

# Interactions Intelligentes à Base de Mouvements

Mathieu Barnachon, Mathilde Ceccaroli, Amélie Cordier, Erwan Guillou, Marie Lefevre

Université de Lyon, CNRS  
Université Lyon 1, LIRIS, UMR5205, F-69622, France  
prenom.nom@liris.cnrs.fr  
<http://liris.cnrs.fr/iibm>

## Résumé

Cet article décrit des travaux réalisés dans le cadre du projet IIBM (Interactions Intelligentes à Base de Mouvements). L'objectif de ce projet est d'explorer le champ des interactions homme-machine dites naturelles selon une double perspective : reconnaissance de mouvements et intelligence artificielle. Dans le projet IIBM, nous utilisons le raisonnement à partir de l'expérience tracée pour modéliser les interactions sous forme de traces d'interactions. En nous appuyant sur de solides résultats en capture de mouvements et en exploitant le principe du raisonnement à partir de l'expérience tracée, nous proposons ainsi une approche permettant à l'utilisateur et au système de co-construire des connaissances partagées sur les « façons » d'interagir. Dans cet article, nous décrivons notre approche et nous illustrons sa mise en œuvre au sein du logiciel d'annotation d'images IIANNOTATE. Nous discutons des premiers résultats obtenus et des pistes de recherche ouvertes par ce travail ainsi que des perspectives d'application dans différents domaines porteurs, tels que ceux du jeu vidéo et des jeux sérieux (serious games).

## Mots Clef

Raisonnement à partir de cas ; Raisonnement à partir de l'expérience tracée ; Traces d'interactions ; Interprétation de gestes ; Interactions Homme-Machine ; Co-construction de connaissances.

## Abstract

*This paper describes an ongoing research conducted within the IIBM project (Intelligent Interactions Based on Motion). The aim of this project is to investigate the field of so-called Natural Interactions from a twofold point of view : motion recognition and artificial intelligence. The term "Natural Interactions" defines the new interaction modalities that surround us nowadays (motion, touch, etc.). These new interaction modalities enable users to control applications through more natural gestures rather than by using a dedicated device such as a mouse. In the IIBM project, we focus on the opportunity of using artificial intelligence techniques in order to improve the motion interpretation phase. We use the trace-based reasoning paradigm*

*in order to model interaction sequences in the form of interaction traces. Modelled interaction traces, associated to the trace-based system, provide us with an additional support for interactions between users and systems. We propose an approach allowing users to interact with software systems in order to co-construct shared knowledge about "how to interact". For that purpose, we rely on solid results in motion capture that we combine to the trace-based reasoning theory. In this paper, we describe our approach and we illustrate its implementation with IIANNOTATE, an image annotation tool. We analyse the first results and we draw preliminary conclusions on the different research issues raised by this work. We also discuss possible applications of this an approach in different promising application fields such as video games and serious games.*

## Keywords

Case-Based Reasoning ; Trace-Based Reasoning ; Interaction traces ; Gesture Recognition ; Human Computer Interaction ; Co-construction of Knowledge.

## 1 Introduction

Dans le domaine des Environnements Informatiques pour l'Apprentissage Humain (EIAH), les jeux sérieux (serious games) sont de plus en plus utilisés [1, 2]. La conception de jeux sérieux comporte deux phases principales : la définition du contenu pédagogique et la conception du jeu lui-même (environnement, *gameplay*, modalités d'interactions, etc.). Concevoir un jeu facile à manipuler est crucial pour permettre une exploitation efficace du contenu pédagogique qu'il contient.

Dans notre travail, nous nous intéressons à un aspect particulier de la conception du jeu : celui des interactions entre le joueur et l'environnement. Plus particulièrement, nous nous intéressons aux interactions gestuelles riches. En effet, depuis quelques années, les éditeurs de jeu vidéo rivalisent de créativité et d'ingéniosité pour proposer au grand public des modalités d'interaction de plus en plus riches. Nous sommes passés du système de bouton à celui de la souris, puis très rapidement à des modalités orientées « mouvements ». Les dispositifs tels que la *WiiMote*, ou le *PS Move* offrent au joueur une nouvelle façon d'en-

visager le jeu. Il ne s'agit plus de cliquer pour effectuer une action, mais, le plus souvent, de « simuler l'action » avec le dispositif de contrôle (tourner un volant, faire un revers au tennis, *etc.*). Naturellement, la tendance vise à faire disparaître totalement les dispositifs de contrôle qui se placent comme des intermédiaires entre l'utilisateur et le jeu. C'est le cas par exemple avec *Kinect* : l'utilisateur interagit directement avec son corps, et le concept de manette disparaît complètement. Nous assistons donc à un véritable changement de paradigme dans la façon d'interagir avec les systèmes. Le niveau d'indirection qui était imposé par l'utilisation de dispositifs de contrôle (clavier, souris, manette) est levé. Les interactions ne sont plus limitées à la combinatoire des boutons disponibles sur la manette, mais la nouvelle limite est définie par la capacité des systèmes à interpréter les gestes des utilisateurs.

