



HAL
open science

Modélisation de l'interaction multimodale 3D

Louis-Pierre Bergé

► **To cite this version:**

Louis-Pierre Bergé. Modélisation de l'interaction multimodale 3D. 25ème conférence francophone sur l'Interaction Homme-Machine, IHM'13, Nov 2013, Bordeaux, France. hal-00879635

HAL Id: hal-00879635

<https://inria.hal.science/hal-00879635>

Submitted on 4 Nov 2013

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

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

Modélisation de l'interaction multimodale 3D

Louis-Pierre Bergé

IRIT, Université Paul Sabatier

118 Route de Narbonne

F-31062 TOULOUSE CEDEX 9, France

Louis-Pierre.Berge@irit.fr

RESUME

Dans cet article, nous nous intéressons à la nécessité de réduire l'écart entre les considérations de conception en Informatique Graphique (IG) et en Interface Homme-Machine (IHM) afin de démocratiser l'usage et l'utilisation des environnements virtuels 3D (EV3D) pour des utilisateurs non-experts ou occasionnels. Notre approche consiste en une modélisation de l'interaction multimodale 3D (3DIM). Notre modèle 3DIM contribue à ramener l'ensemble des problématiques 3D et les exigences de l'utilisateur sur la compréhension et la caractérisation du lien entre un ensemble de tâches utilisateur et le graphe de scène de l'EV3D.

Mots Clés

Conception ; modèle ; technique d'interaction 3D ; interaction multimodale avancée ; réalité mixte.

ACM Classification Keywords

D.2.2. Software engineering: Design Tools and Techniques – User interfaces; H.5.2 Information interfaces and presentation: User Interfaces – Theory and methods, User-centered design, I.3.6 Computer Graphics: Methodology and Techniques – Interaction techniques.

INTRODUCTION

Avec l'évolution des technologies, des capacités de calcul et de rendu, les environnements virtuels 3D (EV3D) sont utilisés dans de multiples domaines d'applications incluant des situations industrielles [1], des contextes quotidiens de mobilité [2], le web et même dans les musées [3].

En parallèle, l'interaction Homme-Machine (IHM) évolue rapidement et se complexifie notamment avec l'apparition d'interaction multimodale et d'autres formes avancées comme l'interaction en mobilité, l'informatique ambiante [4], l'interaction spatiale ou encore tangible [5]. Ces formes avancées d'interaction sont aussi multimodales [6] car elles combinent de multiples flux d'entrée et sortie pour utiliser simultanément des entités présentes dans l'environnement physique de l'utilisateur et des entités numériques présentes dans les systèmes

manipulés. L'ensemble des ces interactions multimodales avancées sont potentiellement des solutions interactives plus faciles à appréhender par des utilisateurs novices ou grands-publics comme pour des visites virtuelles de musées [7].

Cependant, lorsqu'il s'agit de concevoir des solutions d'interaction multimodale avancées pour des EV3D, il apparaît un écart important entre les considérations prises en compte par les communautés travaillant en informatique graphique (IG) et celles travaillant en IHM. En effet pour structurer l'espace de conception, la communauté IG s'appuie principalement sur les tâches de Bowman [8] destinées à décrire les actions possibles applicables aux éléments de l'EV3D : navigation, sélection, manipulation, contrôle d'application. Au niveau de l'implémentation de l'interaction, différentes boîtes à outils existent : 3-DIC [9] ou IFFI [10]. Mais jusqu'à présent, aucun consensus n'a été atteint pour définir des méthodes d'implémentation des techniques d'interaction pour un EV3D. En conséquence la réutilisation logicielle des techniques existantes est complexe et une grande variété de techniques d'interaction, difficilement comparables, sont mises en œuvre. A l'opposé, la conception d'IHM et d'interaction multimodale avancée repose fréquemment sur un processus itératif en quatre étapes [11]. Chacune peut profiter de supports méthodologiques bien établis couvrant les différentes facettes de l'IHM [12], comme l'architecture logicielle [13] ou les outils de prototypage [14]. Les approches de conception en IHM ont, du point de vue général, l'intention de comprendre et caractériser les besoins et les tâches de l'utilisateur final afin d'augmenter l'utilisabilité et l'adéquation du système.

