

# Modèle d'Interaction Mixte : la Réalité Mixte à la Lumière des Modalités d'Interaction

Céline Coutrix, Laurence Nigay, Philippe Renevier

Équipe IHM, CLIPS-IMAG

385, rue de la Bibliothèque, BP 53 38041 Grenoble Cedex 9

+33 4 76 51 44 40

Celine.Coutrix@imag.fr, Laurence.Nigay@imag.fr, Philippe.Renevier@imag.fr

## RÉSUMÉ

Malgré la multiplicité des études et démonstrateurs de réalité mixte, le cadre de compréhension des systèmes mariant les mondes numérique et physique reste flou. Dans cet article, nous proposons d'adopter un point de vue unificateur sur la réalité mixte en considérant les modalités d'interaction mises en jeu par le système, focalisant ainsi sur l'interaction. Nous présentons un nouveau modèle d'interaction, noté modèle de l'interaction mixte. Comme tout modèle d'interaction, nous étudions son pouvoir descriptif mais aussi génératif de nouvelles formes d'interaction mixte.

## Mots Clés

Réalité Mixte, Modèle d'Interaction, Modalité d'Interaction, Multimodalité.

## ABSTRACT

Mixed Reality seeks to smoothly link the physical and data processing (digital) environments. In recent years, mixed reality has been the subject of growing interest. Although mixed reality systems are becoming more prevalent, we still do not have a clear understanding of this interaction paradigm. This article introduces a new interaction model called Mixed Interaction model. It adopts a unifying point of view on mixed reality by considering the interaction modalities that are involved for defining mixed environments. This article presents the model after explaining its foundations, applies it to describe and compare existing systems and shows its generative power for designing new forms of mixed interaction.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – User-centered design, theory and methods.

D.2.2 [Software Engineering]: Design Tools and Techniques – User Interfaces.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UBIMOB 05, May 31. - June 3, 2005, Grenoble, France. Copyright 2005 ACM X-XXXXX-XXX-X/XX/XXXX \$5.00.

## General Terms

Design; Human Factors.

## Keywords

Mixed Reality, Interaction Model, Interaction Modality, Multimodality.

## 1. INTRODUCTION

Le paradigme de réalité mixte en Interaction Homme-Machine (IHM) repose sur la fusion des mondes physique et numérique, affranchissant ainsi les utilisateurs des barrières séparant les deux mondes. Le monde ainsi obtenu est qualifié de *mixte*. Depuis une quinzaine d'années et la présentation du Digital Desk[14], de nombreux démonstrateurs et systèmes de réalité mixte ont été développés. Néanmoins la multiplicité des systèmes ne s'est pas accompagnée d'avancée conceptuelle : ainsi nous constatons une prolifération de termes et de modèles pour décrire la réalité mixte qui traduit sans nul doute des points de vue différents et une absence de consensus. Dans cet article, nous proposons d'adopter un point de vue unificateur sur la réalité mixte en considérant les modalités d'interaction mises en jeu par le système focalisant ainsi sur l'interaction. Nous proposons un nouveau modèle noté modèle de l'interaction mixte. Comme tout modèle d'interaction [2], nous étudions son pouvoir descriptif mais aussi génératif de nouvelles formes d'interaction mixte

Après avoir exposé les travaux à l'origine de notre modèle d'interaction, nous présentons deux de nos systèmes de réalité mixte qui appuieront les illustrations du modèle proposé. Nous expliquons ensuite notre modèle d'interaction mixte avant d'examiner son pouvoir de description et de classification ainsi que ses capacités génératives.

## 2. FONDEMENTS

Plusieurs travaux sont à l'origine de notre modèle d'interaction mixte. Ces travaux s'inscrivent dans différents axes de recherche de l'IHM, tels que la Réalité mixte, l'Interaction Instrumentale et la Multimodalité. Nous rappelons d'abord notre classification des systèmes de Réalité Mixte [5] puis nous exposons deux modèles d'interaction qui mettent en exergue les notions d'objet et outil de la tâche pour finir par des définitions des termes modalité d'interaction et multimodalité.

## 2.1 Espace de classification des systèmes de réalité mixte

Dans [5], nous proposons une classification des systèmes de réalité mixte en considérant deux caractéristiques intrinsèques à la tâche-utilisateur : l'objet de la tâche et le type d'augmentation.

Nous considérons d'abord l'objet de la tâche : si cet objet appartient au monde réel, alors il s'agit d'un système de Réalité Augmentée (RA). A l'inverse, si l'objet de la tâche est numérique, alors il s'agit de Virtualité Augmentée (VA). De notre distinction entre objet de la tâche physique et numérique, deux types de tâches sont clairement identifiés : dans un cas une entité physique est au centre de l'interaction, dans l'autre c'est une entité numérique. Pour affiner notre distinction entre RA et VA, dans [5] nous proposons deux continua traduisant le degré d'intrusion d'un monde dans l'autre : le degré d'intrusion du monde numérique dans le monde réel pour la RA et respectivement le degré d'intrusion du monde réel dans le monde numérique pour la VA. Les deux continua, RA et VA distinguent des systèmes qui respectivement augmentent soit l'interaction avec le monde réel grâce à l'ordinateur, soit l'interaction avec l'ordinateur en exploitant des éléments du monde réel. Nous nous intéressons maintenant aux types d'augmentation de l'interaction fournis par le système, deuxième dimension de classification.

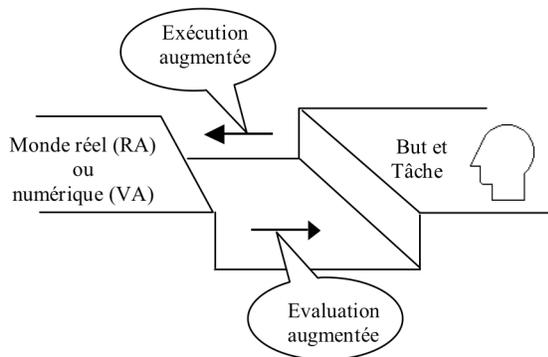


Figure 1. Classification des systèmes de Réalité Mixte selon l'objet de la tâche (RA et VA), et le type d'augmentation (exécution et évaluation). Figure traduite de [5].