Cette révolution dans la façon d'envisager le jeu a un impact considérable : nous voyons apparaître de nouveaux joueurs et, avec eux, se profilent de nouvelles attentes. En effet, bien que les joueurs invétérés décrivent ces nouvelles approches, force est de constater que ces nouvelles modalités d'interactions attirent l'attention d'un public différent, parfois inattendu. Les jeux s'ouvrent à de nouveaux joueurs qui ne trouvaient pas naturel l'ancien paradigme, mais qui se familiarisent mieux, et très rapidement, avec ces modes d'interactions plus naturels. Mais ces nouveaux joueurs sont exigeants : les jeux se doivent d'être plus riches, plus réactifs, plus créatifs, plus intuitifs.

Ces nouvelles modalités de contrôle, associées à l'apparition de nouveaux joueurs, nous conduisent naturellement à repenser la façon donc nous concevons les interactions. En effet, le recours à des mouvements augmente considérablement le nombre d'interactions possibles. Face à cette multiplicité, deux situations sont alors envisageables : soit les systèmes limitent *a priori* les interactions qu'ils sont capables d'interpréter, soit ils se dotent de la capacité « d'apprendre » de nouvelles interactions, au fur et à mesure de leur utilisation. Dans le premier cas (le plus fréquent), la richesse offerte par les interactions gestuelles n'est pas exploitée à sa juste valeur. Les jeux sont conçus *a priori* pour interpréter les gestes « stéréotypés » des utilisateurs et ne présentent aucune faculté d'adaptation à l'utilisateur. Le second cas offre des perspectives riches, mais reste plus difficile à mettre en œuvre. C'est à ce second cas que nous nous intéressons ici. Nous cherchons à construire des outils permettant aux systèmes de s'adapter aux utilisateurs et surtout, à leurs façons d'interagir. Ces outils permettront aux concepteurs de jeux d'envisager de nouvelles façons de concevoir les jeux et donc de jouer.

Dans cet article, nous décrivons des travaux exploratoires autour de la question des interactions intelligentes à base de mouvements. Nous dépassons le concept classique de reconnaissance de mouvements (une action reconnue implique une réaction par le système) en ajoutant l'étape d'interprétation, qui permet plus de souplesse dans le pro-

cessus. En effet, pour aller au delà de la simple identification d'un geste correspondant à un modèle, le système s'appuie sur une base de connaissances pour « interpréter » les gestes effectués par les utilisateurs et propose donc des réactions adaptées au contexte et à l'utilisateur. Ce processus permet également au système d'apprendre de nouveaux gestes, en collaboration avec l'utilisateur.

Pour développer ce concept d'interactions intelligentes, nous combinons deux domaines de recherche : l'acquisition de mouvements et l'intelligence artificielle (et plus précisément le raisonnement à partir de l'expérience tracée [3]). L'acquisition de mouvements nous permet de nous outiller pour envisager des modalités de communication riches entre un utilisateur et un outil informatique. Le raisonnement à partir de l'expérience tracée nous permet de doter nos systèmes de mécanismes et de connaissances pour interpréter les mouvements des utilisateurs. Ces deux outils sont combinés afin de proposer une approche d'interprétation de mouvements.

Cet article est organisé de la façon suivante. Dans la section 2, nous décrivons le projet IIBM, qui constitue le cadre de ces travaux de recherche. Nous détaillons les deux aspects principaux du projet : la reconnaissance de mouvements et l'utilisation du raisonnement à partir de l'expérience tracée. Dans la section 3, nous illustrons une mise en œuvre de notre approche à l'aide de l'application IIANNOTATE, une application d'annotation d'images entièrement contrôlable par l'intermédiaire de gestes. Dans la section 4, nous discutons de l'intérêt d'appliquer cette approche dans le domaine du jeu vidéo, et par extension, à celui des jeux sérieux, ainsi que des problématiques qui sont liées. Une discussion sur la base des premiers résultats obtenus et sur les perspectives de recherche envisageables conclut l'article.

## 2 Le projet IIBM

Le projet IIBM (Interactions Intelligentes à Base de Mouvements) a pour but de combiner l'interprétation de mouvements et l'intelligence artificielle afin de développer des interactions naturelles [4]. Le terme « interactions naturelles » caractérise les nouvelles modalités d'interactions (mouvements, toucher, vue, ouïe, *etc.*) qui permettent aux utilisateurs de contrôler des applications non plus *via* un dispositif d'indirection classique tel qu'une souris, mais par des gestes plus naturels. Dans notre contexte, nous nous intéressons aux interactions gestuelles permettant aux utilisateurs d'interagir avec des applications en faisant des mouvements, et sans utiliser de dispositif de contrôle particulier (clavier, souris, manette). Dans ce travail, nous ne nous focalisons pas uniquement sur la capture et la reconnaissance de gestes. Nous mettons avant tout l'accent sur l'exploitation intelligente des interactions gestuelles.

Le projet IIBM rassemble des chercheurs de deux domaines différents : la capture de mouvement et l'intelligence artificielle. Il explore de nouvelles façons de per-

mettre les interactions homme-machine (IHM) en essayant d'apporter une réponse à la question suivante : comment l'intelligence artificielle peut contribuer à enrichir les interactions entre un utilisateur et un outil numérique ? Le couplage de ces deux approches devrait conduire à l'élaboration de solutions permettant aux utilisateurs et aux systèmes de négocier afin de co-construire des connaissances partagées sur la façon d'interagir ensemble. Le principal défi consiste à orchestrer des interactions, afin de faciliter cette négociation de sens [5].

