelle les schémas IRVO sont dessinés « à la main » par le concepteur en suivant les règles du formalisme graphique. Un embryon d'éditeur a été développé sous la forme d'un gabarit d'extension du logiciel Visio de Microsoft ; il s'agit d'une simple bibliothèque d'éléments graphiques facilitant l'édition de schémas propres (tous les schémas de ce rapport ont été réalisés avec cet éditeur). Il nous semblerait donc intéressant, soit d'étendre ce gabarit, soit de refaire un autre éditeur qui assisterait mieux le concepteur en proposant des outils de vérification automatique de la syntaxe et qui pourrait aussi l'aider à vérifier la cohérence et la complétude des schémas produits de manière semi-automatique voire automatique.

### *Etudier l'apport des Sciences Humaines*

Nous avons fondé IRVO sur une approche centrée sur l'interaction de l'utilisateur avec le système de Réalité Mixte. Même si nous prenons en compte certains aspects ergonomiques, cette approche reste centrée système. Il serait intéressant de prendre une autre approche venant des Sciences Humaines et centrée sur l'humain et de la confronter à nos travaux. Nous pensons tout particulièrement à la phénoménologie de la perception qui a beaucoup étudié la relation à l'objet et à la Théorie de l'Activité.

### *Extensions du formalisme d'IRVO*

De nombreuses extensions du formalisme sont envisageables.

La première concerne la prise en compte des aspects liés à la mobilité comme c'est déjà le cas pour ASUR et le formalisme de Renevier. Il existe déjà dans le formalisme d'IRVO une représentation partielle de la mobilité des objets et des utilisateurs et il s'agira essentiellement d'affiner ces concepts.

La deuxième extension concerne l'ajout de relations temporelles dans le formalisme d'IRVO pour prendre en compte les aspects critiques liés à la boucle sensori-motrice dans l'interaction de l'utilisateur avec le système au niveau de la tâche. Un premier pas dans cette direction est déjà fait

dans les transducteurs par la caractérisation du temps de traversée qui est un facteur essentiel dans la qualité de l'interaction.

D'autres extensions sur les caractéristiques des objets et des outils pourraient aussi être représentées par le formalisme. Il s'agit cependant de ne pas rendre trop complexe IRVO en multipliant les entités au-delà de ce qui est nécessaire sinon la capacité de synthèse qu'offre ce formalisme serait perdue.

#### *Affiner la méthodologie de conception*

La méthodologie de conception que nous proposons est encore à l'état d'ébauche. Cependant nous partons d'une méthodologie existante, CoCsys et il s'agit donc essentiellement de s'insérer dans le processus de développement en établissant des liens entre les différents modèles ce que nous avons déjà faits pour deux modèles :

- Avec l'arbre de tâches et plus spécifiquement CTT, au niveau du modèle comportemental de l'application,
- Avec le modèle AMF-C, au niveau de l'architecture du logiciel.

De nombreux aspects restent encore à travailler pour prendre en compte tous les impacts de l'insertion de ce nouveau modèle dans la méthodologie en particulier dans la phase de « spécialisation, contextualisation, adaptation » et au niveau de l'architecture du logiciel.

Une direction importante pour compléter la méthodologie est de constituer une base de patterns d'interaction permettant de concevoir plus facilement des applications nouvelles. Cette base pourrait facilement être bâtie à partir d'une généralisation des modélisations que nous avons faites des applications de la littérature.

#### *Tester la méthodologie de conception*

Même complétée et affinée, la méthodologie est encore à l'état théorique et il nous faut donc la valider en la testant sur une application d'une certaine envergure ce que ne nous permettait pas les applications du projet ACCORÉA. Cette opportunité nous sera sans doute offerte dans le cadre du projet IMERA (Cf. § VIII.6).

#### *Extension de IRVO à l'informatique ubiquitaire et pervasive*

IRVO a été conçu pour supporter la conception de systèmes de Réalité Mixte. Nous avons montré qu'il pouvait aussi être utilisé pour modéliser des interactions non standard comme l'interaction bi-manuelle avec des *toolglases*. Il serait donc intéressant de voir si d'autres domaines d'applications peuvent être envisagés et en particulier l'informatique ubiquitaire et pervasive. Les questions qui se posent sont en particulier de voir comment prendre en compte la mobilité des acteurs et des données ainsi que l'environnement, ou plus globalement le contexte de l'interaction. De plus, IRVO est centré autour de la notion de tâche ce qui pose le problème de la modélisation de « l'interaction opportuniste » mise en avant par l'*Ambient Intelligence* (AmI) où il n'y a pas de tâche *a priori*. Là encore nous comptons sur le projet IMERA pour nous offrir l'occasion d'explorer ces voies nouvelles et d'enrichir éventuellement le modèle IRVO.



# Bibliographie

*La bibliographie se fait après et non avant d'aborder un sujet de recherche.*  
Jean Perrin (1870 – 1942).

*The secret to creativity is knowing how to hide your sources.*  
Albert Einstein (1879 – 1955).

- [Ahlers 1995] **Ahlers K. H., Kramer A., Breen D.E., Chevalier P.Y., Crampton C., Rose E., Tuceryan M., Whitaker R.T., Greer D.** *Distributed Augmented Reality for Collaborative Design Applications*. ECRC Technical Report, ECRC-95-03, 1995. (Also appeared in Eurographics '95 Proceedings, pp. C-03 - C-14, September 1995). Available from Internet: <http://www.cs.iupui.edu/~tuceryan/AR/ECRC-95-03.ps.gz>
- [Allen 1983] **Allen, J. F.** Maintaining knowledge about temporal intervals. In *Communications of the ACM*, Vol. 26, n° 11, pp 832 - 843, November 1983.
- [Azuma 1993] **Azuma R.** Tracking Requirements for Augmented Reality. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 50-51.
- [Azuma 1997] **Azuma R.T.** A Survey of Augmented Reality. In *Presence: Teleoperators and Virtual Environments (Special Issue on Augmented Reality)*, August 1997, 6 (4), pp. 355-385.
- [Azuma 2001] **Azuma R., Baillot Y., Behringer R., Feiner S., Julier S. MacIntyre B.** Recent Advances in Augment Reality. In *IEEE Computer Graphics and Applications*. 21, 6, Nov/Dec 2001, pp 34-47.
- [Baecker 1993] **Baecker R.M., Nastos D., Posner I.R., Mawby K.L.** The User-Centred Iterative Design of Collaborative Writing Software. In *Proceedings of INTERCHI '93*, 1993, pp. 399-405, ISBN:0-89791-575-5.
- [Barfield 1995] **Barfield W., Rosenberg C. Lotens W.A.** Augmented-Reality Displays. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 542-575. ISBN 0-19-507555-2.
- [Baudel 1993] **Baudel T., Beaudouin-Lafon M.** Charade: Remote control of objects using free-hand gestures. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 28-35.
- [Baudel 1995] **Baudel T.** Réalité Augmentée : pour une définition et une nouvelle dénomination du domaine, 1995. Available from Internet : <http://www-ihm.lri.fr/~thomas/Docs/ra.html> [consulté le 27/01/2000]
- [Beaudouin-Lafon 2000a] **Beaudouin-Lafon M.** Instrumental Interaction: an Interaction Model for Designing Post-WIMP Interfaces. In *Proceedings of ACM Human Factors in Computing Systems, CHI'2000*, The Hague (Netherlands), April 2000, ACM Press, pages 446-453. Online: <http://www.daimi.aau.dk/CPnets/CPN2000/download/chi2000.pdf>
- [Beaudouin-Lafon 2000b] **Beaudouin-Lafon M. et Lassen H.M.** The Architecture and Implementation of CPN2000, a Post-WIMP Graphical Application. In *Proc. ACM Symposium on User Interface Software and Technology, UIST 2000, San Diego (USA)*, Novembre 2000, CHI Letters 2(2):181-190, ACM Press.
- [Beaudoux 2001] **Beaudoux O., Beaudouin-Lafon M.** DPI: A Conceptual Model Based on Documents and Interaction Instruments. In *People and Computers XV - Interaction without frontiers* (Joint proceedings of HCI 2001 and IHM 2001, Lille, France), pp 247-263, Springer Verlag, 2001.
- [Beldame 2001] **Beldame M.** *Projet ACCORÉA: Etude des apports d'un système de réalité augmentée au processus de conception*, Mémoire de DEA, Université du Maine, Le Mans, septembre 2001, 93 pages.