La seconde dimension de classification, orthogonale à la première, distingue l'exécution augmentée de l'évaluation augmentée. Nous reprenons ici les termes d'exécution et d'évaluation issus de la Théorie de l'Action [11]. Avec l'exemple d'un système de RA, l'exécution augmentée consiste à permettre à l'utilisateur d'effectuer une tâche dont l'objet cible est physique par des moyens qui n'existent pas dans le monde réel. L'évaluation augmentée, quant à elle, consiste à ajouter des informations numériques ancrées dans le monde réel : l'évaluation se trouve aidée par l'informatique. Adaptant un schéma de [11] au cas de la réalité augmentée, la Figure 1 résume les deux dimensions de notre classification [5].

## 2.2 Objet et outil de la tâche

Plusieurs approches ont distingué l'objet de l'outil dans la réalisation d'une tâche. La première présentée ici est ASUR (Adaptateur, Système, Utilisateur, entités Réelles), une notation de conception pour la réalité mixte que nous avons introduite dans [5], détaillée dans [6] et qui a été affinée dans [8]. Selon ASUR,

l'interaction avec un système de réalité mixte est modélisée par 6 composants mis en relation :

- l'utilisateur U ;
- le système S, comprenant les éléments du système utiles à l'interaction en entrée  $S_{\text{outil}}$  et les éléments du système utiles à la perception de l'état interne du système  $S_{\text{presentation}}$  (dont l'objet de la tâche  $S_{\text{objet}}$  et ses attributs  $S_{\text{info}}$ );
- les objets réels R, comprenant les outils  $R_{\text{outil}}$  et les objets de la tâche  $R_{\text{objet}}$  ;
- les adaptateurs A, comprenant les adaptateurs en entrée  $A_{\text{in}}$  et en sortie  $A_{\text{out}}$ , qui transfèrent les données d'un monde à l'autre.

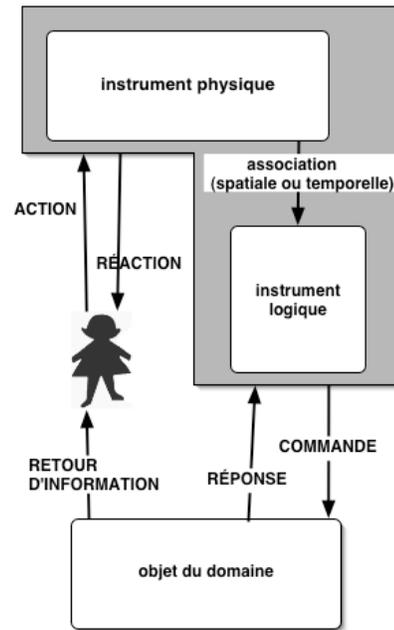


Figure 2. Modèle instrumental de l'interaction. Figure traduite de [1][2].

La notation ASUR repose donc sur la distinction entre objet et outil de la tâche. Cette distinction a été exploitée dans [4] pour étendre la classification du paragraphe précédent en considérant outre l'objet de la tâche, l'outil comme axes de classification. Nous retrouvons cette distinction entre objet et outil dans le modèle d'interaction instrumentale [1][2]. Ce modèle présente un cadre de compréhension de l'interaction entre un utilisateur et une interface graphique. Nous le schématisons dans la figure 2. Il y a trois pôles dans ce modèle de l'interaction : l'utilisateur, l'objet du domaine sur lequel se porte l'attention de l'utilisateur, et l'instrument. L'utilisateur agit sur l'instrument. Son action sur l'instrument est transformée en commande par l'instrument, la commande étant ciblée sur l'objet du domaine. Cet objet modifie en conséquence son état interne et en informe l'utilisateur par un retour d'information et l'instrument par une réponse. L'instrument lui-même informe l'utilisateur de son état par une réaction, et permet ainsi à l'utilisateur de contrôler son action. De plus, le modèle d'interaction instrumentale distingue la partie

logique de l'instrument (sa représentation dans le système informatique, par exemple une barre de défilement) de la partie physique (le dispositif physique, par exemple une souris).

### 2.3 Modalité et multimodalité

La notion d'instrument du modèle d'interaction exposée ci-dessus conditionne comment l'utilisateur interagit avec un système informatique. Cette notion est donc proche de celle de modalité d'interaction.

Dans [10], nous définissons une modalité d'interaction comme un couple  $(d,l)$ , avec  $d$  un dispositif d'interaction et  $l$  un langage d'interaction. Une modalité  $(d,l)$  couvre les distances articulatoires et partiellement sémantiques à la fois d'exécution (modalité en entrée du système) et d'évaluation (modalité en sortie du système) de la Théorie de l'Action [11]. Par exemple la manipulation directe au sein des interfaces graphiques est décrite par le couple  $(souris, Manipulation\ d'objets\ graphiques)$ . De même la modalité parole se décrit par le couple  $(Microphone, Langage\ pseudo\ Naturel)$ .

Basé sur cette définition d'une modalité d'interaction et dans le cadre de l'interaction multimodale, il convient d'étudier la composition de modalités. Dans [13] nous déclinons ces compositions de modalités selon cinq aspects. Les compositions *spatiales* décrivent les relations dans l'espace (lieu d'interaction et lieu de perception). La composition *temporelle* est aussi considérée. Elle caractérise les enchaînements possibles de modalités. La composition *articulatoire* des modalités dénote une composition à un niveau d'abstraction plus bas, entre les flux d'information transitant par les dispositifs associés aux modalités. La composition *syntactique* considère les relations entre les langages d'interaction utilisés. Enfin, la composition *sémantique* concerne la composition des informations véhiculées par les modalités.

