

Gestaction 3D: combiner commandes gestuelles et interactions 3D

Diane Lingrand^{1*} Sébastien Meunier¹ Philippe Renevier¹ Anne-Marie Pinna-Déry¹ Julien Soula^{2 †}
Michel Riveill¹

¹ Équipe Rainbow
Laboratoire I3S (UNSA/CNRS) UMR 6070
B.P. 121 - 06903 Sophia Antipolis - FRANCE
<http://www.i3s.unice.fr/~lingrand>

² Salle Immersive le Corbusier
CSTB
B.P. 209 - F 06904 Sophia Antipolis - FRANCE

Résumé

Nous proposons un prototype permettant d'interagir de façon gestuelle avec des applications. Nous nous intéressons particulièrement aux déplacements et déformations d'objets 3D mais avons bâti ce prototype afin de permettre également les interactions 2D. Dans une étude utilisateurs préliminaire, basée sur un magicien d'Oz, nous avons tenté de dégager des gestes intuitifs. Nous avons ensuite développé ce prototype afin de pouvoir faire de réelles études.

Keywords: Interactions gestuelles, interactions 3D.

1 Introduction

L'informatique s'est depuis longtemps intéressée aux représentations 3D du monde réel, malgré les limitations imposées par nos écrans et périphériques d'entrée courant (clavier, souris...) essentiellement conçus pour des représentations et une navigation dans un espace 2D. Avec la banalisation des cartes graphiques permettant une accélération matérielle 3D sur nos ordinateurs et l'émergence de nouveaux périphériques de visualisation 3D, les représentations 3D deviennent courantes, que ce soit dans les jeux, les applications ludo-éducatives, les logiciels de modélisation ou de simulation et même les catalogues de produits. Ces technologies permettant de redonner à l'utilisateur une perception du monde 3D, plus proche de sa réalité quotidienne, sont sans aucun doute promises à se généraliser davantage. On peut d'ailleurs voir apparaître depuis peu des tentatives d'applications exploitant la troisième dimension dès le système de fenêtrage 3D (remplaçant nos bureaux KDE, Gnome ou Windows).

Cependant, claviers et souris sont aujourd'hui indissociables de nos ordinateurs fixes ou portables et, même s'ils permettent effectivement de naviguer et d'interagir avec les mondes 3D, ils ne sont pas adaptés à ceux-ci puisque conçus pour des applications 2D. Il suffit, pour s'en convaincre, de construire un objet 3D à l'aide d'un modèleur 3D (tels Blender, Moonlight, 3DSmax ou Maya) : chaque modification dans l'espace doit être effectuée indirectement, en travaillant dans les différents sous-espaces de projection.

Depuis plusieurs années, de nombreuses tentatives de périphériques 3D ont vu le jour [3]. Pourtant, et malgré la difficulté de navigation dans les mondes 3D à partir d'un périphérique 2D, aucun d'entre-eux n'a rencontré un succès comparable à celui de la souris. Construits la plupart du temps pour des applications précises, il n'existe pas de périphérique 3D générique.

Gestaction 3D a pour but la construction d'un prototype évolutif pour tester des interactions 3D. Nous nous attachons en premier lieu à des interactions basées sur les mouvements des mains, très utilisés déjà dans le langage parlé. Le prototype présenté ici a été construit à l'aide de composants bon marché et couramment disponibles sur le marché afin d'être aisément adaptable, modifiable et de donner

la possibilité de construire plusieurs prototypes utilisables et diffusables.

Le présent article fait suite aux travaux déjà publiés [7] et [17].

2 Travaux antérieurs

La dextérité de nos mains nous permet d'exécuter dans la vie courante des tâches avec précision et rapidité. Pour des tâches bien spécifiques, nous avons de bonnes performances d'apprentissage (travaux manuels, instruments de musique, ...). Les mouvements des mains seules ou utilisant différents objets ou outils ont été utilisés en IHM grâce aux techniques de Vision par Ordinateur depuis plusieurs années [22, 19, 2, 4, 3, 27, 18, 8, 13].