- [Berard 2000] **Bérard F., Coutaz J., Crowley J.L.** Le Tableau Magique. In *Actes de la conférence Ergonomie et Informatique Avancée - Interaction Homme-Machine (ERGO-IHM)*, Biarritz, France, 3-6 octobre 2000. CRT ILS & ESTIA, ISBN 2-9514772-0-1, pp 33-40
- [Bernsen 1994a] **Bernsen N. O.** Foundations of Multimodal representations. A taxonomy of representational modalities. In *Interacting with Computers*, vol. 6, n. 4, Elsevier, 1994, pp. 347-371.
- [Bernsen 1994b] **Bernsen N. O.** *A Revised Generation Of The Taxonomy Of Output Modalities*. TM/WP11, Amodeus-2 project, ESPRIT Basic Research Action 7040, 12 pages. On-line: [http://www.mrc-cbu.cam.ac.uk/amodeus/abstracts/tm/tm\\_wp11.html](http://www.mrc-cbu.cam.ac.uk/amodeus/abstracts/tm/tm_wp11.html)
- [Bernsen 1995] **Bernsen N. O.** *A taxonomy of input modalities*. TM/WP22, Amodeus-2 project, ESPRIT Basic Research Action 7040, 11 pages. On-line: [http://www.mrc-cbu.cam.ac.uk/amodeus/abstracts/tm/tm\\_wp22.html](http://www.mrc-cbu.cam.ac.uk/amodeus/abstracts/tm/tm_wp22.html)
- [Bianchi 2004] **Bianchi J.-P., Brusseau T., Clédat R., Michailovsky A., Sato M.** *Table Interactive pour le travail collaboratif (TableGate)*. Rapport final de projet de Projet d'étude promotion 2005, Ecole Centrale de Lyon, janvier 2004, 82 pages.
- [Bier 1993] **Bier E. A., Stone M., Pier K., Buxton W., DeRose T.** Toolglass and magic lenses: the see-through interface. In *Proceedings of SIGGRAPH '93*, pp. 73-80.
- [Billinghurst 2001a] **Billinghurst M., Kato H., Poupyrev I.** The MagicBook: a transitional AR interface. In *Computer & Graphics*, vol. 25, n. 5, October 2001, pp. 745-753, Elsevier.
- [Billinghurst 2001b] **Billinghurst M., Kato H., Poupyrev I.** The MagicBook: Moving Seamlessly between Reality and Virtuality. In *IEEE Computer Graphics and Applications*, May/June, 2001, pp. 2-4.
- [Bimber 2002] **Bimber O., Gatesy S. M., Witmer L. M., Raskar R., Encarnaçao L. M.** Merging Fossil Specimens with Computer-Generated Information. In *Computer*, vol. 35, No 9, September 2002, IEEE Computer Society, ISSN 0018-9162, pp. 25-30.
- [Bimber 2004] **Bimber O.** Combining Optical Holograms with Interactive Computer Graphics. In *Computer*, vol. 37, n. 1, January 2004, IEEE Computer Society, ISSN 0018-9162, pp. 85-91.
- [Bouchet 2002] **Bouchet J., Nigay L., Renevier P., Pasqualetti L.** Troc: un jeu collaboratif sur support mobile exploitant des techniques de réalité augmentée. Colloque sur la mobilité, GDR-I3, 6 décembre 2002, LORIA, Nancy. 2 pages.
- [Boud 1999] **Boud A.C., Haniff D.J., Baber C., Steiner S.J.** Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks. In *International Conference on Information Visualisation (IV'99)*, London, England, July 14 - 16, 1999, pp. 32-36.
- [Bourguin 2000] **Bourguin G.** *Un support informatique à l'activité coopérative fondé sur la Théorie de l'Activité : le projet DARE*. Thèse Informatique : Université des sciences et technologies de Lille, 2000, 215 pages.
- [Bowskill 1997] **Bowskill J., Morphett J., Downie J.** A taxonomy for Enhanced Reality Systems. In *Proceedings of the First International Symposium on Wearable Computers (ISWC '97)*, IEEE Computer Society, 1997.
- [Brun 1998] **Brun P.** *XTL : Une logique temporelle pour la spécification formelle des systèmes interactifs*. Thèse en informatique, Université de Paris XI Orsay. 1998, 179 pages.
- [Burbeck 1992] **Burbeck S.** *Applications Programming in Smalltalk-80(TM): How to use Model-View-Controller (MVC)*. Available from Internet : <http://st-www.cs.uiuc.edu/users/smarch/st-docs/mvc.html> [consulté le 11/06/2004]
- [Bush 1945]. **Bush V.** As We May Think. In *The Atlantic Monthly*. July, 1945, Vol. 176, No. 1, pp. 101-108. on-line: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm> .
- [Butz 2004] **Butz A., Groß M., Krüger A.** TUISTER: a Tangible UI for Hierarchical Structures. In *Proceedings of 2004 International Conference on Intelligent User Interfaces (IUI'04)*, Funchal, Madeira, Portugal, Jan 13-16, 2004, ACM Press, ISBN 1-58113-815-6, pp. 223-225.
- [Buxton 1983] **Buxton, W.** Lexical and Pragmatic Considerations of Input Structures. In *Computer Graphics*, 1983, vol. 17, n. 1, pp. 31-37.