Les approches exposées dans ce paragraphe, concernant la réalité mixte, l'interaction instrumentale, et la multimodalité, servent de fondement à notre modèle d'interaction mixte. Avant de le présenter, nous décrivons deux systèmes de réalité mixte.

## 3. APPLICATIONS EXEMPLES

Afin d'illustrer notre modèle d'interaction mixte qui fait l'objet du paragraphe suivant, nous exposons au préalable deux systèmes de réalité mixte.

### 3.1 Razzle

Dans la continuité de nos systèmes TROC [9] et MAGIC [12], le jeu Razzle (Réalité Augmentée puZZLE), est un système de réalité augmentée mobile et collaboratif. Par équipes, les utilisateurs du jeu se déplacent à l'intérieur ou à l'extérieur, à la recherche des pièces dispersées d'un puzzle numérique. Comme le montre la Figure 3, les utilisateurs sont équipés d'un capteur de position et d'orientation, et d'un casque de vision semi transparent, où des pièces numériques, augmentées d'une position dans l'espace, sont affichées. Ainsi les pièces numériques du puzzle augmentent l'environnement réel de l'utilisateur mobile. L'utilisateur ramasse une pièce en s'arrêtant à proximité et en regardant la pièce. Une fois saisie, la pièce du puzzle numérique n'est plus affichée dans l'environnement. Ainsi, la recherche des

pièces nécessite l'implication physique de l'utilisateur : il n'est plus dans un univers virtuel comme dans un jeu vidéo classique.

Razzle est un système de VA, puisque l'objet de la tâche de l'utilisateur est la collecte de pièces de puzzle numériques. L'évaluation est dite augmentée, puisque l'utilisateur perçoit les pièces du puzzle ancrées dans le monde réel. L'exécution est aussi augmentée puisque l'utilisateur recherche les pièces du puzzle en se déplaçant et peut les ramasser simplement en les regardant.



Figure 3 : Un enfant jouant avec Razzle lors de la fête de la science en octobre 2004.

### 3.2 CASPER

Le système CASPER [5] (Computer ASsisted PERicardial puncture) est un système de chirurgie assistée par ordinateur. La tâche du chirurgien (utilisateur de CASPER) consiste en une ponction péricardiale. Pour cette opération, le chirurgien fait une incision dans la poitrine juste assez grande pour introduire l'aiguille de ponction vers le fluide en excès près du cœur. Une trajectoire de ponction optimale est planifiée avant l'opération à partir d'images échographiques du cœur. En temps réel pendant l'opération, le chirurgien voit la position courante de l'aiguille par rapport à la trajectoire optimale planifiée. Un système de localisation (caméra et diodes attachées à l'aiguille) permet de connaître la position courante de l'aiguille, selon la partie émergée de l'aiguille. Plusieurs versions du système ont été développées : dans la première, les positions courantes et prévues sont affichées sur un écran comme le montre la Figure 4 ; dans une autre version du système, la trajectoire prévue est affichée dans un casque à vision semi transparent. Enfin une troisième version de CASPER consiste à augmenter l'aiguille d'un bras nommé PADyC : PADyC n'agit pas de manière autonome avec l'aiguille sur le patient, mais permet au chirurgien de percevoir la trajectoire à suivre, grâce à un retour d'effort.

Dans les trois versions de CASPER, le chirurgien effectue simultanément sa tâche (une opération), et est aidé dans la correction de la trajectoire par le système. Aussi CASPER est un système de RA, puisque l'objet de la tâche est dans le monde réel. Dans la version 2 de CASPER avec un casque, le champ opératoire est augmenté de la trajectoire planifiée. L'évaluation est donc augmentée. De même dans la version 3 avec retour d'effort, l'aiguille de ponction est augmentée : l'évaluation est donc aussi augmentée.



Figure 4 : Le système CASPER avec affichage sur un écran d'ordinateur.

Les deux systèmes Razzle (VA) et CASPER (RA) sont utilisés pour illustrer notre modèle d'interaction mixte

#### 4. MODÈLE D'INTERACTION MIXTE

Notre modèle d'interaction repose sur la notion d'objet mixte que nous définissons au paragraphe suivant. Un objet mixte étant défini, nous modélisons ensuite les techniques d'interaction mixte.

##### 4.1 Objet mixte

Un objet physique se caractérise par un ensemble de propriétés. Ces propriétés sont perceptibles par les cinq sens humains. Un objet numérique, de façon similaire, contient un ensemble de propriétés, qui peuvent être stockées et traitées en machine. Un objet mixte est un objet qui appartient aux deux mondes : il est composé d'une partie physique et d'une partie numérique comme le schématise la Figure 5. Nous reprenons la définition de [4] : *Un objet mixte est défini par le lien entre un objet physique et un objet numérique.* Notre modèle vise à définir ce lien qui unit ces deux parties en une entité mixte.

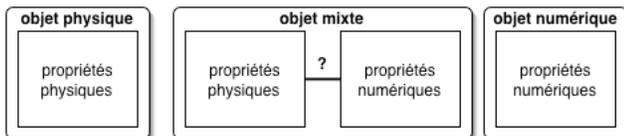


Figure 5. Objet physique, objet mixte et objet numérique.

La partie physique de l'objet mixte comprend un ensemble de propriétés, appelées *propriétés physiques* de l'objet mixte. De même, la partie numérique de l'objet mixte comprend un ensemble de propriétés, appelées *propriétés numériques* de l'objet mixte. En considérant la définition d'une modalité d'interaction comme un couple  $(d, l)$ , l'association entre ces deux ensembles de propriétés peut ensuite se définir de la façon suivante :

- En entrée du système, une modalité de liaison  $(d_o^e, l_o^e)$  permet de :
  1. capter certaines données parmi les *propriétés physiques*, à l'aide d'un dispositif  $d_o^e$ , dispositif en entrée de l'objet,
  2. puis traduire ces *données physiques captées* en *propriétés numériques*, grâce à un langage  $l_o^e$ .
- En sortie du système, une autre modalité de liaison permet de :

1. générer certaines données à partir des *propriétés numériques*, à l'aide d'un langage  $l_o^s$ ,
2. rendre ces *données physiques générées* perceptibles par des *propriétés physiques*, grâce à un dispositif  $d_o^s$ .