Les défis à relever se situent à la fois au niveau de la capture de mouvements et de l'intelligence artificielle (dans ce cas précis, nous parlons du paradigme du raisonnement à partir de l'expérience tracée). Au niveau de la capture de mouvements, il s'agit de développer des algorithmes de reconnaissance de mouvements fonctionnant sans apprentissage automatique (ni semi-supervisé) et utilisables en temps réel (voir section 2.1). Le raisonnement à partir de l'expérience tracée est utilisé pour faciliter la réutilisation de l'expérience des utilisateurs mais également pour permettre le partage d'expériences entre utilisateurs. Cela nécessite de capturer cette expérience, de trouver des moyens dynamiques de la structurer dans des « cas », et d'organiser ces cas pour une réutilisation future (voir section 2.2). Les résultats attendus sont des applications capables non seulement de reconnaître des actions utilisateurs prédéfinies et de déclencher des réponses systèmes en conséquence, mais aussi de découvrir de nouvelles actions faites par les utilisateurs, de leur associer un sens et une action ou un ensemble d'actions à effectuer.

La figure 1 résume le principe du projet IIBM. Un *utilisateur* interagit avec une *application* (jeu ou toute autre application logicielle) par l'intermédiaire de gestes. Ces gestes sont reconnus et *interprétés* par l'application qui réagit en conséquence. Au cours des interactions avec l'environnement, l'utilisateur laisse des *traces d'interactions*. Ces traces sont stockées dans un *système de gestion de traces*<sup>1</sup> (une sorte de base de données spécialisée pour ce type d'objet). Les traces sont exploitées par un processus de *raisonnement* (le raisonnement à partir de l'expérience tracée) afin d'enrichir la phase d'interprétation des gestes de l'utilisateur. Ainsi, au fur et à mesure de son utilisation, l'application devient de plus en plus efficace et apte à reconnaître les différents gestes effectués par l'utilisateur, même si ceux-ci n'ont pas été prévus lors de l'initialisation du système.

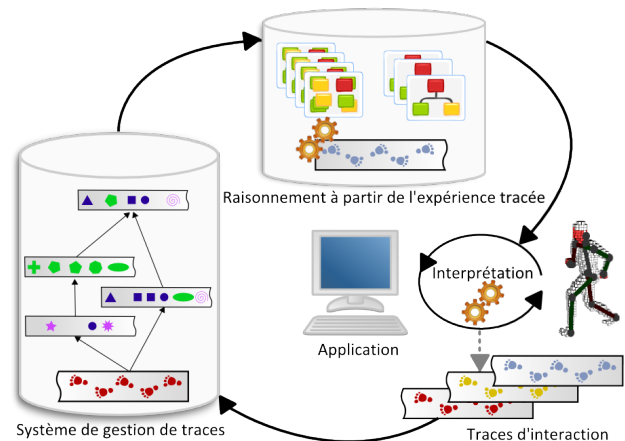


FIGURE 1 – Principes du projet IIBM.

Naturellement, notre domaine d'application cible est celui du jeu vidéo. En effet, dans ce domaine, prendre en compte les spécificités des joueurs serait une formidable valeur ajoutée en matière de jouabilité. Une telle approche permettrait également aux concepteurs de jeux d'ajouter plus de créativité dans leurs jeux. Cependant, cette approche est utile pour d'autres applications où les interactions des utilisateurs jouent un rôle tout aussi important. L'application d'annotation d'images, IANNOTATE, présentée dans la section 3, en témoigne.

Notre approche est liée au champ de recherche sur les interactions homme-machine, et plus précisément, à celui de interactions à base de mouvements [6]. À notre connaissance, il existe peu de travaux combinant l'intelligence artificielle et la capture de mouvements pour construire des systèmes intelligents et adaptatifs. Certaines approches puisent leur inspiration dans des techniques telles que les K plus proches voisins (KNN), les classificateurs ou les modèles de Markov cachés (HMM) pour perfectionner les algorithmes de capture de mouvements (cf. [7, 8, 9]). Dans ces travaux, l'objectif est uniquement d'améliorer les performances en matière de reconnaissance de mouvements prédéfinis.

D'autres approches, telles que [10], utilisent l'intelligence artificielle pour mettre en œuvre des comportements intelligents dans les applications, mais elles s'appuient sur l'observation d'interactions classiques (clavier/souris) et ne considèrent absolument pas les interactions gestuelles. Dans ces travaux, le champ des interactions envisageables est limité par les possibilités de l'interface, et est donc connu à l'avance.

Notre approche repose sur le paradigme du raisonnement à partir de l'expérience tracée. Ce paradigme présente de très nombreuses similarités avec le raisonnement à partir de cas (CBR [11]). Dans le domaine du jeu vidéo, le raisonnement à partir de cas a été utilisé à de nombreuses occasions [12, 13, 14]. Dans toutes ces approches, ce paradigme est utilisé soit pour apprendre les comportements des utilisateurs et les reproduire dans le jeu, soit pour choi-

1. <http://liris.cnrs.fr/sbt-dev/ktbs/>

sir des stratégies appropriées en fonction du contexte. Il s'agit donc essentiellement d'utiliser ce mode de raisonnement pour faire évoluer l'intelligence artificielle des personnages du jeu, et non la façon dont le joueur et le jeu interagissent. Dans notre cas, nous utilisons une approche inspirée du CBR pour enrichir les interactions, et en particulier les interactions basées sur le mouvement, entre l'utilisateur et le jeu.