- [Buxton 1992] **Buxton W.** Telepresence: integrating shared task and person spaces. In *Proceedings of Graphics Interface '92*, pp. 123-129.
- [Calvary 1997] **Calvary G., Coutaz J., Nigay L.** From single-user architectural design to PAC\*: a generic software architecture model for CSCW. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'97)*, Atlanta, USA, 1997, pp 242-249.
- [Carroll 1997] **Carroll J. M.** Scenario-Based Design. In Helander M., Landauer T.K., Prabhu P. (eds), *Handbook of Human-Computer Interaction*, Second, completely revised edition, Elsevier Science B.V., pp. 383-406.
- [Chalon 2001] **Chalon R., David B. T., Beldame M., Cherief N., Lasalle J., Moinard J.** L'oculomètre comme support d'évaluation et d'interaction. In *Actes de la conférence conjointe IHM-HCI 2001*, Lille, France, vol. II, Cépaduès, ISBN 2-85428-564-6, pp 117-122.
- [Chalon 2004a] **Chalon R., David B. T.** IRVO: an Architectural Model for Collaborative Interaction in Mixed Reality Environments. In *Proc. of the Workshop MIXER'04*, Funchal, Madeira, January 13, 2004, CEUR Workshop Proceedings, ISSN 1613-0073, online: CEUR-WS.org/Vol-91/paperD6.pdf
- [Chalon 2004b] **Chalon R. David B. T.** Modélisation de l'interaction collaborative dans les systèmes de Réalité Mixte. In *Actes de la 16<sup>ème</sup> conférence francophone sur l'Interaction Homme-Machine (IHM'04)*. Namur, 1-3 septembre 2004, ACM Press, ISBN 1-58113-926-8, pp. 37-44.
- [Charbonnier 1995] **Charbonnier C.** *La commande oculaire - Etude et validation expérimentale d'interfaces homme-machine contrôlées par la direction du regard*. Thèse de Doctorat, Université Joseph Fourier - Grenoble 1, 26 octobre 1995, 143 pages.
- [Choi 1999] **Choi H. I., Rhee P. K.** Head Gesture recognition using HMM. In *Expert Systems with Applications*, vol 17, Elsevier Science, 1999, pp. 213-221.
- [Cohen 1995] **Cohen M., Wenzel E. M.** The Design of Multidimensional Sound Interfaces. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 542-575. ISBN 0-19-507555-2.
- [Cohen 2004] **Cohen P. R., McGee D. R.** Tangible Multimodal Interfaces for Safety-Critical Applications. In *Communications of the ACM*, vol. 47, n. 1, January 2004.
- [Colombo 1997] **Colombo C., Del Bimbo A.** Interacting through eyes. In *Robotics and Autonomous Systems*, vol. 19, 1997, Elsevier, pp. 359-368.
- [Coutaz 2001] **Coutaz J., Nigay L.** Architecture logicielle conceptuelle des systèmes interactifs. In Kolski C. (Ed.), *Analyse et conception de l'IHM*. Hermès, Paris, chapitre 7, ISBN 2-7462-0239-5, pp. 207-246.
- [Crowley 1995] **Crowley J.L., Bérard F., Coutaz J.** Finger Tracking as an Input Device for Augmented Reality. In *Proceedings of the International Workshop on Face and Gesture Recognition*, Zurich, Switzerland, June 1995.
- [David 1996] **David B.T., Tarpin-Bernard F., Vial C.** Ergonomie du Travail coopératif en conception. In *Actes de la conférence ERGO'LA 1996*.
- [David 2001a] **David B. T.** IHM pour les collecticiels. In *Réseaux et Systèmes Répartis*, Hermès, Paris, vol. 13, novembre 2001, pp. 169-206.
- [David 2001b] **David B.T., Chalon R., Beldame M.** Oeil et IHM: suivi du regard et interaction "à l'oeil". In *Actes de la Conférence JIM 2001*, Metz, Juillet 2001.
- [David 2003a] **David B., Chalon R., Delotte O., Vaisman G.** Scenarios in the model-based process for design and evolution of cooperative applications. In Jacko J., Stephanidis C. (Eds.) *Human-Computer Interaction Theory and Practice*, Vol. 1, LEA, London, pp. 68-72.
- [David 2003b] **David B., Chalon R., Vaisman G., Delotte O.** Capillary CSCW. In Stephanidis C., Jacko J., (Eds.) *Human-Computer Interaction Theory and Practice* Vol. 2, LEA, London, pp. 879-883.
- [David 2003c] **David B., Chalon R., Delotte O., Ros J., Boutros N.** Travail coopératif capillaire en dépannage, maintenance et interventions de crise. In *Actes du 5<sup>ème</sup> Congrès International de Génie Industriel*, Québec, Canada. Octobre 2003.

- [David 2004] **David B., Delotte O., Seffah A., Tarpin-Bernard F., Desmarais M., Chalon R., Samaan K.** Scenarios in Design of Capillary Collaborative Systems. In *Proceedings of COOP04, 6th International Conference on the Design of Cooperative system*, May 11-14, 2004.
- [Delotte 2003] **Delotte O., David B., Chalon R.** Environnement et démarche pour la construction et l'évolution de systèmes collaboratifs. In *Actes de la conférence IHM 2003*, Caen, France, novembre 2003
- [Delotte 2004] **Delotte O., David B. T., Chalon R.** Task Modelling for Capillary Collaborative Systems based on Scenarios. In *Proceedings of Tamodia'2004 (3rd International Workshop on TASK MODELS and DIAGRAMS for user interface design)*, Prague, Czech Republic, November 15-16, 2004, pp. 25-31.
- [Demeure 2002] **Demeure A., Calvary G.** Jeu et Réalité Mixte : Retours d'expérience. In *Actes de IHM 2002*, 26-29 novembre 2002, Poitiers, France, ACM Press, ISBN : 1-58113-615-3, pp. 89-96.
- [Dewan 1998] **Dewan P.** Architectures for Collaborative Applications. In Beaudouin-Lafon M. (Ed.), *Collaborative Systems*, John Wiley & Sons Ltd, 1998.
- [Dourish 1999] **Dourish P.** *Embodied Interaction: Exploring the Foundations of a New Approach to HCI*. Unpublished paper, on-line: <http://www.ics.uci.edu/~jpd/publications/misc/embodied.pdf>.
- [Dubois 1999] **Dubois E., Nigay L., Troccaz J., Chavanon O., Carrat L.** Classification Space for Augmented Surgery, an Augmented Reality Case Study. In A. Sasse and C. Johnson (eds.), *Proceedings of Interact'99*, IOS Press, (1999), Edinburgh (UK), p. 353-359.
- [Dubois 2000] **Dubois E., Nigay L., Troccaz J.** Combinons le monde virtuel et le monde réel: classification et principes de conception. In *Actes des Premières Rencontres Jeunes Chercheurs en Interaction Homme-Machine (RJC-IHM)*, Ile de Berder, France, 3-5 mai 2000, pp. 31-34.
- [Dubois 2001a] **Dubois E., Nigay L., Troccaz J.** Consistency in Augmented Reality Systems. In *Proceedings of EHCI'01*, IFIP WG2.7 (13.2) Conference, Toronto, May 2001, LNCS 2254, Springer-Verlag. Available from Internet: [http://iuhm.imag.fr/pubs/2001/EHCI01\\_DuboisNT\\_ConsistencyInAR.pdf](http://iuhm.imag.fr/pubs/2001/EHCI01_DuboisNT_ConsistencyInAR.pdf)
- [Dubois 2001b] **Dubois E.** *Chirurgie Augmentée, un Cas de Réalité Augmentée; Conception et Réalisation Centrées sur l'Utilisateur*. Thèse de Doctorat, Université Joseph Fourier - Grenoble 1, 6 juillet 2001, 275 pages.
- [Dubois 2002a] **Dubois E., Gray P., Nigay L.** Asur++: A Design Notation for Mobile Mixed Systems. In Paternò F. (Ed.) *Proceedings of Mobile HCI 2002*, LNCS 2411, 2002, Springer-Verlag, pp. 123-139.
- [Dubois 2002b] **Dubois E.** Asur: un point de départ pour fédérer différents aspects de la conception d'un système interactif mobile. *Colloque sur la mobilité*, GDR-13, 6 décembre 2002, LORIA, Nancy. 2 pages.
- [Dubois 2002c] **Dubois E., Pinheiro da Silva P., Gray P.** Notational Support for the Design of Augmented Reality Systems. In Forbrig, P., Limbourg, Q., Urban, B., Vanderdonck, J. (eds.), *Proceedings of the 9th international workshop conference DSV-IS'02*, Rostock - Germany, June 2002, pp. 95-114.
- [Dubois 2003a] **Dubois E., Nigay L., Troccaz J.** Un Regard Unificateur sur la Réalité Augmentée : Classification et Eléments de Conception. In *Revue d'Interaction Homme-Machine*, vol. 4, n° 1, Europa, 2003, ISSN 1289-2963, pp.85-118.
- [Dubois 2003b] **Dubois E., Gray P., Nigay L.** ASUR++: Supporting the design of mobile mixed systems, In *Interacting with Computers*, Volume 15, Issue 4, 1 August 2003, Elsevier, pp 497-520. (<http://www.sciencedirect.com/science/article/B6V0D-48YW2SD-4/2/2366d79aa09ead5c4107e249d4408ac6>)
- [Dubois 2004a] **Dubois E., Gray P., Trevisan D., Vanderdonck J.** Workshop on Exploring the Design and Engineering of Mixed Reality Systems. In *Proceedings of 2004 International Conference on Intelligent User Interfaces (IUI'04)*, Funchal, Madeira, Portugal, Jan 13-16, 2004, ACM Press, ISBN 1-58113-815-6, pp. 374-375.