Le lien entre les propriétés physiques et numériques d'un objet mixte peut donc se définir comme une modalité en entrée et une modalité en sortie. La Figure 6 représente cette définition de l'objet mixte.

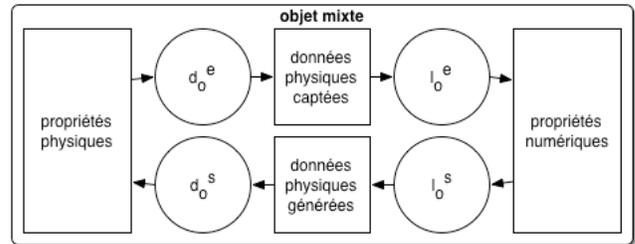


Figure 6. Objet Mixte.

Il convient de remarquer que la modélisation autorise plusieurs niveaux d'abstraction. Le choix du niveau d'abstraction est laissé au concepteur. La souris par exemple peut être modélisée de deux façons différentes. Elle peut être objet mixte, alors sa modalité de liaison en entrée est la composition de deux modalités, chacune ayant un dispositif différent (les boutons et le capteur de localisation). Mais la souris peut aussi être modélisée comme un dispositif d'entrée  $d_o^e$ , comme c'est généralement le cas. Le modèle permet donc au concepteur d'adapter le niveau d'abstraction à ses besoins.

Pour illustrer notre définition d'un objet mixte, considérons les différentes versions de CASPER.

L'aiguille est un objet mixte. Parmi ses *propriétés physiques*, nous trouvons sa position dans l'espace en trois dimensions. Ses *propriétés numériques* sont la position dans le référentiel. Entre ces deux ensembles de propriétés, nous définissons deux modalités. En entrée du système, le dispositif  $d_o^e$  est l'ensemble caméras et diodes, les *données physiques captées* sont les chiffres bruts. Le langage  $l_o^e$  transforme les données brutes en leur donnant un sens relatif au référentiel utilisé. En sortie du système, dans la version avec le bras PADyC attaché à l'aiguille, la position de l'aiguille mixte par rapport à la trajectoire pré-calculée (*propriétés numériques*) est interprétée par le langage  $l_o^s$  pour générer un ordre de commande pour le bras, le bras  $d_o^s$  rendant tangible la trajectoire.

La Figure 6 présente le cas le plus général d'un objet mixte. En effet, il convient de noter que la modélisation d'un objet mixte peut comporter deux parties physiques distinctes. Par exemple dans le cas de la version de CASPER de la Figure 4, où les informations de guidage incluant la position de l'aiguille sont affichées sur un écran, l'aiguille de ponction est un objet mixte qui comporte deux parties physiques, l'une l'outil physique (aiguille) associé à une modalité d'entrée  $(d_o^e, l_o^e)$ , l'autre les pixels à l'écran associés à une modalité de sortie  $(d_o^s, l_o^s)$ , où  $d_o^s$  désigne l'écran. De plus, la modélisation d'un objet mixte peut ne comporter qu'une seule modalité en entrée ou en sortie.

Considérons par exemple la version de CASPER avec un casque à vision semi transparent dans lequel n'est affichée que la trajectoire pré-calculée. L'aiguille de ponction est alors un objet mixte augmenté qu'en "entrée", sa modélisation ne comportant qu'une modalité d'entrée ( $d_o^e, l_o^e$ ).

Une fois la notion d'objet mixte modélisée, nous pouvons en apprécier le pouvoir descriptif. Comme nous l'avons démontré avec les trois versions de CASPER, le modèle autorise une compréhension plus fine de l'augmentation, parce qu'il intègre des "modalités de liaison" entre les parties physiques et numériques. Ces modalités de liaison explicitent deux niveaux d'abstraction entre les parties physiques et numériques d'un objet mixte. Nous examinons maintenant comment interagir avec un tel objet mixte.

## 4.2 Modalité d'interaction mixte

Pour analyser l'interaction, nous généralisons le modèle de l'interaction instrumentale [1][2] aux systèmes mixtes en exploitant notre définition d'une modalité d'interaction. Dans le cas de l'interaction en entrée, selon le modèle d'interaction instrumentale, nous considérons donc un instrument ou outil et un objet de la tâche. L'outil est nécessairement lié au monde physique, car l'être humain interagit dans le monde physique. L'outil est donc un objet mixte, que nous notons outil d'interaction.

(1) Si l'objet de la tâche est physique, l'interaction reste donc dans le monde physique et la partie numérique de l'outil mixte (Figure 6) est exploitée par le système comme une donnée et non une technique d'interaction. C'est le cas de CASPER avec l'aiguille comme objet mixte.

(2) Si l'objet de la tâche est numérique ou mixte, nous appliquons alors le modèle d'interaction instrumentale présenté au paragraphe 2.2.

### 4.2.1 Outil d'interaction

Dans le cas général, l'outil est alors composé de deux parties :

- un outil mixte, objet mixte, comme une souris ou la position du regard dans RAZZLE,
- un outil logique, objet numérique, composé uniquement de propriétés numériques.

L'utilisateur fait une *action* modifiant les propriétés physiques de l'objet mixte, et celui-ci lui permet d'ajuster son *action* grâce à sa *réaction*. Les propriétés numériques de l'objet mixte sont interprétées par un langage d'interaction de l'outil en entrée  $l_o^e$  pour mettre à jour les propriétés numériques de l'outil logique. Les propriétés numériques de l'outil logique sont traduites en une tâche élémentaire par un langage de l'interaction en entrée  $l_i^e$ , et celle-ci s'applique à l'objet de la tâche. L'état interne de l'objet de la tâche est alors modifié, et celui-ci en rend compte à l'utilisateur par un retour d'information et à l'outil mixte par une réponse. Cette réponse peut éventuellement participer à l'élaboration par le système de la réaction de l'outil mixte. La Figure 7 schématise cette interaction avec un objet de la tâche (pour l'instant indifféremment mixte ou numérique), selon une modalité d'interaction en entrée modélisée comme un couple associant un outil mixte à un outil logique.

