

# Le Modèle d'Interaction Mixte : Un cadre pour la conception des systèmes mixtes

Céline Coutrix, Laurence Nigay

CLIPS-IMAG

CLIPS-IMAG Laboratory, University of Grenoble 1, BP 53, 38041 Grenoble Cedex 9, France

+33 4 76 51 44 40

celine.coutrix@imag.fr, laurence.nigay@imag.fr

## RÉSUMÉ

La réalité mixte marie les mondes physique et numérique. Malgré la multiplicité des études et démonstrateurs de réalité mixte, le cadre de compréhension de ces systèmes reste flou. Dans cet article, nous proposons d'adopter un point de vue unificateur sur la réalité mixte en considérant les modalités d'interaction et la multimodalité mises en jeu par le système mixte, focalisant ainsi sur l'interaction. Nous présentons un nouveau modèle d'interaction, appelé modèle d'interaction mixte. Cet article présente le modèle et ses fondements. Comme tout modèle d'interaction, nous étudions ensuite ses pouvoirs génératif et comparatif en les appliquant à la conception de techniques d'interaction dans RAZZLE, un jeu mobile en réalité mixte dans lequel le joueur cherche des pièces de puzzle d'origine numérique dispersées sur le terrain de jeu.

## Categories and Subject Descriptors

D.2.2 [Software Engineering]: Design Tools and Techniques - *User interfaces*. H.5.2 [Information Interfaces And Presentation] User Interfaces - *Interaction styles, User-centered design*.

## General Terms

Design, Human Factors.

## Mots-clés

Réalité mixte, Réalité/Virtualité augmentée, Modèle d'interaction, Modèle instrumental, Multimodalité, Modalité d'interaction.

## 1. INTRODUCTION

Les systèmes mixtes marient les monde physique et numérique. Malgré la multiplicité des études et démonstrateurs de réalité mixte, le cadre de compréhension de ces systèmes reste flou. Historiquement, beaucoup de systèmes mixtes superposaient des informations visuelles par dessus l'environnement physique. C'est pourquoi une des premières taxonomies proposées définissait un continuum réalité-virtualité pour classer les dispositifs d'affichage des systèmes mixtes. Pourtant, la conception et la réalisation de la fusion des mondes physiques et numériques peuvent prendre en compte d'autres modalités d'interaction que des modalités

visuelles, comme le montre la recherche dans le domaine des interfaces tangibles. Puisqu'elle repose sur un nouveau paradigme d'interaction, la conception des systèmes mixtes nécessite de nouveaux modèles d'interaction pour permettre la conception d'objet mixte en explorant tout l'espace des possibilités. De plus, les objets prenant part à l'interaction entre l'utilisateur et le système mixte peuvent être physiques, mixtes ou numériques, et avoir des rôles différents. Pour pouvoir imaginer de nouvelles modalités d'interaction et de nouvelles formes de multimodalité pour interagir avec de tels objets, le concepteur a donc besoin de nouveaux modèles d'interaction.

Un modèle d'interaction a pour but de fournir un cadre de travail pour la conception de systèmes interactifs [1]. Un tel modèle se caractérise par trois dimensions [1] : son pouvoir descriptif permet de mieux comprendre les systèmes déjà existants, son pouvoir génératif permet de créer des techniques d'interaction en explorant l'espace de conception de façon systématique, et le pouvoir comparatif d'un modèle permet de choisir entre deux techniques.

Après avoir présenté un exemple de système mixte que nous avons développé et appelé RAZZLE (Réalité Augmenté puzzle), nous présenterons la définition du modèle d'interaction. Un tel modèle doit être capable de décrire des systèmes mixtes tels que RAZZLE, donc nous illustrerons sa définition avec RAZZLE. Enfin nous présenterons les pouvoirs génératifs et comparatifs du modèle d'interaction mixte sur l'exemple de la conception de RAZZLE.