Dans ce contexte, notre contribution au développement d'une approche de conception plus rationnelle de technique d'interaction pour les EV3D vise à réduire l'écart entre les approches en IG et en IHM en intégrant au mieux les problématiques 3D et les besoins utilisateur. A ces fins, nous proposons une approche de description abstraite pour servir de langage de référence partagé entre les différents acteurs. Notre approche de description de l'interaction multimodale avancée pour les EV3D permet d'identifier les modalités impliquées en analysant les liens entre les tâches utilisateurs impactant de la 3D et les éléments structurant l'EV3D mis en jeu. Afin de caractériser finement chaque modalité, nous introduisons 3DIM (3D Interaction Modality) : le modèle d'une modalité d'interaction 3D.

Dans les sections suivantes nous présentons notre approche de modélisation, les éléments structurant le modèle 3DIM et nous discutons le positionnement de nos travaux au regard de l'existant et ses apports.

NOTRE APPROCHE DE MODELISATION

L'approche adoptée pour décrire l'interaction multimodale avancée pour les EV3D consiste en premier lieu à identifier les modalités impliquées dans l'interaction 3D. Pour cela nous exploitons deux piliers conceptuels : 1) l'analyse de tâche (issue de l'IHM) afin de déterminer les sous-tâches et les besoins de l'utilisateur, et 2) le graphe de scène (issus de l'IG) qui propose un point de vue sur la structure interne de l'EV3D.

L'analyse des tâches

Grace à une décomposition de l'objectif de l'utilisateur en tâches et sous-tâches, cette analyse vise à comprendre, décrire la logique et la dynamique des tâches et structurer l'activité de l'utilisateur [15]. Nous avons adopté HTA [15] pour représenter le résultat de cette analyse car son formalisme est épuré. En pratique les feuilles de l'arbre des tâches impactant l'EV3D sont proches des tâches de Bowman [8] mais aucun des éléments 3D touchés par les tâches n'est spécifié. C'est là que le graphe de scène interactif rentre en action.

Le graphe de scène interactif

Le graphe de scène décrit et structure les éléments essentiels d'un EV3D en faisant abstraction de la géométrie et de la topologie des objets 3D et en établissant des relations entre eux. Il permet aussi de faciliter la manipulation de ces éléments par les moteurs 3D, leurs rendus, la gestion des collisions, etc. Dans notre approche nous retenons uniquement l'aspect structuration de l'espace numérique 3D que nous apporte le graphe de scène. Cependant, le graphe de scène n'indique pas quels composants de l'EV3D sont susceptibles d'être impactés par l'interaction de l'utilisateur. Pour répondre à ce manque, nous introduisons la notion de « graphe de scène interactif ». Son objectif est de mettre en évidence et de caractériser les objets **manipulables** ou **non-manipulables**, en d'autres termes, les objets impactés ou non par l'une des tâches de l'utilisateur identifiées lors de l'analyse des tâches. Sur la base des éléments décrits dans la norme X3D¹, nous distinguons deux types d'objets manipulables : les **composants 3D** et les **composants de rendu**.

Les **composants 3D** sont les objets géométriques 3D de l'EV3D (maillage, forme géométrique basique). L'interaction de l'utilisateur affecte deux classes d'attributs des composants 3D : les attributs d'*état* (couleur, texture, visibilité) et les attributs de *manipulation* (position, orientation, mise à l'échelle).

Les **composants de rendu** sont les objets prenant part au rendu de l'EV3D (sources de lumières, caméras). L'interaction de l'utilisateur affecte deux classes d'attributs des composants de rendu : les attributs d'*état* (couleur, activation) et les attributs de *manipulations* (position, orientation de la lumière ou du point de vue).

Mise en évidence des modalités d'interaction 3D

Mettre en évidence les modalités d'interaction 3D requises pour atteindre l'objectif utilisateur dans la scène 3D considérée, consiste à relier chaque feuille de l'arbre de tâche aux attributs modifiables impactés dans le graphe de scène interactif. L'ensemble des liens entre l'arbre de tâche et le graphe de scène interactif décrit alors l'interaction multimodale pour un EV3D dans sa globalité.

LE MODELE 3DIM : MODALITE D'INTERACTION 3D

Pour enrichir cette vue globale, nous introduisons le modèle 3DIM décrivant finement chaque modalité.