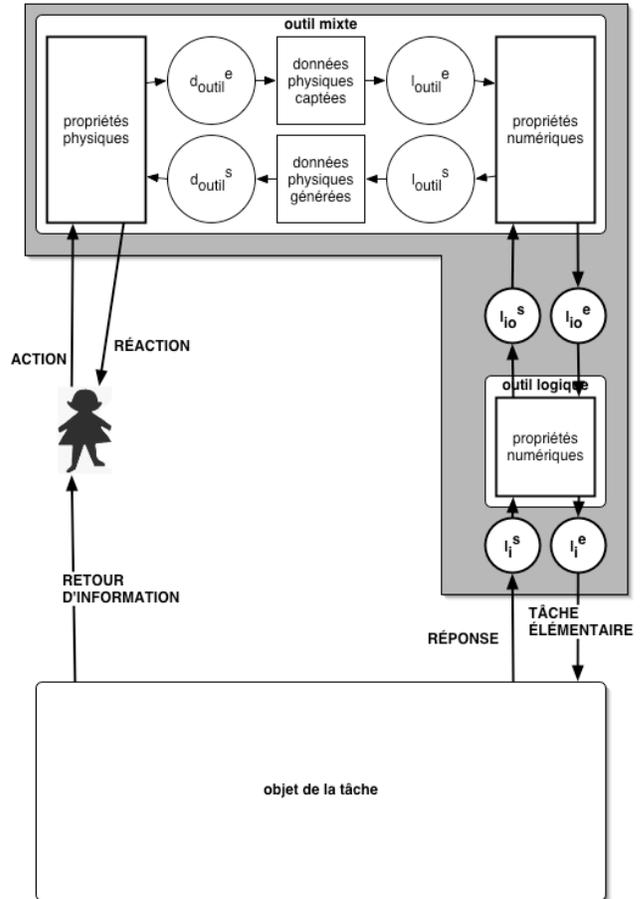


Figure 7. Outil d'interaction : outil mixte, outil logique.

Ce cas général permet de modéliser les interactions du type présenté dans [1][2] : un outil mixte peut activer différents outils logiques. Nous avons néanmoins identifié un cas particulier de l'outil d'interaction : lorsque l'outil mixte est dédié, c'est-à-dire qu'il ne sert qu'à une unique tâche élémentaire, alors il est inutile de l'associer au même outil logique dans le modèle. En effet, cela revient à un langage  $l_{io}$  neutre et à des propriétés numériques de l'outil logique qui sont la réplique exacte des propriétés numériques de l'outil mixte. Par exemple, une molette dédiée au réglage de la hauteur du son est un outil mixte dédié : en agissant sur ses propriétés physiques, on modifie ses propriétés numériques et celles-ci agissent directement sur le volume du son, objet de la tâche. L'intermédiaire est alors inutile.

Il convient de remarquer que ce modèle, même si ce n'est pas son propos, comprend également les modalités d'interaction avec les systèmes informatiques graphiques. Reprenons l'exemple de la barre de défilement issu de [1]. En entrée, l'outil mixte qu'est la souris a certaines propriétés physiques, captées par les capteurs de pression sous les boutons et le capteur de position,  $d_o^e$ . Ces données sont interprétées et deviennent une position et des valeurs booléennes, grâce au langage de la souris  $l_o^e$ , et enfin, à leur tour, ces propriétés numériques sont interprétées en propriétés numériques de la barre de défilement par le langage  $l_{io}^e$ . L'état interne de la barre est traduit en tâche élémentaire à appliquer au document, objet de la tâche. En sortie, les propriétés numériques

de la barre de défilement sont interprétées et la *réaction* de l'outil d'interaction mixte, outre la réaction physique de la souris, mettra en surbrillance la flèche sur l'écran. La *réponse* de l'objet de la tâche sera le décalage de la vue du document. Cette réponse participera, par l'intermédiaire de l'outil logique, à la *réaction* de l'instrument, par la mise à jour de l'ascenseur de la barre à l'écran. La souris est donc modélisée comme un outil mixte, celui-ci comprenant deux parties physiques disjointes : la souris objet physique et le curseur affiché à l'écran. Cet outil mixte correspond donc au cas de l'aiguille dans CASPER lorsque sa position par rapport à la trajectoire pré-calculée est affichée sur l'écran (Figure 4). Dans [7], nous avons défini la notion de continuité perceptuelle en étudiant la relation spatiale entre l'outil mixte, l'aiguille affichée à l'écran et l'objet de la tâche. Nous pouvons rapprocher cette notion de continuité perceptuelle de celle de degré d'indirection dans le modèle d'interaction instrumentale [1] : la relation spatiale entre l'outil logique et l'objet de la tâche y est présentée comme une propriété de l'outil.

Considérons maintenant le système Razzle, un système de VA où l'objet de la tâche, la pièce du puzzle, est mixte. La modélisation de la Figure 7 s'applique donc. Cette application est présentée à la Figure 10. La technique d'interaction avec la pièce de puzzle pour le ramasser est composée d'un outil mixte et d'un outil logique. Les *propriétés physiques* de l'outil mixte comprennent la position et l'orientation du joueur. Celles-ci sont captées par les capteurs de position et d'orientation, et les données ainsi obtenues sont interprétées par le langage qui leur donne un sens dans le référentiel du jeu. Ces coordonnées de position et d'orientation font ainsi partie des *propriétés numériques* de l'outil. Le langage d'interaction de l'outil  $l_{io}^e$  traduit ces propriétés en temps d'immobilité accumulé. Ces *propriétés numériques* sont ensuite interprétées en *tâche élémentaire* « ramasser la pièce ». L'objet de la tâche répond par un *retour d'information* à l'attention de l'utilisateur : il disparaît du casque de vision semi transparent. Nous notons que dans la version actuelle, il n'y a ni *réponse* vers l'outil logique, ni *réaction* de l'outil mixte. Nous pourrions envisager une autre version de Razzle où l'utilisateur déplace un curseur dans le casque de vision semi transparent pour l'amener sur la pièce de puzzle. L'outil mixte est alors différent et comprend une partie en sortie avec le casque comme dispositif  $d_{outil}^s$ .