Ainsi, même si différents outils issus de l'intelligence artificielle croisent régulièrement le domaine du jeu vidéo, nous n'avons pas identifié de travail considérant spécifiquement le couplage entre capture de mouvements et CBR dans une approche intégrée.

## 2.1 Capture et interprétation de mouvements

La plupart des travaux relevant du domaine des Interfaces Homme-Machine (IHM) s'intéressent à des situations dans lesquelles le nombre d'interactions possibles est très restreint. Le plus souvent, le repérage des interactions passe par l'utilisation de marqueurs afin de faciliter le suivi et la correspondance des différents éléments entre chaque image considérée. Parmi les applications de ce type, on peut citer Zhang *et al.* [15], qui ont utilisé une feuille de papier pour déplacer la souris, cliquer, tourner des objets et dessiner. Le principal inconvénient de leur solution est l'utilisation de marqueur, qui bien que simple, limite considérablement le nombre d'actions disponibles et les conditions dans lesquelles le dispositif peut être utilisé.

D'autres solutions se sont affranchies des marqueurs [16, 17]. N'utilisant pas de marqueurs, elles sont plus adaptées au contexte des IHM intuitives mais restent limitées en termes d'actions de par les capacités d'acquisition des mouvements. Si ces solutions sont suffisamment simples pour être des substituts à la souris ou à quelques raccourcis claviers, elles ne permettent pas d'atteindre l'objectif des interactions naturelles à base de mouvements.

L'interprétation de mouvements a été très largement étudiée dans le domaine de la surveillance [18, 19, 20]. Ces méthodes ont de bonnes performances lorsqu'il s'agit de détecter des activités intrinsèquement différentes les unes des autres – comme l'action de sauter par rapport à l'action de marcher – mais elles requièrent de nombreuses variations d'un mouvement (grande variance intra-classe). Cela s'explique par le fait qu'elles sont basées sur des méthodes d'apprentissage statistique. De ce fait, elles sont peu flexibles, lourdes à créer et chaque nouvelle action impose une réexécution complète de l'algorithme d'apprentissage, souvent non incrémental. Un état de l'art récent des méthodes de surveillance d'activités est disponible dans [21].

## 2.2 Raisonnement à partir de traces et interactions intelligentes

Le raisonnement à partir de cas (CBR) est un paradigme de résolution de problème qui consiste à réutiliser des cas passés pour résoudre de nouveaux problèmes. Un cas est un problème dont on connaît la solution. Lorsque l'on cherche à résoudre un nouveau problème, on recherche un cas passé similaire (c'est-à-dire un problème semblable) et on en réutilise la solution, en l'adaptant éventuellement au problème que l'on cherche à résoudre. Les systèmes de raisonnement à partir de cas ne se contentent pas de résoudre des problèmes : ils acquièrent également des connaissances de résolution de problèmes. Le plus souvent, cette acquisition passe par des interactions avec l'utilisateur qui guide le système dans le processus de résolution de problème.

Le raisonnement à partir de l'expérience tracée (*Trace-Based Reasoning*, TBR) peut être envisagé comme une forme de raisonnement à partir de cas [22]. L'objectif du TBR est en effet de fournir des mécanismes de raisonnement permettant de réutiliser les expériences des utilisateurs dans différents contextes (résolution de problèmes, assistance à l'utilisateur, apprentissage, découverte de connaissances, *etc.*). Le système étudié est instrumenté de sorte à ce que les expériences des utilisateurs soient naturellement capturées sous forme de traces d'interactions. Il est possible d'identifier, dans ces traces d'interactions, des *épisodes* qui correspondent à des actions (ou tâches) spécifiques réalisées par les utilisateurs [23, 24]. Ces épisodes peuvent ensuite être réutilisés dans différents contextes. De manière analogue au CBR, le TBR permet donc d'acquérir des connaissances, notamment sur les interactions des utilisateurs.

Dans le contexte du projet IIBM, nous utilisons le TBR pour faciliter les interactions entre l'utilisateur et le système. Les traces se présentent comme un support de négociation entre l'utilisateur et le système, tant au niveau des gestes que des actions. Le système observe en permanence l'utilisateur et cherche à identifier les actions que ce dernier est en train d'effectuer. Pour cela, il recherche dans la base de traces des épisodes similaires. Si un épisode est retrouvé, le système cherche à le réutiliser pour proposer de l'aide à l'utilisateur (compléter une action par exemple). À tout moment, l'utilisateur peut interagir avec le système pour négocier de nouveaux comportements (associer un geste à une action, corriger une mauvaise interprétation effectuée par le système, *etc.*). De façon analogue à ce que l'on observe en CBR, dans ce cas, le système saisit l'opportunité pour acquérir davantage de connaissances sur les façons d'interagir de l'utilisateur.

Dans IIBM, le TBR est également utilisé pour permettre le partage d'expériences entre les différents utilisateurs [3]. En effet, le système peut réutiliser les traces d'un utilisateur pour en aider un autre. Ainsi, il fait bénéficier le second de l'expérience acquise par le premier.

Le scénario détaillé dans la section suivante illustre la façon dont le raisonnement à partir de l'expérience tracée est mis en œuvre dans IIBM.