La définition existante d'une modalité

Nigay définit une technique d'interaction ou une modalité d'interaction comme une paire constituée d'un dispositif et d'un langage d'interaction [16]. Un dispositif est un artefact requis par le système, pour acquérir (dispositif d'entrée) ou fournir (dispositif de sortie) de l'information. Le dispositif fait donc partie du niveau matériel et il utilise la perception et l'observation d'une réalité physique. Un langage d'interaction est un langage utilisé par le système pour échanger des informations, par exemple un langage pseudo-naturel ou un langage de geste physique. En opposition avec le dispositif, le langage d'interaction fait partie du niveau cognitif et propose une représentation, un formalisme que le système doit interpréter. Cette définition, en plus de poser un cadre pour décrire l'interaction, permet de structurer le raisonnement au cours d'un processus de conception. En basant notre modèle sur cette définition, nous gardons ces avantages en les adaptant à la modalité d'interaction en 3D. Nigay adopte un point de vue côté système et applique sa définition d'une modalité en entrée et en sortie [16]. Dans la suite de ce papier, le modèle 3DIM est limité à l'adaptation d'une modalité d'entrée telle que définie par Nigay.

Adaptation de la définition d'une modalité

En exploitant la définition de Nigay, nous introduisons et justifions progressivement la structure et les concepts formant le modèle 3DIM résumé dans la Figure 1. Pour commencer, nous établissons que : « *un utilisateur manipule une modalité d'interaction 3D qui impacte le système* ». C'est la façon la plus grossière de décrire l'interaction de l'utilisateur avec un EV3D.

Dans le contexte de l'interaction multimodale avancée pour un EV3D, il apparaît qu'une modalité d'interaction 3D est constituée de deux aspects complémentaires : la **modalité physique** regroupe les éléments de l'environnement de l'utilisateur jouant un rôle dans la modalité : l'utilisateur, ses actions physiques, son comportement et

¹ <http://www.web3d.org/x3d/>

les objets physiques ; la **modalité 3D** regroupe les dispositifs spécifiques pour la 3D et les considérations logicielles pour conduire l'interaction avec la 3D. Une façon plus fine de décrire l'interaction de l'utilisateur avec un EV3D devient : « un utilisateur manipule une **modalité physique** qui affecte une **modalité 3D** qui impacte le système ».

La **modalité 3D** étant reliée au système, la définition d'une modalité selon Nigay est applicable. La modalité 3D peut être décomposée en un *dispositif 3D* et un *langage 3D*. Le *dispositif 3D* capte des propriétés physiques comme la position 3D ou l'orientation 3D fournie par un tracker. Le *langage 3D* est un algorithme qui adapte les données captées pour le système et permet d'effectuer les actions adéquates. Nous considérons le *dispositif 3D* comme le **périphérique d'entrée** utilisé dans la modalité 3D et le *langage 3D* comme le **comportement 3D** qui adapte les données pour les EV3D.

Il semble opportun d'appliquer également la définition de Nigay pour affiner la **modalité physique**. Prenons un exemple d'interaction tangible où l'utilisateur manipule des *objets physiques*. Ces derniers appartiennent au niveau physique car ils sont observables et manipulables. Avec la définition de Nigay, ils peuvent être considérés comme le dispositif d'une modalité. De plus, pour manipuler l'objet physique, l'utilisateur doit planifier ses mouvements et ses actions. Il doit passer par une étape de cognition et d'expression de son mouvement. Par rapport à Nigay, l'action physique et les mouvements sont considérés comme un *langage physique* produit par l'utilisateur pour manipuler l'*objet physique*. La **modalité physique** se décompose donc en **action physique** et en **objet physique**. Dans le cas d'une interaction non tangible, aucun **objet physique** n'est présent, mais le *langage physique* reste : l'utilisateur doit effectuer directement ses **actions physiques** sur le **périphériques d'entrée** de la **modalité 3D**.

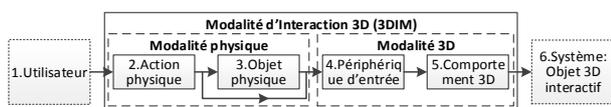


Figure 1. Modèle de l'interaction multimodale 3D

Enfin, le système impacté par une modalité d'interaction 3D est au final l'ensemble des objets 3D interactifs touchés par une tâche de l'utilisateur. Ainsi la façon la plus raffinée de décrire une modalité d'interaction 3D est la suivante : « un utilisateur effectue une **action physique** qui affecte un **objet physique**. Les attributs de celui-ci sont captés par un **périphérique d'entrée**. Un **comportement 3D** interprète les données acquises pour impacter un objet 3D interactif » (voir Figure 1).