- [Dubois 2004b] **Dubois E., Mansoux B., Bach C., Scapin D., Masserey G. Viala J.** Un modèle préliminaire du domaine des systèmes mixte. In *Actes de la 16<sup>ème</sup> conférence francophone sur l'Interaction Homme-Machine (IHM'04)*. Namur, 1-3 septembre 2004, ACM Press, ISBN 1-58113-926-8, pp. 61-68.
- [Ellis 1991] **Ellis C., Gibbs S.J., Rein G.L.** Groupware: some issues and experiences. In *Communications of the ACM*, vol. 34, n° 1, Jan. 1991, pp. 38-58.
- [Ellis 1994] **Ellis C., Wainer J.** A conceptual model of Groupware. In *Proceedings of CSCW'94*, ACM Press, pp. 79-88.
- [Ellis 1997] **Ellis S.R., Begault D.R., Wenzel E.M.** Virtual Environments as Human-Computer Interfaces. In Helander M., Landauer T.K., Prabhu P. (eds), *Handbook of Human-Computer Interaction*, Second, completely revised edition, Elsevier Science B.V., pp. 163-201.
- [Engelbart Symposium 1998] *Engelbart's Unfinished Revolution*, A Symposium at Stanford University, December 9, 1998. on-line: <http://unrev.stanford.edu/index.html> [consulté le 05/04/2001]
- [Feiner 1993] **Feiner S., MacIntyre B., Seligmann D.** Knowledge-based Augmented Reality. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 52-62.
- [Feiner 1997] **Feiner S., MacIntyre B., Höllerer T., Webster, T.** A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In *Proceedings of the Int. Symp. on Wearable Computers (ISWC '97)*, Cambridge, USA, October 13-14, 1997. Available from Internet: <http://www.cs.columbia.edu/graphics/publications/iswc97.pdf> [consulté le 31/01/2001]
- [Fitzmaurice 1993] **Fitzmaurice G.** Situated Information Spaces: Spatially Aware Palmtop Computers. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 38-49.
- [Fjeld 1999] **Fjeld M., Voorhorst F., Bichsel M., Lauche K., Rauterberg M., Krueger H.** Exploring Brick-Based Navigation and Composition in an Augmented Reality. In H.-W. Gellersen (ed.) *Handheld and Ubiquitous Computing (HUC 99)*, Lecture Notes in Computer Science, Vol. 1707. Berlin/Heidelberg: Springer-Verlag, pp. 102-116.
- [Fjeld 2000] **Fjeld M., Voorhorst F., Bichsel M., Krueger H., Rauterberg M.** Navigation Methods for an Augmented Reality System. In *the video program / extended abstracts of CHI 2000*, pp. 8-9.
- [Fjeld 2002a] **Fjeld M., Lauche K., Bichsel M., Voorhorst F., Krueger H., Rauterberg M.** Physical and Virtual Tools: Activity Theory Applied to the Design of Groupware. In Nardi B. A. & Redmiles D. F. (eds.): *Activity Theory and the Practice of Design*, Computer Supported Cooperative Work (CSCW), vol 11, 2002, Kluwer Academic Publishers, pp. 153-180.
- [Fjeld 2002b] **Fjeld M., Voegtli B. M.** Augmented Chemistry: An Interactive Educational Workbench. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR'02)*, September 30 - October 01, 2002, Darmstadt, Germany, pp. 259-260.
- [Florins 2004] **Florins M., Trevisan D. G., Vanderdonckt J.** (2004) The continuity property in mixed reality and multiplatform systems: a comparative study. In R. Jacob et al. (eds), *Computer-Aided Design of User Interfaces IV, Proceedings of CADUI'2004*, Funchal, Madeira, Portugal, Jan 13-16, 2004, Kluwer, ISBN 1-4020-3145-9, pp. 323-334.
- [Furness 1995] **Furness III T.A., Barfield W.** Introduction to Virtual Environments and Advanced Interface Design. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 542-575. ISBN 0-19-507555-2.
- [Gaver 1991] **Gaver W. W.** Technology Affordances. In *Proceedings of CHI'91*, ACM, New York, pp. 79-84.
- [Goose 2003a] **Goose S., Sudarsky S., Zhang X., Navab N.** Speech-Enabled Augmented Reality Supporting Mobile Industrial Maintenance. In *Pervasive Computing*, January-March 2003, vol. 2, n° 1, IEEE Computing Society, pp. 65-70.

- [Goose 2003b] **Goose S., Guven S., Zhang X, Sudarsky S. Navab N.** Mobile 3D Visualization an Interaction in an Industrial Environment. In Stephanidis C. (Eds), *Universal Access in HCI, Inclusive Design in the Information Society*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 4, pp. 379-383.
- [Graham 2000] **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, In *Proceedings of DIS2000*, 17-19 August 2000, ACM, pp 406-416.
- [GRAM 1996] **Groupe de Recherche des Arts Médiatiques.** *Dictionnaire des arts médiatiques*, GRAM, Université du Québec à Montréal, 1996. on-line : <http://www.comm.uqam.ca/~GRAM/>
- [Hannaford 1995] **Hannaford B., Venema S.** Kinesthetic Displays for Remote and Virtual Environments. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 415-436. ISBN 0-19-507555-2.
- [Harper 1995] **Harper R., Sellen A.** *Paper-Supported Collaborative Work*. Technical Report EPC-1995-109, Cambridge:Rank Xerox Research Center, 1995, 16 p. Available from Internet: <http://www.unifr.ch/perso/corti/docs/epc-1995-109.pdf> [consulté le 05/04/2001]
- [Hasu 2000] **Hasu M., Engeström Y.** Measurement in action: an activity-theoretical perspective on producer–user interaction. In *International Journal of Human-Computer Studies*, Volume 53, Issue 1, Elsevier, July 2000, pp. 61-89.
- [Hightower 2001] **Hightower J., Boriello G.** Location Systems for Ubiquitous Computing. In *Computer*, 34(8), IEEE Computer Society, August 2001, pp. 57-66.
- [Höllerer 1999a] **Höllerer T., Feiner S., Pavlik J.** Situated Documentaries: Embedding Multimedia Presentations in the Real World. In *Proceedings of Third Int. Symp. on Wearable Computers (ISWC '99)*, San Francisco, USA, October 18-19, 1999, pp. 79-86. Available from Internet: <http://www.cs.columbia.edu/graphics/publications/iswc99.pdf> [consulté le 31/01/2001]
- [Höllerer 1999b] **Hollerer T., Feiner S., Terauchi T., Rashid G., Hallaway D.** Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. In *Computers & Graphics*, vol. 23, Elsevier Science Limited, 1999, pp. 779-785.
- [Hosoi 2003] **Hosoi Kazuhiro, Sugimoto Masanori.** Integrating Shared and Personal Spaces to Support Collaborative Learning. In Jacko J. & Stephanidis C. (Eds), *Human-Computer Interaction Theory and Practice*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 1, pp. 956-960.
- [Hutchins 1986] **Hutchins E. L., Hollan, J. D., Norman D.A.** Cognitive Engineering. In Norman D.A., Draper S. W. (Eds), *User Centered System Design. New Perspectives on Humann-Computer Interaction*, Lawrence Erlbaum Associates, 1986, chapter 5, pp. 87-124.
- [Hutchinson 1993] **Hutchinson T. E.** Eye-gaze computer interfaces: Computers that sense eye position on the display. In *Computer*, vol. 26, n° 7, July 1993, pp. 65,67.
- [Ishii 1997] **Ishii H., Ullmer B.** Tangible Bits: Towards Seamless Interfaces between People. Bits and Atoms. In *Proceedings of CHI'97*, Atlanta, 1997, p. 234-241.
- [Ishii 2002] **Ishii H., Ben-Joseph E., Underkoffler J., Yeung L., Chak D., Kanji Z., Piper B.** Augmented Urban Planning Workbench: Overlaying Drawings, Physical Models and Digital Simulation. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR'02)*, Darmstadt, Germany, September 30 - October 01, 2002, pp. 203-211.
- [Jacob 1993] **Jacob R.** Eye-gaze computer interfaces: What you look at is what you get. In *Computer*, vol. 26, n° 7, July 1993, pp. 65,67.
- [Jacob 1995] **Jacob R. J. K.** Eye-tracking in Advanced Interface Design. In Barfield W. & Furness III T. A. (Eds), *Virtual Environments and Advanced Interface Design*, Oxford University Press, 1995, ISBN 0-19-507555-2, pp. 258-288.
- [Jambon, 1996] **Jambon F.** *Erreurs et interruptions du point de vue de l'ingénierie de l'interaction Homme-Machine*. Thèse en informatique, Université Grenoble I, 1996, 228 pages.