### 3 Une étude de cas avec IANNOTATE

Afin de réaliser de premières expérimentations, nous avons conçu et implémenté une application contrôlable par les gestes. Cette application, nommée IANNOTATE, permet la gestion et l'annotation d'images. La figure 2 présente l'interface principale de l'application. Celle-ci est composée d'un menu classique ainsi que d'une visionneuse affichant :

- une ou plusieurs photos ;
- une barre de miniatures permettant aux utilisateurs de visionner les photos ;
- une barre de tags permettant aux utilisateurs de gérer les tags qui peuvent être utilisés pour annoter les photos.

Au sein de la visionneuse, des onglets permettent à l'utilisateur de travailler simultanément sur plusieurs groupes de photos.

Les fonctionnalités de l'application sont nombreuses (annotation d'une photo ou un groupe de photos, création d'un tag ou super-tag, gestion des utilisateurs, *etc.*). Une des principales caractéristiques de cette application est qu'elle a été conçue afin d'être complètement manipulée par des gestes. L'acquisition des mouvements se fait à travers un *middleware* s'appuyant sur *Kinect*.

L'application dispose d'un mécanisme qui enregistre toutes les interactions entre l'utilisateur et le système sous la forme de traces d'interactions. Les traces d'interactions sont ensuite exploitées pour améliorer l'interprétation des mouvements, visant ainsi à offrir une meilleure ergonomie de l'application. Afin d'expérimenter cette idée, nous avons défini plusieurs scénarios. Un de ces scénarios est détaillé ci-dessous pour donner une vue plus claire des avantages de l'application. Le scénario met en lumière la contribution du TBR dans l'interprétation des actions des utilisateurs.

Alex utilise l'application IANNOTATE pour annoter ses photos. Il annote une première sélection de photos en leur affectant les tags suivants :

- *vacances, été, 2011* et *oiseaux* pour les deux premières photos ;
- *vacances, été, 2011* et *amis* pour la troisième photo ;
- *vacances, été, 2011* et *paysages* pour la quatrième photo.

Alors qu'il commence à annoter la cinquième photo avec les tags *vacances, été* et *2011*, le système identifie qu'il a effectué des actions similaires dans le passé (étape de remémoration du CBR). En s'appuyant sur sa base de connaissances, le système recommande une action appropriée dans cette situation : « créer un super-tag » (un super-tag est un raccourci pour un groupe de tags, cette recommandation s'appuie sur une heuristique connue par le système). Alex accepte cette recommandation et crée un super-tag pour le

groupe de tags « *vacances+été+2011* ». Alex est maintenant capable d'annoter ses photos avec trois tags en une seule action. Il peut même associer un geste personnel à ce super-tag.

Dans cette première partie de notre scénario, le TBR est utilisé à deux fins différentes. Premièrement, il permet au système d'identifier une action répétitive de l'utilisateur et de lui fournir une aide appropriée. Ensuite, il est utilisé pour apprendre le geste que l'utilisateur souhaite associer à une action spécifique (à savoir, annoter une image avec un super-tag spécifique).

Alex annote ensuite un second dossier contenant des photos prises à Noël. Toutes les photos sont étiquetées avec les tags *Noël* et *vacances*, et cinq d'entre elles avec le tag *Arbre de Noël*. Un deuxième utilisateur, Marc, utilise la même application pour annoter ses photos. Il convient d'observer que le système permet le partage d'expériences entre les utilisateurs. Ainsi, même si Marc dispose de son espace privé avec ses propres photos et ses propres tags, il peut tout de même bénéficier des expériences des autres utilisateurs en matière d'interactions. Marc annoté ses photos de vacances prises pendant l'été 2011. Il annoté la première photo avec *vacances, été, 2011* et *Papa*. Quand il annoté la photo suivante en utilisant *vacances, été, 2011* pour la deuxième fois, le système propose immédiatement de créer un super-tag. Ici, les actions réalisées par Marc ont été analysées à la volée, et le système a identifié une action similaire dans sa base de traces. En effet, dans une situation antérieure (le cas d'Alex), le système a fourni à l'utilisateur une recommandation qui a été acceptée. Par conséquent, le système repropose plus rapidement la même recommandation. Si Marc est satisfait de l'aide proposée et qu'il accepte de créer le super-tag recommandé, le système lui proposera alors une deuxième forme d'assistance. Le système montrera à Marc le geste appris par Alex et lui demandera s'il veut utiliser le même geste ou en créer un autre. Cette deuxième partie du scénario montre comment le TBR permet le partage d'expériences entre les utilisateurs et améliore l'efficacité du processus d'annotation en fournissant une assistance pertinente.

Ensuite Marc annoté un second dossier de photos. Il commence par utiliser les tags *Arbre de Noël* et *vacances*. Le système recommande de compléter ces tags avec le tag *Noël* parce qu'à chaque fois qu'Alex a utilisé les tags *Arbre de Noël* et *vacances*, il a également utilisé le tag *Noël*. Dans ce cas, le système utilise l'expérience d'un autre utilisateur pour fournir des annotations supplémentaires.