Il en résulte donc que le modèle 3DIM propose un découpage en 6 blocs afin de caractériser le lien entre une feuille de l'arbre de tâche de l'utilisateur et un élément du graphe de scène interactif.

Les attributs principaux contenus dans chaque bloc sont : **Utilisateur** : les parties du corps impliquées dans l'interaction, le niveau d'expertise de l'utilisateur (expert, avancée ou débutant) ainsi que les dispositifs en sorties requis (HMD, écran,...).

Action physique : des mouvements discrets ou continus (mouvement de pression sur un bouton vs. translation de la main), le type de mouvement (translation, orientation, geste), mouvement libre ou contraint (mouvement dans l'espace vs mouvement sur une surface tactile).

Objet physique : les propriétés visuelles (forme, taille, couleur, texture), les contraintes physiques avec d'autres objets physiques ou l'environnement.

Périphérique d'entrée : le type d'information captée (son, image, position, force, angle, déformation, état), les degrés de libertés captés et utiles pour la tâche.

Comportement 3D : le nom de la transformation logicielle (changement de repère, détection d'une collision), le type de données en entrée et en sortie.

Objet 3D interactif : les propriétés visuelles (forme, couleur, texture), les attributs du graphe de scène interactif modifiés.

DISCUSSIONS

Les attributs et le découpage de 3DIM s'appuient sur des approches de la littérature, tant en IG qu'en IHM, relatives à la modélisation et la conception d'interaction avancées et d'EV3D interactif. Le découpage a été amené par l'adaptation de la définition de Nigay et les attributs évoqués sont également présents dans des modèles existant. La Figure 2 résume les parties de 3DIM couvertes par des modèles de la littérature et positionne ainsi 3DIM au regard de l'existant.

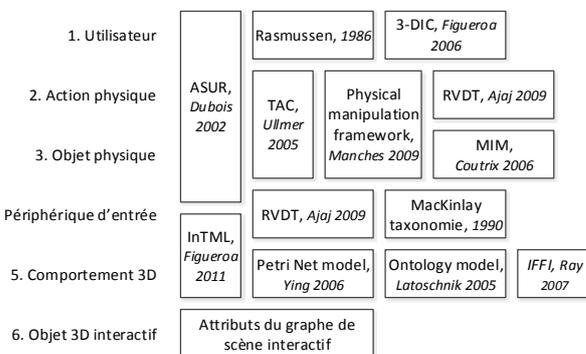


Figure 2 : Positionnement de 3DIM dans l'existant

Nous avons appliqué notre approche (analyse des tâches, graphe de scène interactif et modélisation avec 3DIM des liens) sur de nombreux exemples de la littérature. Il en ressort que nos travaux permettent d'identifier et de caractériser certaines différences entre les techniques d'interaction. 3DIM souligne ces différences au niveau des sources (tâches) et cibles (graphe de scène interactif) des liens : une modalité peut impacter un attribut d'état ou un attribut de manipulation d'un objet 3D interactif.

3DIM attire l'attention sur les différents liens constituant l'interaction multimodale 3D : identification d'éléments

impliqués dans un des six blocs qui peut être partagé par une autre solution interactive ou par d'autres tâches ; identification du nombre d'actions physiques et de périphériques d'entrée en se basant sur le nombre de connexions entre blocs.

Ces critères de comparaison peuvent à partir d'une modélisation existante, permettre la génération de nouvelle solution : par exemple, en ajoutant un objet physique (bloc 3), en limitant l'ajout d'élément 3D perturbant le réalisme de l'EV3D (bloc 5) ou en se posant des questions sur la cohérence et la complétude des tâches présentes dans la solution.

Enfin 3DIM offre une vue globale sur l'activité de l'utilisateur avec un EV3D sans se focaliser sur une seule des tâches 3D atomiques identifiées par Bowman [8]. 3DIM permet de mettre l'accent sur des aspects spécifiques à la technique d'interaction afin de l'optimiser par rapport au contexte général de l'application. Par exemple un concepteur peut se focaliser sur les actions physiques (bloc 2) réalisées et voir le traitement déclenché au niveau du comportement 3D (bloc 5). De plus, à travers les attributs liés aux contraintes physiques de l'objet physique (bloc 3) et du périphérique d'entrée (bloc 4), 3DIM apporte un appui sur la façon dont l'utilisateur interagit avec l'espace qui l'entoure, les objets qu'il utilise afin d'exploiter l'interaction pas seulement sur l'objet physique mais aussi autour de celui-ci.