- [Jambon 2001] **Jambon F., Brun P., Aït-Ameur Y.** Spécification des systèmes interactifs. In Kolski C. (Ed.), *Analyse et conception de l'IHM*. Hermès, Paris, chapitre 6, ISBN 2-7462-0239-5, pp. 175-206.
- [Johnson 1996] **Johnson C.W.** The Namur Principles: Criteria for the Evaluation of User Interface Notations. In Bodart F. and Vanderdonk J. (Eds) *Proceedings of the Design, Specification and Verification of Interactive Systems '96*, Springer Verlag, 1996.
- [Juras 2004] **Juras D.** *Modélisation des systèmes mixtes*. Examen probatoire en informatique du CNAM, Centre de Grenoble, 14 décembre 2004, 42 pages.
- [Kaczmarek 1995] **Kaczmarek K.A., Bach-Y-Rita P.** Tactile Displays. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 349-414. ISBN 0-19-507555-2.
- [Karsenty 1994] **Karsenty A.** Le collecticiel: de l'interaction homme-machine à la communication homme-homme. In *Technique et Science Informatique (TSI)*, vol. 13, n° 1, pp. 105-127, 1994.
- [Kato 2000] **Kato H., Billingham M., Poupirev I., Imamoto K., Tachibana K.** Virtual Object Manipulation on a Table-Top AR Environment. In *Proceedings of ISAR 2000*, October 5th-6th, 2000.
- [Kato 2003] **Kato H., Tachibana K., Nakajima T., Fukuda Y., Tananbe M., Cheok A. D.** Collaborative City-Planning System based on Augmented Reality. In Jacko J. & Stephanidis C. (Eds), *Human-Computer Interaction Theory and Practice*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 1, pp. 1153-1157.
- [Kim 1999] **Kim H., Kim J., Park S.** A Bird's-eye View System using Augmented Reality. In *Proceedings of the Thirty-Second Annual Simulation Symposium (SS)*, San Diego, USA, 11-15 April, 1999. IEEE, 1999.
- [Kiyokawa 1999] **Kiyokawa K., Takemura H., Yokoya N.** SeamlessDesign: A Face-to-face Collaborative Virtual/Augmented Environment for Rapid Prototyping of geometrically constrained 3-D Objects. In *Proceedings of the IEEE International Conference on Multimedia and Systems (ICMCS)*, volume II, 7-11 June 1999, Florence, Italy.
- [Klinker 1997] **Klinker G., Ahlers K. H., Breen D.E., Chevalier P.Y., Crampton C., Greer D., Koller D., Kramer A., Rose E., Tuceryan M., Whitaker R.T.** Confluence of Computer Vision and Interactive Graphics for Augmented Reality. In *Presence : Teleoperators and Virtual Environments (Special Issue on Augmented Reality)*, August 1997, 6 (4), pp. 433-451.
- [Kocian 1995] **Kocian D. F., Lee Task H.** Visually Coupled Systems Hardware and the Human Interface. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 175-257. ISBN 0-19-507555-2.
- [Kuzuoka 1994] **Kuzuoka H., Kosuge T., Tanaka M.** GestureCam: a Video Communication System for Sympathetic Remote Collaboration. In *Proceedings of CSCW'94*, Chapel Hill, USA, ACM, 1994, ISBN:0-89791-689-1, pp. 35 – 43.
- [Laurillau 2002a] **Laurillau Y., Nigay L.** Le Modèle d'Architecture Clover pour les Collecticiels. In *Proceedings of IHM 2002*, International Conference Proceedings Series, ACM, Poitiers, France, ISBN 1-58113-615-3/02/0011, pp. 113-120.
- [Laurillau 2002b] **Laurillau Y., Nigay L.** Clover architecture for groupware. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, New Orleans, USA, November 16 - 20, 2002, pp 236-245.
- [Laurillau 2002c] **Laurillau Y.** *Conception et réalisation logicielles pour les collecticiels centrées sur l'activité de groupe : le modèle et la plate-forme Clover*, Mémoire de thèse de l'Université Joseph Fourier, Grenoble, France, Septembre 2002.
- [Leigh 2002] **Leigh D., Dietz P.** DiamondTouch Characteristics and Capabilities. In *Workshop on Collaboration with Interactive Walls and Tables, UbiComp 2002*, September 29 - October 1, 2002, Göteborg, Sweden.
- [Lewis 2004] **Lewis J. R.** In the Eye of the Beholder. In *IEEE Spectrum*, May 2004. Available from Internet : <http://www.spectrum.ieee.org/WEBONLY/publicfeature/may04/0504reti.html>