Marc traite enfin un troisième dossier en utilisant trois fois de suite les tags *Savoie, montagne* et *été*. Le système utilise l'expérience d'Alex pour suggérer, plus rapidement que la première fois, la création d'un super-tag et d'un geste associé. En effet, IANNOTATE est capable de généraliser les expériences acquises. Dans ce cas, la recommandation proposant la création d'un super-tag, quand trois tags sont souvent utilisés ensemble, a été généralisée à partir de

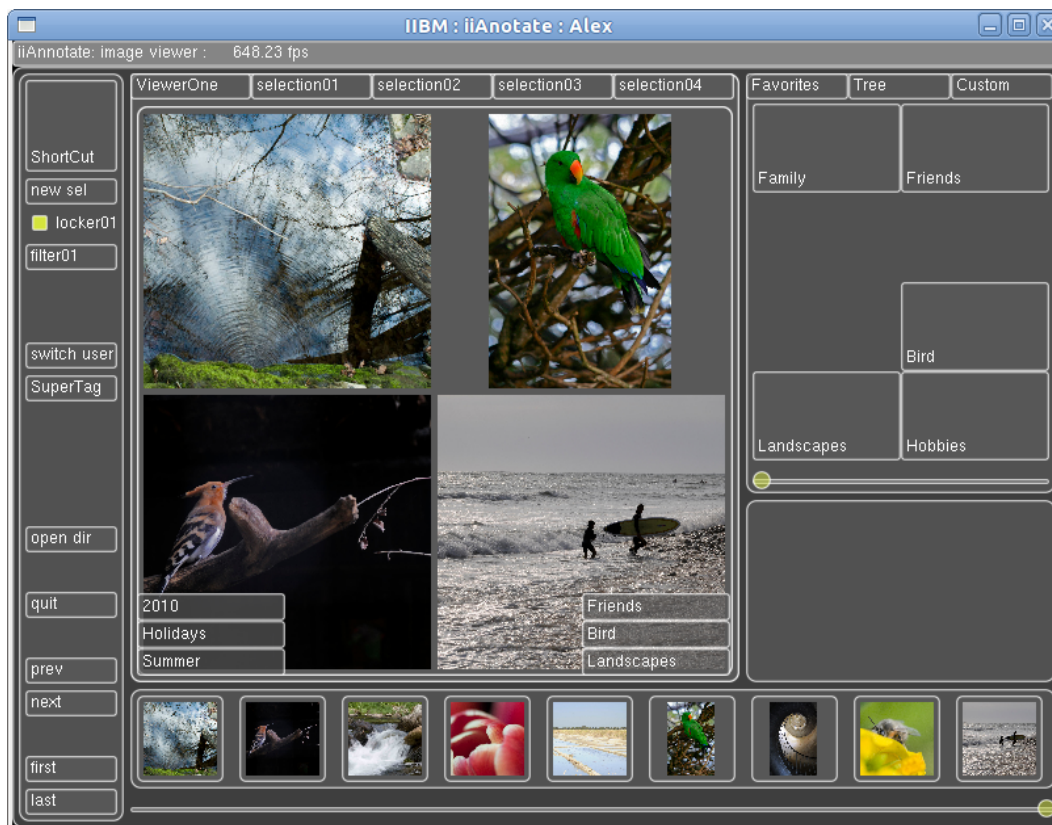


FIGURE 2 – Interface de l'application d'annotation de photos IIANNOTATE.

l'assistance qui a été faite à Alex au préalable.

IIANNOTATE est encore en développement, et certaines parties de ce scénario ne sont pas implémentées. Nos premiers résultats montrent que l'approche de raisonnement à partir de l'expérience tracée que nous avons appliquée fournit des résultats pertinents et aide les utilisateurs dans leur tâche d'annotation d'images. D'autres développements nous permettront d'obtenir davantage de résultats qui seront publiés sur le site web<sup>2</sup> du projet IIBM dès qu'ils seront disponibles.

## 4 Des interactions intelligentes aux jeux vidéo

Bien que nous travaillons pour l'instant sur un outil d'annotation de photos, notre objectif principal au sein du projet IIBM reste d'appliquer les interactions intelligentes aux jeux vidéo. Nous pensons en effet que ce domaine constitue un terrain de démonstrations particulièrement riche. Dans cette section, nous discutons de quelques arguments qui justifient ce point de vue.

Grâce aux nouveaux modes de contrôle, les jeux vidéo actuels pourraient proposer des interactions beaucoup plus riches que celles qu'ils offrent actuellement. Ce faisant, ils pourraient toucher un plus large public, et proposer de

nouvelles expériences aux joueurs, même les plus agueris. Malheureusement, le potentiel des interactions intelligentes n'est pas, à l'heure actuelle, suffisamment utilisé au sein des jeux. La limitation principale provient de la capacité des jeux à *interpréter* les actions des utilisateurs et à les exploiter de façon intelligente. Dans la suite, avec plusieurs exemples, nous montrons comment de nouvelles façons d'interagir pourraient enrichir les expériences de jeu.

Une utilisation naturelle des interactions gestuelles est visible dans les jeux de sport (danse, foot, ...). Nous avons mis en place une démonstration de jeu de danse "sans manette" en 2009. Dans cette démonstration, le joueur est filmé par une simple webcam, et il lui suffit de bouger pour jouer<sup>3</sup>. En 2010, cette démonstration a été adaptée avec une caméra *Time-Of-Flight*<sup>4</sup>, et ceci avant la sortie du périphérique *Kinect* de Microsoft. Cette démonstration est similaire au jeu *Dance Central* [25], commercialisé à la sortie de *Kinect*. Bien que ce jeu soit très impressionnant d'un point de vue technique, il reste très basique au niveau des interactions possibles (il n'y a que quatre boutons à actionner).