## 2. LE CAS D'ÉTUDE : RAZZLE

Le système mixte qui nous servira tout au long de cet article pour illustrer le modèle d'interaction mixte est un jeu de réalité augmentée mobile, où l'utilisateur est équipé d'un laptop dans un sac qu'il porte sur le dos, de lunettes semi transparentes où s'affichent des informations par dessus le monde physique, d'un magnétomètre pour capter l'orientation de son regard, d'un touchpad et enfin d'un micro (Figure 1). Ainsi équipé, l'utilisateur doit se déplacer physiquement sur le terrain de jeu, un espace modélisé au préalable dans le système. Le joueur se déplace pour rechercher et ramasser des pièces de puzzle d'origine numérique éparpillées sur le terrain de jeu. Quand l'utilisateur a trouvé une pièce de puzzle, il vérifie que celle-ci est convenablement orientée en la comparant au puzzle modèle. Il la tourne si besoin pour qu'elle s'insère dans le puzzle. Enfin il peut la ramasser. Une fois toutes les pièces ramassées, le joueur peut retourner au lieu de départ pour signifier au système que le jeu est terminé.

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.

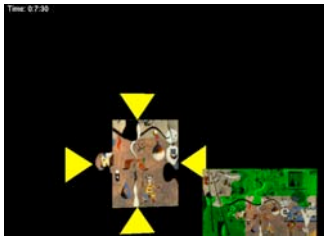
Conference '04, Month 1–2, 2004, City, State, Country.

Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.



**Figure 1. Une joueuse équipée de lunettes semi transparentes et d'un capteur d'orientation.**

L'interface a été spécifiée de la façon suivante : dans les lunettes semi transparentes, l'utilisateur voit le monde physique, mais aussi en bas à droite l'image estompée du puzzle à obtenir (Figure 2). À chaque ramassage d'une pièce, la partie correspondante dans le puzzle se trouve affichée clairement (Figure 2). De plus, puisque le but du jeu est de ramasser les pièces en un minimum de temps, le temps qui s'écoule est affiché en haut à gauche (Figure 2).



**Figure 2. Vue dans les lunettes semi transparentes du joueur (les pixels noirs sont transparents)**

Lorsque l'utilisateur se déplace physiquement sur le terrain de jeu, si une pièce de puzzle placée là est dans son champ de vision, le système affiche cette pièce avec la taille, la position et l'orientation adéquates pour qu'elle semble faire partie du monde physique (Figure 2). L'utilisateur doit alors s'en approcher pour la sélectionner et agir sur elle. Lorsque le joueur est suffisamment proche, alors la pièce est sélectionnée, et le système le signifie à l'utilisateur en affichant des flèches jaunes qui pointent sur les côtés de la pièce (Figure 2).

Une fois la pièce sélectionnée, le joueur compare l'image de la pièce à ramasser avec le modèle de puzzle affiché dans les lunettes semi transparentes. Si la pièce doit être tournée pour entrer dans le puzzle, le système offre trois modalités différentes pour tourner une pièce dans le sens horaire ou antihoraire : nous décrivons la conception des modalités plus précisément grâce au modèle d'interaction mixte dans une partie suivante. En réponse à l'action de l'utilisateur, RAZZLE signifie à l'utilisateur le changement d'état de la pièce en l'affichant correctement orientée.

Une fois la pièce correctement orientée pour entrer dans le puzzle, le joueur doit la ramasser. Pour cela, le système offre encore trois modalités. Si le ramassage n'était pas possible, parce que la pièce n'était pas bien orientée pour entrer dans le puzzle, alors les flèches de sélection autour de la pièce de puzzle deviennent rouges un instant. Dans le cas contraire, si le joueur peut ramasser la pièce, alors elle est ramassée. Elle disparaît donc du terrain de jeu, et apparaît clairement à sa place dans le puzzle affiché dans les lunettes (au lieu d'être estompée).

Après avoir exposé RAZZLE, nous allons maintenant présenter le modèle d'interaction mixte, en illustrant sa définition avec le cas de RAZZLE.