Avec l'exemple de Razzle, nous soulignons aussi qu'une technique d'interaction n'implique pas forcément la manipulation d'un artefact. Le regard comme dans Razzle ou un geste non instrumenté sont des exemples de techniques d'interaction sans artefact. Ce cas, certes aux limites du modèle d'interaction instrumentale, est donc intégré au modèle que nous proposons.

Parmi les cas aux limites du modèle, nous trouvons aussi l'interaction « directe », c'est-à-dire l'action de l'utilisateur sur l'objet de la tâche sans outil mixte intermédiaire. Quand l'utilisateur manipule l'objet directement, nous le modélisons par un langage d'interaction  $l_i^e$  qui agit sur l'objet mixte qui est aussi outil mixte.

#### 4.2.2 Objet de la tâche mixte

Après avoir focalisé sur l'outil mixte (Figure 7), nous considérons maintenant l'objet de la tâche. Pour qu'il y ait un retour d'information de l'objet vers l'utilisateur, comme le montre la Figure 7, il est nécessaire que l'objet de la tâche soit mixte, c'est-à-dire qu'il ait un langage interprétant ses *propriétés numériques* en *propriétés physiques*, via le dispositif de l'objet en sortie. Si

l'objet n'est pas augmenté en sortie, c'est-à-dire si l'objet est purement numérique, son état n'est pas observable. Seules les *propriétés physiques* sont perceptibles par l'utilisateur. La Figure 9 étend la modélisation de la Figure 7 avec un objet de la tâche mixte.

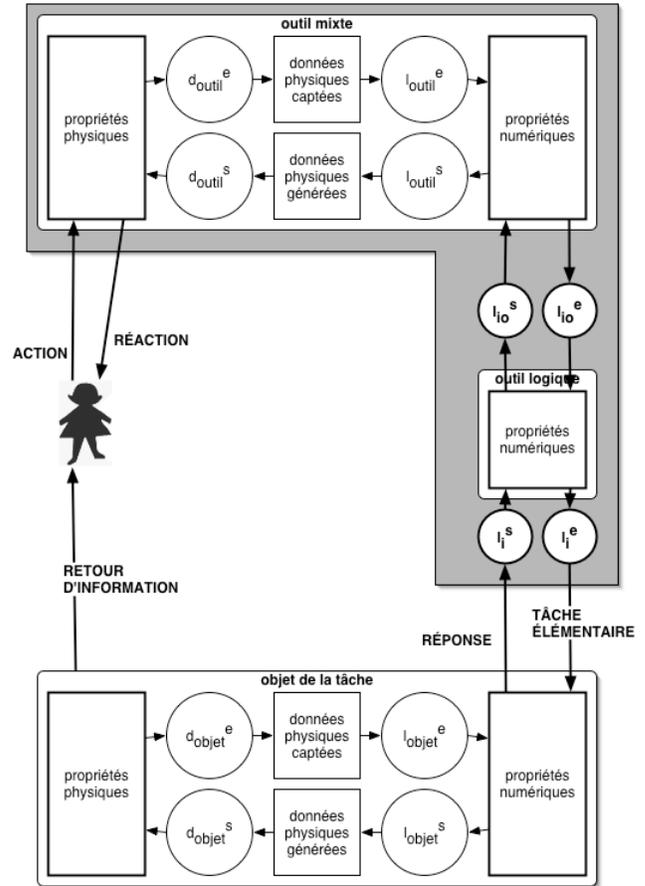


Figure 9. Technique d'interaction mixte sur un objet mixte.

Appliquant le modèle de la Figure 9 au cas de Razzle, nous affinons à la Figure 10 la modélisation partielle précédente. L'objet mixte de la tâche est la pièce de puzzle. Commençons par décrire l'objet de la tâche selon le modèle : ses *propriétés physiques* sont la position dans l'espace et l'orientation. Ses *propriétés numériques* sont (1) les coordonnées de position et d'orientation de la pièce dans le référentiel lié à l'utilisateur, (2) l'attribut booléen précisant si elle a été ramassée, (3) l'image numérique qui lui est associée et (4) sa place dans le puzzle. Entre ces deux ensembles de propriétés, nous trouvons deux modalités : En entrée du système, le dispositif  $d_{objet}^e$  est le capteur de position et d'orientation, les *données physiques captées* sont 3 données brutes, 2 pour la position et 1 pour l'angle. Le langage  $l_{objet}^e$  transforme les deux premières données en abscisse et ordonnée relatives au référentiel utilisé, et la dernière en orientation. En sortie du système, les *propriétés numériques* de la pièce mixte sont interprétées par le langage  $l_{objet}^s$  pour générer une image à la bonne taille, distance et orientation. Cette image est ensuite rendue observable grâce au casque à vision semi transparent.

Alors quand le langage  $l_i^e$  a généré la tâche élémentaire « ramasser la pièce », cette tâche est effectuée sur les *propriétés numériques* de la pièce mixte : le booléen précise que la pièce est ramassée. Le langage  $l_{objet}^s$  ne génère plus d'image localisée, c'est-à-dire une image vide, rendue observable grâce au casque à vision semi transparent. L'observabilité est soulignée, grâce au langage  $l_{objet}^s$ , par l'affichage en couleur de la pièce ramassée dans une zone dédiée dans le casque. La pièce était grise dans cette zone auparavant. L'utilisateur conclut alors que la pièce a été ramassée : c'est le retour d'information.