- [Mackay 1993] **Mackay W., Velay G., Carter K., Ma C., Pagani D.** Augmenting Reality: Adding Computational Dimensions to Paper. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 96-97.
- [Mackay 1996] **Mackay W.E.** Réalité Augmentée : le meilleur des deux mondes. *La Recherche*, mars 1996, n° 285, pp. 32-37. Available from Internet [english translation]: <http://www-ihm.lri.fr/~mackay/psfiles/LaRecherche.English.ps>
- [Mackay 1997] **Mackay W. E., Fayard A.-L.** Radicalement nouveau et néanmoins familier : les strips papiers revus par la réalité augmentée. In *Actes des neuvièmes journées francophones sur l'Interaction Homme Machine (IHM'97)*, 10-12 septembre 1997, Poitiers.
- [Mackay 1998a] **Mackay W.E.** Augmented Reality: Linking real and virtual worlds - A new paradigm for interacting with computers. In *Proceedings of AVI'98, ACM Conference on Advanced Visual Interfaces, l'Aquila, Italy, May 1998*. New York : ACM Press, 1998. Available from Internet : <http://www-ihm.lri.fr/~mackay/psfiles/AVI98.AugmentedReality.ps>
- [Mackay 1998b] **Mackay W. E., Fayard A.-L., Frobert L., Médini L.** Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control. In *Proceedings of Conference on Human Factors in Computing Systems (CHI'98)*, April 18-23, 1998, Los Angeles.
- [Mackenzie 1995] **Mackenzie I. S.** Input Devices and Interaction Techniques for Advanced Computing. In *Virtual Environments and Advanced Interface Design*. Edited by Barfield W. and Furness III T.A. Oxford University Press, 1995. pp 437-470. ISBN 0-19-507555-2.
- [Mackinlay 1990] **Mackinlay, J. D., Card, S. K. and Robertson, G. G.** A Semantic Analysis of the Design Space of Input Devices. In *Human-Computer Interaction 5(2-3)*, Lawrence Erlbaum Associates, pp 145-190.
- [Mansoux 2004] **Mansoux B., Nigay L., Troccaz J.** Abstract and Concrete Interaction with Mixed Reality Systems / The case of the mini screen, a new interaction device in Computer-Assisted Surgery. In *Proc. of the Workshop MIXER'04*, Funchal, Madeira, January 13, 2004, CEUR Workshop Proceedings, ISSN 1613-0073, online: [CEUR-WS.org/Vol-91/paperD3.pdf](http://CEUR-WS.org/Vol-91/paperD3.pdf)
- [Martinez 2004] **Martinez M., Muñoz G.** Designing augmented interfaces for guided tours using multimedia sketches. In *Proc. of the Workshop MIXER'04*, Funchal, Madeira, January 13, 2004, CEUR Workshop Proceedings, ISSN 1613-0073, online: [CEUR-WS.org/Vol-91/paperD1.pdf](http://CEUR-WS.org/Vol-91/paperD1.pdf)
- [Milgram 1994a] **Milgram P., Kishino F.** A taxonomy of mixed reality visual displays. In *IEICE Transactions on Information Systems*, Vol E77-D, n° 12, December 1994. Available from Internet : [http://gypsy.rose.utoronto.ca/people/paul\\_dir/IEICE94/ieice.html](http://gypsy.rose.utoronto.ca/people/paul_dir/IEICE94/ieice.html)
- [Milgram 1994b] **Milgram P., Takemura H., Utsumi A., Kishino F.** Augmented Reality : a class of displays on the reality-virtuality continuum. In *SPIE, Telem manipulator and Telepresence Technologies*, Vol. 2351, 1994. Available from Internet : [http://gypsy.rose.utoronto.ca/people/paul\\_dir/SPIE94/SPIE94.full.html](http://gypsy.rose.utoronto.ca/people/paul_dir/SPIE94/SPIE94.full.html)
- [Milgram 1995] **Milgram P., Drascic D., Grodski J.J., Restogi A., Zhai S., Zhou C.** Merging Real and Virtual Worlds. In *Proceedings of IMAGINA'95*, Monte-Carlo, Monaco, 1-3 Feb. 1995, pp. 218-230. Available from Internet : [http://gypsy.rose.utoronto.ca/people/david\\_dir/IMAGINA95/Imagina95.full.html](http://gypsy.rose.utoronto.ca/people/david_dir/IMAGINA95/Imagina95.full.html) [consulté le 31/01/2001]
- [Millán 2002] **Millán J. del R.** Brain-Computer Interfaces. In M.A. Arbib (ed.) *Handbook of Brain Theory and Neural Networks*, Second edition, Cambridge: MIT Press, 2002.
- [Millán 2004] **Millán J. del R., Renkens F., Mouri no J., Gerstner W.** Brain-Actuated Interaction. In *Artificial Intelligence*, 2004.
- [Mori 2002] **Mori G., Paternò F., Santoro C.** CITE: Support for Developing and Analyzing Task Models for Interactive System Design. In *IEEE Transactions on software engineering*, vol. 28, n. 9, september 2002.

- [Nakanishi 2002] **Nakanishi Y., Sato Y., Koike H.** EnhancedDesk and EnhancedWall: Augmented Desk and Wall Interfaces with Real-Time Tracking of User's Motion. In *Workshop on Collaboration with Interactive Walls and Tables, UbiComp 2002*, September 29 - October 1, 2002, Göteborg, Sweden.
- [Nigay 2002] **Nigay L., Salembier P., Marchand T., Renevier P., Pasqualetti L.** Mobile and Collaborative Augmented Reality: A Scenario Based Design Approach. In Paternò F. (Ed.) *Proceedings of Mobile HCI 2002*, LNCS 2411, 2002, Springer-Verlag, pp. 241-255.
- [Nigay 2003] **Nigay L., Dubois E., Renevier P., Pasqualetti L., Troccaz J.** Mixed Systems: Combining Physical and Digital Worlds. In Jacko J. & Stephanidis C. (Eds), *Human-Computer Interaction Theory and Practice*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 1, pp. 1203-1207.
- [Noma 2003] **Noma H., Yanagida Y., Tetsutani N.** The Proactive Desk : A New Force Display System for a Digital Desk Using a 2-DOF Linear Induction Motor. In *Proceedings of the IEEE Virtual Reality 2003 (VR'03)*.
- [Norman 1986] **Norman D.A.** Cognitive Engineering. In Norman D.A., Draper S. W. (Eds), *User Centered System Design. New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, 1986, chapter 3, pp. 31-61.
- [OMG 2004] *OMG Model-Driven Architecture*. Home page. online: <http://www.omg.org/mda/> [consulté le 21/06/2004]
- [O'Neill 2003] **O'Neill K., de Paor A., MacLachlan M., McDarby G.** An Investigation into the performance of a Virtual Mirror Box for the treatment of Phantom Limb Pain in Amputees using Augmented Reality Technology. In Stephanidis C. (Eds), *Universal Access in HCI, Inclusive Design in the Information Society*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 4, pp. 236-240.
- [Ouadou 1994] **Ouadou K.** *AMF: un modèle d'architecture multi-agents multi-facettes pour Interfaces Homme-Machine et les outils associés*. Thèse de doctorat, Ecole Centrale de Lyon, 1994.
- [Paternò 2000] Paternò F. Model-Based Design and evaluation of Interactive Applications. Springer-Verlag, ISBN 1-85233-155-0.
- [Pierret-Golbreich 1989] **Pierret-Golbreich C., Delouis I., Scapin D. L.** *Un outil d'acquisition et de représentation des tâches orienté-objet*. Rapport de recherche INRIA n° 1063, août 1989, 34 pages. Online: <ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-1063.pdf>
- [Pinheiro da Silva 2000] **Pinheiro da Silva P.** User Interface Declarative Models and Development Environments: A Survey. In *Proceedings DSV-IS'2000*, LNCS, vol. 1946, pp. 207-226.
- [Poupyrev 2002] **Poupyrev I., Tan D.S., Billingham M., Kato H.** Developing a Generic Augmented-Reality Interface. In *Computer*, vol. 35, n. 3, IEEE, March 2002, pp. 44-50.
- [Rabardel 2003a] **Rabardel P., Waern Y.** Editorial. From artefact to instrument. In *Interacting with Computers*, Elsevier, vol. 15, 2003, pp. 641-645.
- [Rabardel 2003b] **Rabardel P., Bourmaud G.** From Computer to instrument system: a developmental perspective. In *Interacting with Computers*, Elsevier, vol. 15, 2003, pp. 665-691.
- [Rayner 1998] **Rayner K.** Eye Movements in Reading and Information Processing: 20 Years of Research. In *Psychological Bulletin*, vol. 124, No 3, 1998, pp. 372-422.
- [Rekimoto 1995] **Rekimoto J., Nagao K.** The World through the Computer: Computer Augmented Interaction with Real World Environments. In *Proceedings of User Interface Software and Technology*, 1995.
- [Rekimoto 2000] **Rekimoto J., Ayatsuka Y.** CyberCode: Designing Augmented Reality Environments with Visual Tags. In *Proceedings of DARE'2000*.
- [Renevier 2001] **Renevier P., Nigay L.** Mobile Collaborative Augmented Reality: the Augmented Stroll. In *Proceedings of EHCI'01*, IFIP WG2.7 (13.2) Conference, Toronto, May 2001, LNCS 2254, Springer-Verlag, pp. 315-334. Available from Internet: <http://iihm.imag.fr/publs/2001/ehci01-renevier.pdf>