### 3. LE MODÈLE D'INTERACTION MIXTE

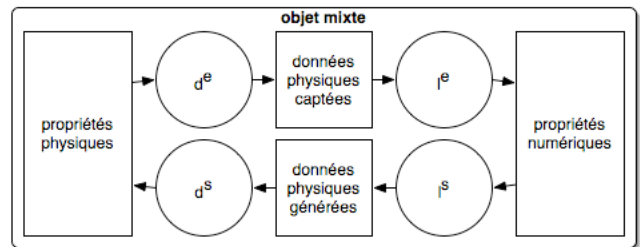
Un modèle d'interaction a pour but d'aider le concepteur d'interface homme-machine dans son travail. Les pouvoirs du modèle lui permettent d'avoir un appui lors de la conception [1].

Le modèle d'interaction mixte [1] fournit ainsi un cadre de travail pour la conception des interfaces mixtes. Il s'appuie sur des travaux venant des interfaces graphiques [1] et des interfaces multimodales. La première notion à définir par rapport à ces travaux est la définition d'un objet mixte, car celui-ci est omniprésent dans l'interaction mixte, en tant qu'objet de la tâche ou en tant qu'outil. Le modèle d'interaction mixte définit donc un objet mixte, à partir des notions de propriétés physiques et numériques. Il étend ensuite le cadre du modèle d'interaction instrumental [1] en considérant que l'utilisateur manipule des objets mixtes grâce à une modalité d'interaction. Nous réutilisons notre définition d'une modalité d'interaction comme un couple  $(d, l)$ , où  $d$  est un dispositif qui délivre ou acquiert l'information, comme un microphone ou des haut-parleurs, et  $l$  est un langage d'interaction, c'est-à-dire un ensemble d'expressions bien formées. Cette définition respecte les distances articulatoires et sémantiques de la théorie de l'action [6]. Nous réutilisons également les résultats concernant la composition des modalités d'interaction de [5][7].

En nous fondant sur ces travaux, nous définissons d'abord l'objet mixte avant de définir le modèle d'interaction mixte.

#### 3.1 Objet mixte

Si nous définissons un objet physique comme un ensemble de propriétés physiques, et un objet numérique comme un ensemble de propriétés numériques, alors un objet mixte est défini comme un ensemble de propriétés physiques reliées à un ensemble de propriétés numériques. Pour décrire le lien entre les deux ensembles de propriétés physiques et numériques, nous considérons les deux niveaux d'abstraction d'une modalité  $(d, l)$ . Une modalité qui lie les propriétés d'un même objet mixte est appelée *modalité de liaison*, à la différence des *modalités d'interaction*. En adoptant un point de vue système, nous décrivons les modalités de liaison d'un objet mixte de la façon suivante (Figure 4) : En entrée du système, une modalité de liaison  $(d^e, l^e)$  permet de (1) capter certaines données parmi les *propriétés physiques*, à l'aide d'un dispositif  $d^e$ , dispositif en entrée de l'objet, (2) puis traduire ces *données physiques captées* en *propriétés numériques*, grâce à un langage  $l^e$ . 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  $l^s$ , (2) rendre ces *données physiques générées* perceptibles par des *propriétés physiques*, grâce à un dispositif  $d^s$ .



**Figure 4. Un objet mixte.**

Nous réutilisons dans le cas des modalités de liaison les résultats sur la composition des modalités d'interaction. Ainsi un objet mixte peut avoir une modalité de liaison composée résultant de la

fusion/fission de plusieurs modalités élémentaires. Par exemple, la pièce de puzzle de RAZZLE (Figure 5) a une position et une orientation du point de vue de l'utilisateur, ces deux modalités de liaison sont donc composées en entrée et véhiculent de l'information pour mettre à jour les propriétés numériques. Au contraire, la modalité de sortie est élémentaire et consiste à afficher l'image de la pièce dans les lunettes semi transparentes.

Un tel objet mixte peut prendre part à l'interaction entre système et utilisateur en tant qu'objet de la tâche ou comme outil. Si l'objet mixte est un outil, alors l'objet mixte est le dispositif  $d$  d'une modalité d'interaction ( $d, l$ ). Maintenant que l'objet mixte est défini, nous abordons dans le paragraphe suivant la notion de modalité d'interaction mixte.