3. 2<sup>ème</sup> vidéo à <http://liris.cnrs.fr/~bmichoud/english/index.shtml>

4. Caméra dite : « temps de vol », fournissant une carte de profondeur plutôt qu'une intensité lumineuse, Mesa Imaging SR4000, <http://www.mesa-imaging.ch>

2. IIBM project : <http://liris.cnrs.fr/iibm>

Un autre type de jeu vidéo qui trouve un nouvel essor grâce aux interactions intelligentes est le *First Person Shooter (FPS)*. En effet, si les joueurs les plus expérimentés préféreront toujours l'efficacité d'une manette, il n'en est pas de même pour les joueurs plus occasionnels. En ajoutant plus d'interactivité et d'interactions dans ses jeux, les joueurs s'identifient d'avantages à leurs avatars. Les combats à l'épée, de boxe, ou de tirs peuvent être enrichis par l'interprétation de gestes et la négociation entre utilisateurs et machine.

Les jeux de rôles ou *Role Playing Game (RPG)*, qui offrent souvent des environnements oniriques et de nombreuses métaphores, sont tout indiqués pour expérimenter ces nouvelles interactions. La négociation de la façon d'interagir peut ici permettre au joueur de personnaliser son environnement, voire son avatar. Prenons l'exemple d'un joueur qui prend le rôle d'un druide dans un *RPG*. L'utilisateur peut interagir avec son avatar afin de lui faire apprendre des sortilèges, et apprendre lui-même les gestes pour les exécuter par la même occasion. En partant d'un sortilège de base, et en interagissant avec le système pour apprendre à l'exécuter, le joueur peut inventer un sortilège différent (plus fort, plus rapide, *etc.*). Un joueur apprenant ses sortilèges à partir d'un modèle, ne cherche plus nécessairement la fidélité absolue de reproduction, mais une manière personnelle de faire ses gestes. La même idée peut aussi être appliquée aux combats à l'épée, aux voleurs, *etc.* qui sont des personnages de tout *RPG*.

Les *RPG* sont une cible intéressante, mais il en existe encore une autre parmi les jeux vidéo : les jeux de stratégies en temps réel ou *Real Time Strategy (RTS)*. Traditionnellement ces jeux sont joués au clavier et à la souris. Si ces jeux sont très populaires chez les utilisateurs de PC, ce n'est pas le cas pour les joueurs sur console. La raison est simple : ils sont très difficiles à manipuler avec des manettes conventionnelles. Une réponse possible à ce problème serait de concevoir de nouvelles façons de contrôler ces jeux, *via* des gestes. Une idée serait de faire un jeu où l'utilisateur peut dynamiquement créer ses « raccourcis gestuels » pour des actions complexes. Ce genre d'interactions offre des opportunités inégalées pour se mouvoir dans les univers des jeux vidéo.

## 5 Conclusion et perspectives

Dans cet article, nous avons défendu l'idée qu'en combinant des techniques d'acquisition de mouvement et des méthodes d'intelligence artificielle, il est possible d'enrichir les interactions entre les utilisateurs et les outils informatiques. L'approche que nous proposons permet à l'utilisateur et au système de négocier afin de co-construire des connaissances sur leur façon d'interagir. Cette négociation s'effectue de manière progressive, au fur et à mesure des utilisations du système, et facilite ainsi l'appropriation de l'application par ses utilisateurs. Nous avons montré l'intérêt du raisonnement à partir de l'expérience tracée pour permettre cette négociation.

Ces nouveaux modes d'interactions trouvent tout naturellement leur place dans le domaine du jeu vidéo. Ils contribuent activement à faciliter la prise en main des jeux par les utilisateurs. Cette remarque est également valable pour les EIAH : des modes d'interaction plus intuitifs facilitent la prise en main des jeux sérieux par les apprenants et leur permettent de se concentrer sur l'apprentissage de nouveaux savoir-faire et de nouvelles connaissances, sans être gênés par des problèmes d'interface. Par ailleurs, l'approche que nous proposons favorise la collaboration entre l'utilisateur et le système et permet l'acquisition de connaissances spécifiques aux utilisateurs. Les systèmes développés selon cette approche possèdent certaines facultés d'adaptation aux utilisateurs. Nous pensons que ces résultats, appliqués dans le cadre des jeux sérieux, contribueraient à la démarche de personnalisation des EIAH.

Ces travaux n'en sont encore qu'à leurs débuts mais nous sommes convaincus du potentiel de l'approche proposée. Les perspectives de recherche ouvertes par ce projet sont nombreuses. Les premiers résultats nous encouragent à poursuivre ce travail et à l'étendre à d'autres applications. Dans un futur proche, nous enrichirons le démonstrateur IIANNOTATE afin de continuer à illustrer l'approche présentée ici. En parallèle, nous envisageons de développer un autre démonstrateur intégré à un jeu vidéo afin de démontrer le potentiel des interactions intelligentes dans un milieu naturellement riche en interactions. Des résultats sont attendus dans le domaine de la reconnaissance de mouvement où l'on espère que la prise en compte de l'utilisateur améliorera la qualité de la reconnaissance. D'autre part, nous espérons que ces travaux nous permettront de faire évoluer la théorie du raisonnement à partir de l'expérience tracée sur laquelle nous travaillons.