- [Renevier 2004a] **Renevier P., Nigay L., Bouchet J., Pasqualetti L.** Generic interaction techniques for mobile collaborative mixed systems. In R. Jacob et al. (eds), *Computer-Aided Design of User Interfaces IV, Proceedings of CADUI'2004*, Funchal, Madeira, Portugal, Jan 13-16, 2004, Kluwer, ISBN 1-4020-3145-9, pp. 309-322. Publié aussi in *Proc. of the Workshop MIXER'04*, Funchal, Madeira, January 13, 2004, CEUR Workshop Proceedings, ISSN 1613-0073, online: CEUR-WS.org/Vol-91/paperD2.pdf
- [Renevier 2004b] **Renevier P., Nigay L.** Notation de Conception pour les Systèmes Mixtes Collaboratifs et Mobiles. In *Actes des Premières Journées Francophones : Mobilité et Ubiquité (UbiMob'04)*, 1-3 juin 2004, Nice, Sophia-Antipolis, pp. 66-73.
- [Renevier 2004c] **Renevier P.** *Systèmes Mixtes Collaboratifs sur Supports Mobiles : Conception et Réalisation*. Thèse de doctorat, Université Joseph Fourier – Grenoble I, 28 juin 2004. 220 pages.
- [Reynolds 1997] **Reynolds C.** A critical Examination of Separable User Interface Management Systems: Constructs for Individualization. In *SIGCHI Bulletin*, vol. 29, n° 3, July 1997.
- [Robinson 2001] **Robinson J. A., Robertson C.** The LivePaper system: augmenting paper on an enhanced tabletop. In *Computer & Graphics*, vol. 25, n. 5, October 2001, pp. 731-743, Elsevier.
- [Roussel 1999] **Roussel N., Nouvel G.** La main comme télépointeur. In IHM'99 - 11ème journées sur l'interaction Homme-Machine, Montpellier, 22-26 novembre 1999. Tome 2, pp. 33-36.
- [Saha 2003] **Saha D., Mukherjee A.** Pervasive Computing: A Paradigm for the 21st Century. In *Computer*, vol. 36, No 3, March 2003, IEEE Computer Society, ISSN 0018-9162, pp. 25-31.
- [Salber 1995] **Salber D., Coutaz J., Decouchant D. Riveill M.** De l'observabilité et de l'honnêteté dans la communication homme-homme médiatisée. In *Proceedings of Septièmes Journées sur l'Ingénierie de l'Interaction Homme-Machine IHM'95*, Toulouse, France, Toulouse: Cépaduès, pp. 27-33..
- [Samaan 2004] **Samaan K., Tarpin-Bernard F.** Task models an Interaction models in a Multiple User Interfaces generation process. In *Proceedings of Tamodia'2004 (3rd International Workshop on TASK MODELS and DIAGRAMS for user interface design)*, Prague, Czech Republic, November 15-16, 2004, pp. 137-144.
- [Santos 2003] **Santos P., Graf H., Flaisch T. Stork A.** 3D Interactive Augmented Reality in Early Stages of Product Design. In Stephanidis C. (Eds), *Universal Access in HCI, Inclusive Design in the Information Society*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 4, pp. 1203-1207.
- [Sato 2000] **Sato K., Lim Y.-K.** Physical Interaction and Multi-Aspect Representation for Information Intensive Environments. In *Proceedings of IEEE International Workshop on Robot and Human Interactive Communication*, September 27-29, 2000, Osaka, Japan, pp. 436-443.
- [Scapin 2001] **Scapin D. L. & Bastien J. M. C.** Analyse des tâches et aide ergonomique à la conception : l'approche MAD\*. In Kolski C. (Ed.), *Analyse et conception de l'IHM*. Hermès, Paris, chapitre 3, ISBN 2-7462-0239-5, pp.85-116.
- [Shen 2002] **Shen C., Vernier F., Forlines C., Lesh N.** Around the Table. In *Workshop on Collaboration with Interactive Walls and Tables, UbiComp 2002*, September 29 - October 1, 2002, Göteborg, Sweden.
- [Stapleton 2002] **Stapleton C., Hughes C., Moshell M., Micikevicius P., Altman M.** Applying Mixed Reality to Entertainment. In *Computer*, vol. 35, No 12, December 2002, IEEE Computer Society, ISSN 0018-9162, pp. 122-124.
- [Starner 1997] **Starner T., Mann S., Rhodes B., Levine J.** Augmented Reality Through Wearable Computing. In *Presence: Teleoperators and Virtual Environments*, vol. 6, n° 4, ISSN 1054-7460, pp. 386-398.
- [Stifelman 1996] **Stifelman L. J.** Augmenting Real-World Objects: A paper-based Audio Notebook. In *Proceedings of Conference companion on Human factors in computing systems*. Vancouver, Canada, April 13-18, 1996, ACM Press, ISBN: 0-89791-832-0, pp. 199-200.

- [Sundareswaran 2003] **Sundareswaran V., Wang Kenneth, Chen Steven, Behringer Reinhold, McGee Joshua, Tam Clement, Zahorik Pavel.** 3D Audio Augmented Reality: Implementation and Experiments. In *Proceedings of The Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*, October 07 - 10, 2003, Tokyo, Japan pp. 296-297.
- [Szalavari 1998] **Szalavari Z., Eckstein E., Gervautz M.** Collaborative Gaming in Augmented Reality. In *Proceedings of VRST'98*, Taipei, Taiwan, ACM, pp. 195-204.
- [Szekely 1996] **Szekely, P.** Retrospective and Challenges for Model-Based Interface Development. In *Proceedings of the 2nd International Workshop on Computer-Aided Design of User Interfaces*, (Vanderdonckt, J. Ed.). Namur University Press, Namur, 1996.
- [Tang 2001] **Tang J.C., Yankelovich N., Begole J., Van Kleek M., Li F., Bhalodia J.** ConNexus to awarenex: extending awareness to mobile users Conference on Human Factors in Computing Systems archive In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 2001)*, Seattle, USA, 2001, pp 221-228.
- [Tarpin-Bernard 1997] **Tarpin-Bernard F.** *Travail coopératif synchrone assisté par ordinateur : Approche AMF-C*, Thèse de doctorat, Ecole Centrale de Lyon, juillet 1997, n° 97-14, 158 pages
- [Tarpin-Bernard 1998] **Tarpin-Bernard F., David B. T., Primet P.** Frameworks and patterns for synchronous groupware : AMF-C approach. In *Proceedings of IFIP Working Conference on Engineering for HCI : EHCI'98*, Kluwer Academic Publishers, Greece, septembre 1998, pp. 225-241.
- [Tarpin-Bernard 1999] **Tarpin-Bernard F., David B. T.** AMF: un modèle d'architecture multi-agents multi-facettes. In *Technique et Science Informatiques*, Hermès, Paris, vol. 18, N° 5, mai 1999, pp. 555-586.
- [Terrenghi 2004] **Terrenghi L., Zimmermann A.** Tailored Audio Augmented Environments for Museums. In *Proceedings of 2004 International Conference on Intelligent User Interfaces (IUI'04)*, Funchal, Madeira, Portugal, Jan 13-16, 2004, ACM Press, ISBN 1-58113-815-6, pp. 334-336.
- [Thevenin 1999] **Thevenin D., Bérard F., Coutaz J.** Capture d'Inscriptions pour la Réalité Augmentée. In *IHM'99 - 11<sup>ème</sup> journées sur l'interaction Homme-Machine, Montpellier, 22-26 novembre 1999*. Toulouse : Cépaduès-Editions, 1999. pp. 174-1181. ISBN 2.85428.513.1.
- [Trevisan 2003] **Trevisan D., Vanderdonckt J., Macq B.** Model-Based Approach and Augmented Reality Systems. In Jacko J. & Stephanidis C. (Eds), *Human-Computer Interaction Theory and Practice*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 1, pp. 1273-1277.
- [Trevisan 2004] **Trevisan D., Vanderdonckt J., Macq B.** Designing Interaction Space for Mixed Reality Systems. In *Proc. of the Workshop MIXER'04*, Funchal, Madeira, January 13, 2004, CEUR Workshop Proceedings, ISSN 1613-0073, online: CEUR-WS.org/Vol-91/paperD5.pdf
- [UbiComp-WS 2002] **Collectif.** Workshop on Collaboration with Interactive Walls and Tables, Summary of Subgroup Work. *UbiComp 2002*, September 29 - October 1, 2002, Göteborg, Sweden. Available from Internet: <http://ipsi.fhg.de/ambiente/collabtablewallws/papers/>,
- [UIMS 1992] **Collectif.** The UIMS Workshop Tool Developers: A metamodel for the runtime architecture of an interactive system, In *SIGCHI bulletin*, vol. 24, n° 1, January 1992, p. 32-37.
- [Ullmer 2000] **Ullmer B., Ishii H.** Emerging Frameworks for tangible user interfaces. In *IBM Systems Journal*, vol. 39, n° 3&4, 2000, pp. 915-931. Available from Internet: <http://www.research.ibm.com/journal/sj/393/part3/ullmer.pdf> [consulté le 18/04/2001]
- [Umemuro 2003] **Umemuro H.** Collaboration Table: An Alternative Medium for Multi-user Multi-site Cooperation. In Stephanidis C. & Jacko J. (Eds), *Human-Computer Interaction Theory and Practice*, Proc. of HCI International, 22-27 June 2003, Crete, Greece, London : Lawrence Erlbaum Associates, 2003, vol. 2, pp. 1029-1033.
- [Vacherand-Revel 1992] **Vacherand-Revel J.** *Interaction cognitive et interface graphique. Le cas de la conception / réalisation de logiciels multimédias de formation*. Thèse de doctorat, Université Lumière - Lyon 2, 17 septembre 1992, 403 pages.