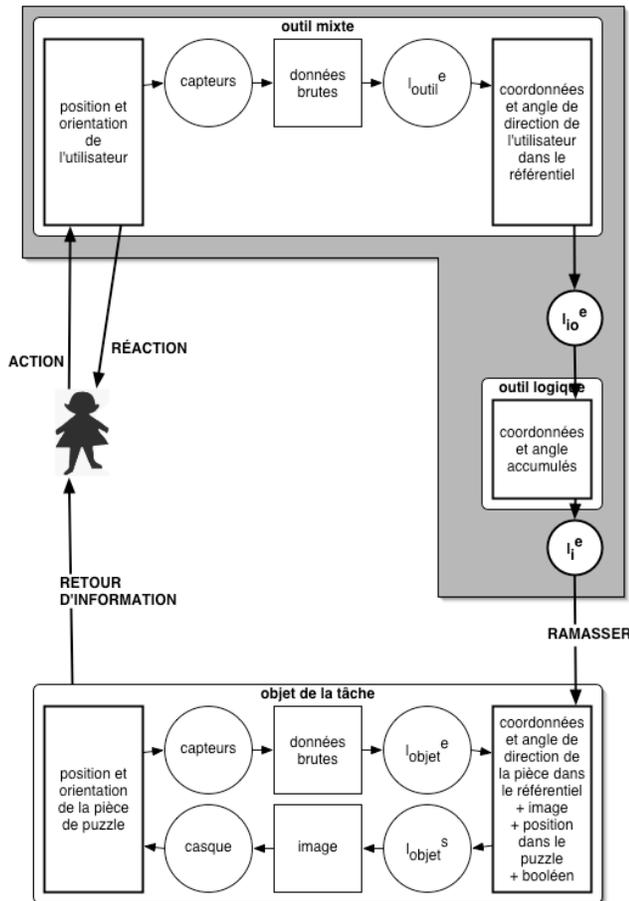


Figure 10. Application du modèle au ramassage d'une pièce mixte de Razzle.

Il est essentiel qu'un modèle d'interaction ait un fort pouvoir descriptif pour faciliter la compréhension de l'interaction. Outre ce pouvoir descriptif, il convient d'analyser le pouvoir comparatif du modèle par rapport aux taxinomies existantes ainsi que son pouvoir génératif de nouvelles formes d'interaction. Ces deux points font l'objet du paragraphe suivant.

## 5. DISCUSSION

### 5.1 Pouvoir taxinomique

Nous comparons notre modèle d'interaction mixte avec la taxinomie présentée au paragraphe 2.1.

#### 5.1.1 Réalité augmentée, virtualité augmentée

La distinction entre la RA et la VA repose sur le type de l'objet de la tâche. Cette distinction est immédiate dans notre modèle si l'objet de la tâche est soit réel (RA), soit numérique (VA).

Dans le cas d'un objet mixte, un système de RA reposera sur une tâche purement physique (interaction physique entre l'outil et l'objet de la tâche). Au contraire un système de VA contiendra dans sa modélisation une tâche numérique et donc un langage  $l_i^e$  pour l'exprimer.

Par exemple dans CASPER (RA), la réalisation de la tâche « faire la ponction » repose sur une interaction physique entre l'outil mixte qu'est l'aiguille de ponction et l'objet de la tâche, le patient. Ce dernier peut être augmenté (comme avec un casque de vision transparent) et sera donc modélisé comme un objet de la tâche mixte. Dans la modélisation de CASPER, aucun langage  $l_i^e$  ne véhiculera la tâche "faire la ponction". Au contraire dans Razzle (VA), comme le montre la Figure 10, un langage  $l_i^e$  véhicule l'expression de la tâche « ramasser la pièce de puzzle ».

#### 5.1.2 Exécution augmentée, évaluation augmentée

L'appréciation du deuxième axe de classification, l'exécution/évaluation augmentée est affinée par notre modèle d'interaction mixte. En effet, le modèle autorise une caractérisation de l'exécution/évaluation fine par l'identification de plusieurs langages qui définissent l'augmentation.

De plus par l'application du modèle d'interaction instrumentale à la réalité mixte, nous identifions plusieurs facettes à l'évaluation :

- l'évaluation de l'état interne de l'objet de la tâche ( $l_{objet}^s$ ),
- l'évaluation de l'état interne de l'outil mixte ( $l_{outil}^s$ ),
- l'évaluation du contrôle de l'outil mixte sur l'outil logique ( $l_{io}^s$ ),
- l'évaluation du contrôle de l'outil logique sur l'objet de la tâche ( $l_i^s$ ).

L'évaluation de l'état interne de l'outil logique passe, quant à lui, par les deux langages  $l_{io}^s$  et  $l_{outil}^s$ . Ces quatre facettes de l'évaluation affinent donc l'évaluation augmentée mais fournit aussi un cadre utile à la conception. Par exemple pour Razzle, nous avons constaté lors de sa modélisation que l'évaluation du contrôle de l'outil mixte sur l'outil logique n'est pas augmentée. En effet, l'utilisateur n'a aucun moyen de savoir le temps passé à fixer du regard la pièce du puzzle. Une solution serait de rendre observable le temps d'immobilité du regard (sorte de compte à rebours) qui constitue l'instrument, par exemple par une couleur variable affichée autour de la pièce de puzzle.

## 5.2 Pouvoir génératif

Comme nous l'avons expliqué au paragraphe précédent, notre modèle permet de raisonner à un niveau fin de description de l'interaction. Ainsi l'absence de certaines dimensions du modèle dans la description d'un système sert de support à la conception. Dans Razzle, nous avons par exemple constaté l'absence d'évaluation du contrôle de l'outil mixte sur l'outil logique, permettant de concevoir des extensions à la technique d'interaction initialement conçue. Le modèle par ses dimensions

sert donc de support à la conception de techniques d'interaction mixte.