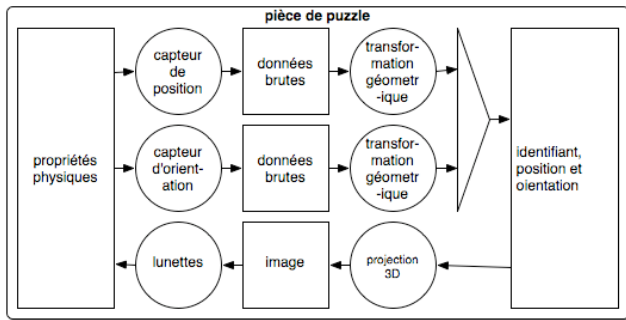


Figure 5. Le modèle mixte de la pièce de puzzle de RAZZLE.

### 3.2 Modalité d'interaction mixte

Comme décrit dans [4], un objet mixte peut participer à l'interaction en tant qu'objet de la tâche ou outil. Pour décrire l'interaction, nous étendons le modèle instrumental d'interaction [1], en considérant des objets mixtes et leurs modalités de liaison, et en considérant aussi les modalités d'interaction mises en jeu.

Dans le modèle instrumental [1], l'interaction entre l'utilisateur et l'objet de la tâche via l'interface graphique se décompose en deux niveaux. Comme le montre la figure 6, l'utilisateur fait une *action* sur l'instrument (outil), qui *réagit* directement. L'outil transforme l'action de l'utilisateur en commande ou *tâche élémentaire* qui s'applique à l'objet de la tâche. L'objet de la tâche renvoie alors une *réponse* vers l'outil et un *retour d'information* directement vers l'utilisateur.

Nous étendons le modèle d'interaction instrumental en affinant les définitions de l'objet de la tâche et de l'outil en objets *mixtes*. L'outil est un objet mixte appelé outil mixte, qui est le dispositif  $d$  d'une modalité d'interaction ( $d, l$ ). On a donc nécessairement un langage  $l$  d'interaction pour former la modalité d'interaction avec l'objet de la tâche. La figure 7 présente une modalité d'interaction mixte où l'outil mixte est le dispositif de la modalité d'interaction, et les langages correspondants permettent d'interagir en entrée et en sortie avec l'objet de la tâche, via la *tâche élémentaire* et la *réponse*. À chaque niveau d'abstraction, les modalités peuvent être composées en suivant les propriétés CARE [5][7].

L'utilisateur fait une action qui modifie les propriétés physiques de l'outil mixte. Les nouvelles propriétés physiques de l'outil mixte sont captées et traduites en entrée pour mettre à jour ses propriétés numériques. Ceci est le dispositif de la modalité d'interaction. En sortie, ces propriétés sont rendues visibles par la réaction de l'outil mixte via la modalité de sortie. La modification des propriétés numériques de l'outil mixte est traduite en tâche élémentaires qui s'applique à l'objet de la tâche via le langage  $l$ .

En sortie, l'objet de la tâche répond à l'outil mixte, via le langage de la modalité en sortie  $l$  (Figure 7).

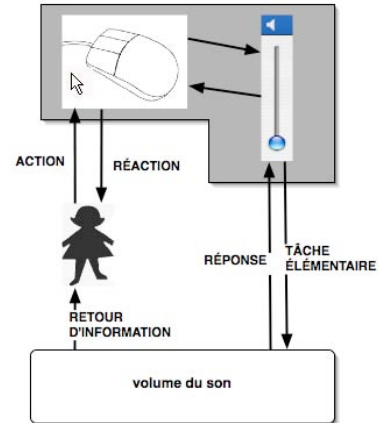


Figure 6. Le modèle d'interaction instrumental appliqué à la tâche élémentaire « augmenter le volume du son ».