- [Vaisman 2002] **Vaisman G.** *TCAO Capillaire : apports du modèle AMF-C*. Mémoire de DEA, Ecole Centrale de Lyon. Juillet 2002. 41 pages.
- [Vallino 1999] **Vallino J., Brown C.** Haptics in Augmented Reality. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems (ICMCS)*, Florence, Italy, 7-11 June 1999, Volume 1, IEEE, 1999.
- [Vanderdonckt 1999] **Vanderdonckt J., Berquin P.** Towards a Very Large Model-Based Approach for User Interface Development. In *Proceedings of User Interfaces to Data Intensive Systems UIDS'99*, September 05-06, 1999, Edinburgh, Scotland, IEEE Computer Society, pp.76-85.
- [van der Veer 2001] **van der Veer G., van Vliet H.** The Human-Computer Interface *is* the System: A Plea for a Poor Man's HCI Component in Software Engineering Curricula. In Ramsey D., Bourque P. and Dupuis R. (Eds) *Proceedings 14th Conference on Software Engineering Education & Training*, February 19-21, 2001, Charlotte, NC, USA, IEEE, pp 276-286.
- [Vigouroux 1996] **Vigouroux N., Poirier F., Nigay L. (Ed.)** *Rapport du GT-M3 (GT5) : Modélisation de la MultiModalité*, 19 pages.
- [Webster 1996] **Webster A., Feiner S., MacIntyre B., Massie W., Krueger. T.** Augmented reality in architectural construction, inspection and renovation. In *Proceedings of the Third Congress on Computing in Civil Engineering (ASCE)*, Anaheim, USA, June 17-19, 1996, pp. 913-919. Available from Internet: <http://www.cs.columbia.edu/graphics/publications/asce.pdf> [consulté le 31/01/2001]
- [Weiser 1991] **Weiser M.** The Computer of the 21st Century, In *Scientific American*, vol. 265, no. 3, sept. 1991, pp. 66-75.
- [Weiser 1993] **Weiser M.** Some Computer Science issues in Ubiquitous Computing. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 74-84.
- [Wellner 1993a] **Wellner P., Mackay W., Gold R. (Ed.)** Computer Augmented Environments: Back to the Real World. Special Issue of *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 24-26.
- [Wellner 1993b] **Wellner P.** Interacting with Paper on the DigitalDesk. In *Communications of the ACM*, vol. 36, n° 7, July 1993, pp 86-96. Available from Internet: <ftp://ftp.parc.xerox.com/pub/euoparc/reports/ddesk-cacm93.ps>.

Total : 192 références bibliographiques.



## Réalité Mixte et Travail Collaboratif : IRVO, un modèle de l'Interaction Homme-Machine

L'un des principaux reproches que l'on fait à l'informatique traditionnelle est que le monde numérique de l'ordinateur est trop déconnecté du monde réel dans lequel les utilisateurs évoluent. La Réalité Mixte (RM) s'intéresse aux systèmes interactifs dans lesquels des objets réels et des données informatiques sont mêlées de manière cohérente.

Dans notre état de l'art nous analysons de nombreuses applications et technologies qui ont été conçues depuis une dizaine d'années dans ce domaine de façon généralement indépendantes et autonomes. Il semble désormais possible de proposer des modèles unificateurs et des méthodologies de conception associées.

Les modèles traditionnels de l'Interaction Homme-Machine (IHM) et du Travail Collaboratif Assisté par Ordinateur (TCAO), comme nous le montrons, ne sont pas adaptés car ils sont conçus pour la modélisation du logiciel et ne prennent pas (ou peu) en compte les artefacts réels.

Nous proposons donc le modèle IRVO (Interacting with Real and Virtual Objects) qui cherche à formaliser l'interaction entre les utilisateurs et les objets réels et virtuels. Pour cela IRVO identifie les entités intervenant dans le système (les utilisateurs, les outils, les objets de la tâche et l'application informatique) et leurs relations. Son pouvoir d'expression nous permet de modéliser les applications de la littérature dans le but de les comprendre, les comparer et les classer.

Nous montrons également que IRVO est un support à la conception d'applications nouvelles : il peut être utilisé soit comme point de départ, soit en fin d'analyse pour aider le concepteur à vérifier un certain nombre de critères (complétude, cohérence, continuité, etc.). Nous étudions la liaison entre IRVO et les modèles classiques de l'IHM et du TCAO et plus spécifiquement avec les modèles AMF-C et Arch. Dans une démarche MBA (Model-Based Approach), IRVO prend naturellement la place aux côtés des modèles de tâches et de dialogues.

**Mots-clés :** Interaction Homme-Machine, Travail Collaboratif Assisté par Ordinateur, Réalité Mixte, Réalité Augmentée, Modèle d'interaction.

---

### Mixed Reality and collaborative work: IRVO, a Human-Computer Interaction model

One of the main weaknesses with traditional computing is the fact that the numerical world of the computer is decoupled from the users' real world. The application of Mixed Reality (MR) can provide interactive systems in which real objects and computer data are combined in a cohesive way.

In our state of the art, we review several applications and technologies that have been developed during the last decade. Most of these applications are home made and were developed in an independent and autonomous way. Therefore, it is important to propose comprehensive model and an associated design methodology.

As we illustrated it, the current models of Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW), are not suitable for Mixed Reality systems. It is because they are limited in terms of modeling the real artefacts.

Therefore, the proposed IRVO model (Interacting with Real and Virtual Objects) aims at formalizing the interaction between the users, the real and virtual objects. It identifies the entities that have a role in the system (users, tools, objects of the task and the application) and their relationships. Its expressiveness help us to model the applications we reviewed while helping us in understanding, comparing and classifying them.

We will also show how IRVO supports the design of a new generation of Mixed Reality based applications. IRVO can be used either as a starting point, or at the end of the analysis phase to help the designer in checking the completeness, coherence, continuity of the design. We investigated also the relation between IRVO and traditional HCI and CSCW models and more specifically with the AMF-C and Arch models. In a Model-Based Approach (MBA), IRVO can supplement the task and dialog models.

**Keywords:** Human-Computer Interaction, Computer Supported Collaborative Work, Mixed Reality, Augmented Reality, interaction model.

---