CONCLUSION

La contribution de ce papier est une méthode support à la conception d'interaction multimodale pour EV3D. En structurant l'approche autour de l'analyse des tâches et du graphe de scène interactif, nous proposons de modéliser chaque modalité d'interaction 3D autour de 6 blocs (le modèle 3DIM). En regroupant ainsi les considérations de conception des communautés en IG et en IHM nos travaux contribuent 1) à obtenir une meilleure compréhension des éléments fondamentaux d'une technique d'interaction 3D ; 2) à identifier systématiquement les questions relatives à la conception d'une nouvelle interaction pour la 3D et 3) à faciliter le développement d'interactions pour la 3D mieux adaptées à des utilisateurs non-experts ou occasionnels.

Afin d'enrichir notre modèle 3DIM, nous prévoyons de caractériser plus finement les liens entre blocs en ajoutant, par exemple, des attributs de relation spatiale. De même ajouter des propriétés ou des opérateurs temporels entre les modalités, comme les propriétés CARE [16] peut être utile pour affiner la description. Enfin notre modèle peut servir de base pour établir un langage formel de description de l'interaction en 3D qui faciliterait l'implémentation logicielle de l'interaction et l'intégration dans une plateforme de prototypage rapide.

BIBLIOGRAPHIE

- [1] A. Gomes De Sá, "Virtual reality as a tool for verification of assembly and maintenance processes," *Comput. Graph.*, vol. 23, no. 3, pp. 389–403, 1999.
- [2] D. Magliocchetti et al., "I-MOVE: towards the use of a mobile 3D GeoBrowser framework for urban mobility decision making," *International Journal on Interactive Design and Manufacturing*, 205–214, 2012.
- [3] L. Pecchioli, et al., "ISEE: Information access through the navigation of a 3D interactive environment," *Journal of Cultural Heritage*, 287–294, 2011.
- [4] H. Ishii, et al., "ambientROOM: integrating ambient media with architectural space", *CHI*, 1998, 173–174.
- [5] C. Williams, et al., "TZee: exploiting the lighting properties of multi-touch tabletops for tangible 3d interactions" *CHI '11*, 2011, 1363–1372.
- [6] B. Dumas, et al., *Multimodal Interfaces: A Survey of Principles, Models and Frameworks*, Springer Berlin Heidelberg, 2009, 3–26.
- [7] E. Hornecker, "Interactions around a contextually embedded system," in *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction - TEI '10*, 2010, 169–176.
- [8] D. A. Bowman, et al., *3D User Interfaces: Theory and Practice*. 2004, p. 478.
- [9] P. Figueroa, et al.; "A Conceptual Model and Specification Language for Mixed Reality Interface Components," *IEEE Virtual Reality*, 2006, 4–11.
- [10] A. Ray, D. Bowman "Towards a system for reusable 3D interaction techniques," *VRST*, 2007, 187–190.
- [11] E. Dubois, et al., "Design and Evaluation of Mixed Interactive Museographic Exhibits," *International Journal of Arts and Technology*, 4, 408–441, 2011.
- [12] S. Basnyat, et al., "Complexity of Design in Safety Critical Interactive Systems: Gathering, Refining, Formalizing Multi-Type and Multi-Source Information while Ensuring Consistency, Reliability, Efficiency and Error-Tolerance," 2005, 45–58.
- [13] "A metamodel for the runtime architecture of an interactive system" *ACM SIGCHI*, 24,1, 32–37, 1992.
- [14] M. Serrano, et al. "A three-dimensional characterization space of software components for rapidly developing multimodal interfaces," *IMCI*, 2008, 149–156.
- [15] J. Annett, "Hierarchical Task Analysis," in in *The Handbook of Task Analysis for Human-Computer Interaction*, Diaper and Stanton, 2004, p. 667
- [16] L. Nigay and J. Coutaz, "Multifeature Systems: The CARE Properties and Their Impact on Software Design," *Multimedia Interfaces: Research and Applications, chapter 9*, 1997