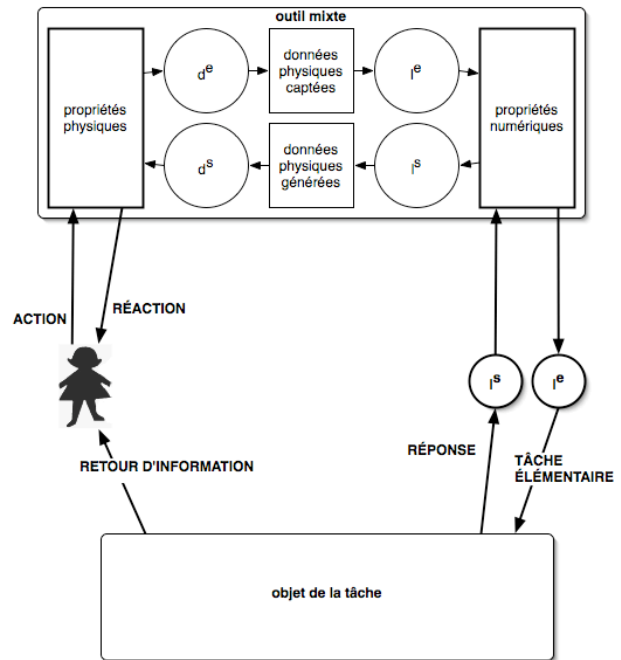


Figure 7. Une modalité d'interaction mixte.

Un des pouvoirs du modèle d'interaction mixte, comme tout modèle d'interaction, est son pouvoir génératif. Celui-ci permet à un concepteur d'explorer de façon systématique l'espace de conception. Nous avons utilisé le modèle d'interaction mixte pour concevoir les modalités d'interaction dans RAZZLE pour tourner et ramasser les pièces de puzzle. Nous étudierons ensuite le pouvoir comparatif du modèle.

## 4. POUVOIRS DU MODÈLE D'INTERACTION MIXTE

Afin d'apprécier les pouvoirs génératifs et comparatifs du modèle d'interaction mixte, nous avons conçu et étudié des modalités d'interaction pour tourner et ramasser les pièces dans RAZZLE.

Les trois modalités d'interaction pour tourner et ramasser les pièces de puzzle ont été conçues suivant le modèle d'interaction mixte. Nous avons choisi trois modalités fonctionnellement équivalentes, pour tourner ou ramasser une pièce de puzzle lors du jeu. Nous avons conçu les trois modalités suivantes à partir du cadre présenté à la Figure 9, en y insérant trois différentes modalités à l'intérieur.

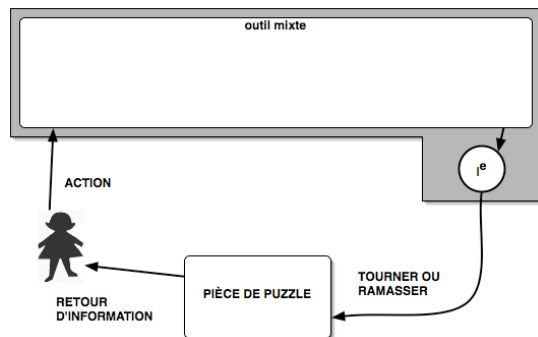


Figure 9. Un cadre pour la conception de RAZZLE.

Chaque modalité est une déclinaison possible de ce cadre général. La première est la modalité tactile qui utilise le touchpad, la seconde la modalité vocale, et la troisième la gestuelle.

Pour utiliser la modalité tactile, l'utilisateur fait un mouvement vers la gauche ou vers la droite sur un touchpad porté au poignet. Le geste vers la gauche fait tourner la pièce dans le sens antihoraire et le geste vers la droite dans le sens horaire. Pour ramasser la pièce qu'il vient d'orienter, le joueur peut cliquer sur le bouton du touchpad. La modalité tactile pour tourner une pièce mixte est modélisée Figure 10.

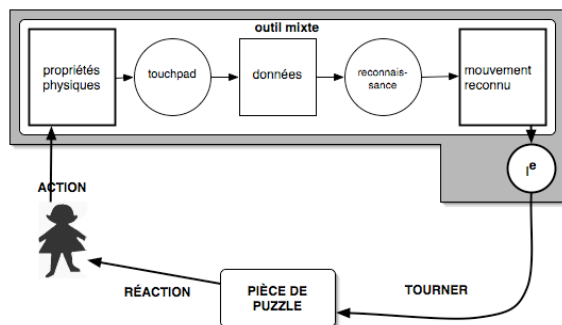


Figure 10. La modalité tactile pour tourner.