## Références

- [1] Martens, A., Diener, H., Malo, S. : Game-based learning with computers : learning, simulations, and games. In : Transactions on edutainment I. (2008) 172–190
- [2] E. Sanchez, M. Ney, J.M.L. : Jeux sérieux et pédagogie universitaire : de la conception à l'évaluation des apprentissages. *Revue Internationale des Technologies en Pédagogie Universitaire* **8** (2011) 48–57
- [3] Cordier, A., Mascret, B., Mille, A. : Extending Case-Based Reasoning with Traces. In : Grand Challenges for reasoning from experiences, Workshop at IJCAI'09. (July 2009)
- [4] OpenNI : The Open Natural Interaction Organization. [www.openni.org](http://www.openni.org)
- [5] Stuber, A. : Co-construction de sens par négociation pour la réutilisation en situation de l'expérience tracée. Thèse de doctorat en informatique, Université Lyon 1 (December 2007)

- [6] Sears, A. and Jacko, J.A. : The human-computer interaction handbook : fundamentals, evolving technologies, and emerging applications (2nd Edition). Human factors and ergonomics. Lawrence Erlbaum Associates (2008)
- [7] Krüger, Björn and Tautges, Jochen and Weber, Andreas and Zinke, Arno : Fast local and global similarity searches in large motion capture databases. In : Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. SCA '10, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2010) 1 – 10
- [8] Okada, R., Stenger, B. : A single camera motion capture system for human-computer interaction. IEICE - Trans. Inf. Syst. (2008) 1855–1862
- [9] Zheng, Yue and Hicks, Yulia and Marshall, Dave and Cosker, Darren : Real-Time Generation of Interactive Virtual Human Behaviours. In Ranchordas, Alpesh-Kumar and Araújo, Hélder J. and Pereira, João Madeiras and Braz, José, ed. : Vision and Computer Graphics. Theory and Applications. Volume 24 of Communications in Computer and Information Science. Springer Berlin Heidelberg (2009) 70 – 82
- [10] Stumpf, Simone and Rajaram, Vidya and Li, Lida and Wong, Weng-Keen and Burnett, Margaret and Dietterich, Thomas and Sullivan, Erin and Herlocker, Jonathan : Interacting meaningfully with machine learning systems : Three experiments. Int. J. Hum.-Comput. Stud. **67** (August 2009) 639–662
- [11] Aamodt, A., Plaza, E. : Case-based reasoning : foundational issues, methodological variations, and system approaches. AI Commun. **7** (March 1994) 39–59
- [12] Molineaux, M. ; Aha, D., Moore, P. : Learning Continuous Action Models in a Real-Time Strategy Environment. In : Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference (FLAIRS 2008). Volume Coconut Grove, USA, May 15-17. AAAI Press. 257–262
- [13] Puga, G.F. ; Díaz-Agudo, B., González-Calero, P. : Experience-Based Design of Behaviors in Videogames. In : Proceedings of the European Conference on Case-Based Reasoning (ECCBR 2008). Volume Trier, Germany, September 1-4. Springer. 180–194
- [14] Hulpus, I. ; Fradinho, M., Hayes, C. : On-the-Fly Adaptive Planning for Game-Based Learning. In : Proceedings of the International Conference on Case-Based Reasoning (ICCBR 2010). Volume Alessandria, Italy, July 19-22. Springer. 375–389
- [15] Zhang, Zhengyou and Wu, Ying and Shan, Ying and Shafer, Steven : Visual panel : virtual mouse, keyboard and 3D controller with an ordinary piece of paper. In : PUI '01 : Proceedings of the 2001 workshop on Perceptive user interfaces, New York, NY, USA, ACM (2001) 1–8
- [16] Shanqing Li and Jingjun Lv and Yihua Xu and Yunde Jia : EyeScreen : A Gesture Interface for Manipulating On-Screen Objects. In : Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments. (2007) 710–717
- [17] Malik, S., Laszlo, J. : Visual touchpad : a two-handed gestural input device. In : Proceedings of the 6th international conference on Multimodal interfaces. ICMI '04 (2004) 289–296
- [18] Ayers, D., Shah, M. : Monitoring human behavior from video taken in an office environment. Image and Vision Computing **19**(12) (2001) 833–846
- [19] Cuntoor, N., Yegnanarayana, B., Chellappa, R. : Activity modeling using event probability sequences. IEEE Trans. Image Processing **17**(4) (April 2008) 594–607
- [20] Ivanov, Y.A., Bobick, A.F. : Recognition of Visual Activities and Interactions by Stochastic Parsing. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 852–872
- [21] Poppe, R. : A survey on vision-based human action recognition. Image and Vision Computing **28**(6) (2010) 976 – 990
- [22] Mille, A. : From case-based reasoning to traces-based reasoning. Annual Reviews in Control **30**(2) (October 2006) 223–232 Journal of IFAC.
- [23] Cram, D., Cordier, A., Mille, A. : An Interactive Algorithm for the Complete Discovery of Chronicles. Technical Report RR-LIRIS-2009-011, LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/École Centrale de Lyon (April 2009)
- [24] Zarka, R., Cordier, A., Egyed-Zsigmond, E., Mille, A. : Rule-Based Impact Propagation for Trace Replay. In Ram, A., Wiratunga, N., eds. : International Case-Based Reasoning Conference (ICCBR 2011). LNAI 6880, Springer-Verlag Berlin Heidelberg (September 2011) 482–495
- [25] MTV Game : Dance Central. <http://www.dancecentral.com>