Certains systèmes utilisent uniquement les mains nues. Parmi eux, GestureVR [22] permet des interactions dans l'espace reconnues par deux caméras. Trois commandes gestuelles peuvent être reconnues. Deux doigts et leur orientations sont détectés. Ce système est utilisé pour du dessin 3D, la manipulation d'un robot et la navigation dans une scène 3D.

Les approches utilisant les techniques de Vision par Ordinateur pour la détection de mouvements libres de mains nues sont toujours au stade de recherche [18, 8] même si certains résultats sont prometteurs [27]. Limiter les mouvements possibles à un petit nombre facilite à la fois la tâche de l'utilisateur (un grand nombre de mouvements impliquant plus d'apprentissage) et celle du module de reconnaissance.

Ainsi, de nombreux travaux utilisent des objets déplacés par les mains, plus facilement détectables. La *Magic Table* [2] utilise la détection de jetons colorés déplacés sur une table par des utilisateurs afin d'interagir avec des objets physiques et virtuels disposés sur ce plan.

La baguette magique visuelle (*VisionWand*) [4] est une baguette dont les deux extrémités sont vivement colorées afin d'être aisément détectées et qui est utilisée pour l'interaction avec des objets 2D et la sélection de menus. Le système permet la reconnaissance de neuf commandes gestuelles.

Les interactions 3D ont également été étudiées dans le cas particulier de visualisateurs volumiques 3D [12], permettant une visualisation et une manipulation tout autour de l'objet étudié. Malheureusement, ce type de visualisateur est encore peu répandu et coûteux.

Les motivations de ces différents travaux concernent la manipulation d'objets [2, 4, 24], le pointage d'objets ou sélection de menus [19], la reconnaissance de gestes pour des commandes gestuelles [25] ou encore la reconnaissance du langage des signes [9].

Smith et ses collègues [24] ont explicités les contraintes permettant la manipulation d'objets 3D en utilisant des périphériques 2D. Ils utilisent leur système pour la manipulation d'objets dans une pièce (chaises, tables, ...).

Moeslund et ses collègues [19] ont développé un système de pointage en utilisant à la fois des techniques de vision par ordinateur et un tracker magnétique monté sur des lunettes stéréo.

*e-mail: diane.lingrand@unice.fr

†e-mail: julien.soula@cstb.fr

Dans les différentes façons d’interagir avec un monde 3D, on se rend compte qu’il existe deux types d’interaction : des interactions réellement spatiale, et des interactions étant plutôt des commandes pouvant être apparentée à une sélection de mode ou de menu, un pointage, ... et qui peuvent être soit linéaire, soit plan. On s’est alors intéressé aux commandes gestuelles.

Pour des commandes gestuelles, on préfère des gestes 2D car leur mémorisation est plus simples. Il est également préférable pour l’utilisateur d’avoir une visualisation de sa gesture [14]. Un geste plan peut être repéré par une caméra seule dirigée vers l’utilisateur, ce qui limite les déplacements de l’utilisateur, ou bien dirigée vers l’écran repérant les traces d’un pointeur laser.

Utiliser un pointeur laser pour interagir devant un écran de projection n’est pas une idée nouvelle. Lors de présentation vidéo-projetée, il est courant de disposer d’un pointeur laser. Une caméra est également présente dans les salles permettant la téléconférence (sinon, des webcams bon marchés peuvent être utilisées). Les interactions utilisant un pointeur laser permettent ainsi, à peu de frais, de se détacher du couple clavier/souris lié au bureau.

Kirstein et Müller [15] utilisent le pointeur laser comme souris à un bouton. Le “clic” est activé lorsque le pointeur reste fixe sur une zone prédéterminée (correspondant à un objet graphique, à un bouton, ...) et est désactivé lorsque le pointeur laser est éteint.

Lapointe et Godin utilisent un écran translucide et une caméra regardant l’écran [16] et ont étudié les caractéristiques de fiabilité, précision de la position et latence de détection du pointeur laser sur l’écran.