Pour utiliser la modalité vocale, le joueur dit qu'il veut faire tourner la pièce en précisant le sens de rotation voulu. Étaient autorisées les expressions telles que « vers la gauche », « gauche », « tourner à gauche », etc. Pour ramasser, l'utilisateur peut signifier au système en disant des mots tels que « attrape », « ramasser », « je prends », etc. qu'il veut ramasser la pièce de puzzle. Dans le cadre du modèle d'interaction mixte Figure 9, la modalité vocale serait composée d'un microphone et d'un langage qui reconnaît les mots précédemment cités.

Pour utiliser la modalité gestuelle, l'utilisateur esquisse un geste dans le sens voulu de rotation à l'aide de sa main dirigé vers la pièce sélectionnée, afin de faire tourner la pièce. Et en faisant un geste de saisie dirigée vers la pièce sélectionnée, l'utilisateur demande au système de ramasser cette pièce. Cette modalité serait représentée par le couple (*caméra, vision par ordinateur*).

Si le concepteur devait choisir entre ces 3 modalités, il pourrait utiliser le pouvoir comparatif du modèle d'interaction mixte, grâce aux propriétés ergonomiques qui sont supportées dans le modèle d'interaction mixte [3]. Par exemple, le critère de continuité nous permet de faire une différence entre la modalité tactile et les deux autres modalités. La continuité de l'outil mixte s'étudie en considérant les modalités de liaisons en entrée et en sortie de l'outil mixte. De façon plus générale, on peut examiner toutes les modalités impliquées dans l'interaction pour accomplir la tâche. Dans le cas de la modalité tactile, la continuité spatiale est moins forte que pour les deux autres modalités, puisque le joueur peut avoir besoin de regarder vers son poignet où est placé le touchpad, au lieu de regarder vers le terrain de jeu où se trouve la pièce de puzzle. En étant capable de distinguer plusieurs options de conception, on donne ainsi la possibilité au concepteur de faire un choix.

## 5. CONCLUSION

Dans cet article, nous avons présenté le modèle d'interaction mixte et nous l'avons illustré sur l'exemple de RAZZLE, un système mixte multimodal. Nous avons introduit la notion de modalité de liaison pour définir un objet mixte, et nous avons fait remarquer que les pouvoirs génératifs et comparatifs du modèle étaient utiles aux concepteurs d'interfaces mixtes, en prenant pour exemple la conception de RAZZLE. Ce modèle cherche un consensus entre les travaux sur les interfaces graphiques [1], les interfaces multimodales [5], et les interfaces mixtes. Afin d'étendre encore ce modèle, nous projetons d'étudier son pouvoir comparatif en capitalisant des propriétés ergonomiques capables d'aider choisir entre plusieurs options de conceptions, pour aboutir à long terme à un guide de conception des systèmes mixtes. Nous explorerons également les liens entre le modèle d'interaction mixte et ICARE [2], un outil de développement pour l'interaction multimodale, basé sur la même définition de modalité. Cet outil pourrait nous permettre de développer rapidement des modalités de liaison et d'interaction.

## 6. RÉFÉRENCES

- [1] Beaudoin-Lafon, Designing Interaction, not Interfaces. *AVI'04*, 15-22.
- [2] Bouchet, Nigay, Ganille, ICARE Software Components for Rapidly Developing Multimodal Interfaces. *ICMI'04*, 251-258.
- [3] Coutrix, Mixed Reality : A model of Mixed Interaction, *AVI'06*, 43-50.
- [4] Dubois, Nigay, Troccaz, Consistency in Augmented Reality Systems. *EHCI'01*, 117-130.
- [5] Nigay, Coutaz, The CARE Properties and Their Impact on Software Design. *Intelligence and Multimodality in Multimedia Interfaces: Research and Applications*, John Lee, AAAI Press, 1997.
- [6] Norman, Cognitive Engineering. *Book chapter of User Centered System Design, New Perspectives on Human-Computer Interaction*, 1986, 31-61.
- [7] Vernier, Nigay, A Framework for the Combination and Characterization of Output Modalities. *DSVIS'00*, 32-48.