De plus notre modèle d'interaction mixte, s'appuyant sur des résultats issus de travaux sur la multimodalité, nous incite naturellement à étudier les techniques d'interaction sous le regard de la fusion de plusieurs modalités. Ainsi en appliquant notre modèle, nous avons conçu deux techniques d'interaction dans Razzle qui s'appuient sur la fusion de deux modalités.

- (1) La première technique implique l'ajout d'un dispositif à la ceinture du joueur. Ce dispositif s'apparente à une souris et constitue un capteur de pression. Nous combinons cette modalité d'interaction avec la position et orientation du joueur par rapport à une pièce du puzzle pour déterminer si la pièce doit être ramassée.
- (2) La deuxième technique repose sur l'ajout d'une modalité analogique [3]. L'utilisateur ramasse une pièce du puzzle par un geste naturel comme pour attraper un objet physique. La capture et l'interprétation du geste est là encore combinée avec sa position relative à celle de la pièce du puzzle.

La conception de ces deux techniques, reposant sur la fusion de modalités, nous pousse à analyser plus en détail les aspects de fusion au sein de notre modèle d'interaction mixte. De plus il convient d'enrichir notre modèle de capacité d'évaluation prédictive afin de guider la conception. Pour cela, nous sommes en cours d'évaluation expérimentale de Razzle. Outre les aspects liés à réalité mixte que nous comparons avec une version de Razzle entièrement physique (pièces physiques du puzzle) et une version de Razzle entièrement numérique (jeu de Razzle en réalité virtuelle), nous étudions aussi comment l'utilisateur ramasse les pièces du puzzle dans la version mixte de Razzle. Trois techniques lui sont offertes, celles développées et modélisées à la Figure 10, et les deux nouvelles techniques conçues. La technique du ramassage par un geste naturel n'a pas été développée : elle est simulée lors des expérimentations par un compère humain (magicien d'oz) qui observe les joueurs. Plusieurs journées de tests ont été menées et nous sommes en cours de dépouillement des données.

## 6. CONCLUSION

Nous avons présenté un modèle de l'interaction mixte, qui tire parti des approches concernant la réalité mixte d'un côté, et la multimodalité de l'autre côté. Ces deux perspectives n'avaient pas été étudiées conjointement jusqu'à aujourd'hui. Nous avons montré que la précision de notre modélisation induit une meilleure compréhension et classification des systèmes existants, mais aussi que le modèle permet de concevoir de nouvelles techniques d'interaction mixte. Les perspectives qu'offre cette capitalisation de plusieurs approches n'ont pas encore été toutes explorées, notamment l'étude de la fusion de modalités et l'utilité du modèle pour la comparaison et l'évaluation de l'utilisabilité (souplesse et robustesse) de l'interaction.

## 7. REFERENCES

- [1] Beaudoin-Lafon M., *Instrumental Interaction : An Interaction Model for Designing Post-WIMP User Interfaces*, CHI'00, CHI Letters, p.446-453.
- [2] Beaudoin-Lafon M., *Designing Interaction, not Interfaces*, AVI'04, May 25-28, 2004, Gallipoli (LE), Italy, p.15-22.
- [3] Bernsen N. O., *Taxonomy of HCI Systems : State of the Art*, ESPRIT BR GRACE, deliverable 2.1, 1993.
- [4] Chalon R., *Réalité mixte et Travail Collaboratif : IRVO, un modèle de l'interaction Homme-Machine*, Thèse de doctorat, décembre 2004.
- [5] Dubois E., Nigay L., Troccaz J., Chavanon O., Carrat L., *Classification space for augmented surgery, an augmented reality case study*, Acte de la conférence INTERACT'99, Chapman and Hall, 1999, p.353-359.
- [6] Dubois E., *Chirurgie Augmentée, un Cas de Réalité Augmentée ; Conception et Réalisation Centrées sur l'Utilisateur*, Thèse de Doctorat, Juillet 2001.
- [7] Dubois E., Nigay L., Troccaz J., *Consistency in Augmented Reality Systems*, Proceedings of EHCI'01, Working Conference, Toronto, Lecture Notes in Computer Science (LNCS), Vol. 2254, M. Reed Little, L. Nigay Eds, Springer-Verlag Publ., May 2001. p.117-130
- [8] Dubois E., Mansoux B., Bach C., *Un modèle préliminaire du domaine des systèmes mixtes*, IHM 04, ACM Press, September 2004, p. 61-68
- [9] Nigay L., Dubois E., Renevier P., Pasqualetti L., Troccaz J., *Mixed Systems: Combining physical and digital worlds*, Conference proceeding of HCI International 03, Theory and Practice (Part I), Vol 1, Edited by Julie Jacko, Constantine Stephanidis, LEA, June 2003. p.1203-1207
- [10] Nigay L., Coutaz J., *A Generic Platform for Addressing the Multimodal Challenge Proceedings*, CHI'95, Denver, May 1995.
- [11] Norman D. (1986), *Cognitive Engineering*, in D.A. Norman & S. W. Draper (eds.), *User Centered System Design, New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, p.31-61.
- [12] Renevier P., Nigay L., *Mobile Collaborative Augmented Reality: the Augmented Stroll*, Acte de la conférence EHCI'01, IFIP WG2.7 (13.2) Working Conference, Toronto, Lecture Notes in Computer Science (LNCS), Vol. 2254, M. Reed Little, L. Nigay Eds, Springer-Verlag Publ. May 2001. p.315-334
- [13] Vernier F., Nigay L., *A Framework for the Combination and Characterization of Output Modalities*, Actes de la conférence DSV-IS'00 (Limerick, Ireland), Lecture Notes in Computer Sciences, Springer-Verlag Publ., 2000, p.32-48.
- [14] Wellner P., *The Digital Desk Calculator : tangible manipulation on a desk top display*, Acte de la quatrième conférence annuelle ACM Symposium on User Interface Software and Technology, Novembre 1991, p.27-33.