En utilisant les zones périphériques de l’écran et en se basant sur la technique de *goal crossing*, les auteurs de [23] définissent différentes commandes au pointeur laser afin de déterminer des actions pour naviguer dans des présentations ou visualiser des cartes géographiques.

Dans ce travail, nous cherchons à permettre l’étude des interactions avec des objets 3D en ce qui concerne leur déformations et leur manipulation ainsi que leur visualisation depuis différents points de vue. Notre prototype se veut suffisamment générique afin de permettre des études en salle immersive (sombre), dans un bureau standard ou même dans une salle de cours. Nous avons construit un périphérique passif à base de gants confortables aux extrémités colorées. On s’intéresse à un espace d’interaction moins limité que celui de systèmes existants tels que les gants P5 de la société *Essential Reality* mais forcément limité par le champ de vue des caméras.

3 Architecture générale

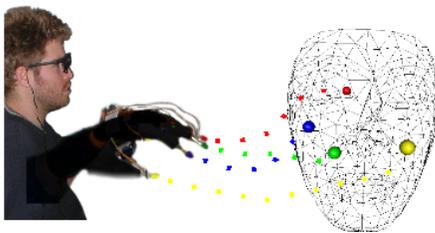


FIG. 1: Interagir avec les doigts sur un maillage 3D : chaque doigt peut interagir avec le maillage selon différents modes.

Le gant est fabriqué à partir d’un sous-gant de soie noir pour le ski sur lequel sont cousues aux extrémités des doigts des diodes de couleurs différentes. L’acquisition stéréo est effectuée au moyen de deux webcams Philips Fun II dans une configuration permettant de simplifier la géométrie épipolaire en géométrie standard[11]. Cela permet de d’exprimer la profondeur des points de façon simple.

Il est bien connu qu’en Vision par Ordinateur, l’extraction des points d’intérêt et la mise en correspondance entre deux vues sont les étapes les plus critiques car difficiles et compromettant la reconstruction en cas d’erreur. Dans notre cas, l’utilisation de diodes colorées permet une localisation aisée et robuste de celles-ci dans les deux images. La mise en correspondance est basée sur la teinte évitant ainsi les difficultés de mise en correspondance [28]. Un étalonnage est nécessaire afin d’exploiter l’intégralité du champ de vision des webcams pour les déplacements souhaités dans la scène 3D. Enfin, la profondeur est directement déterminée en fonction de la disparité 2D entre les points ainsi mis en correspondance.

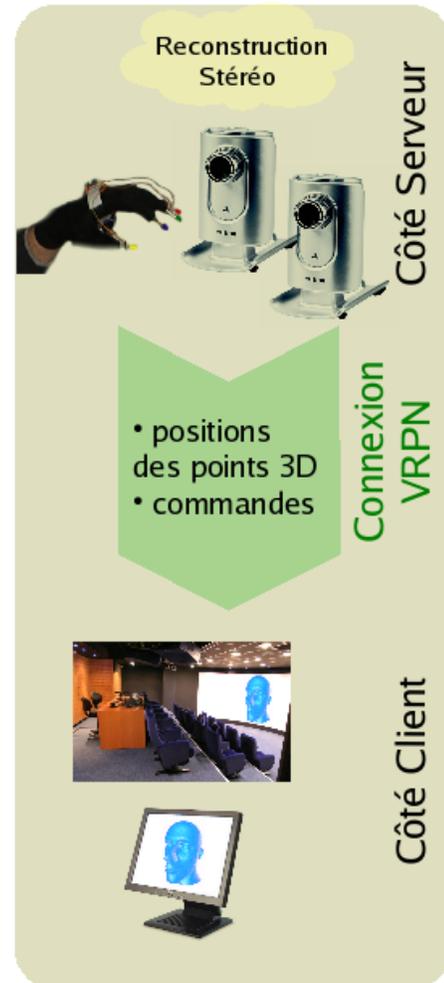


FIG. 2: Client et serveur.

La communication entre la machine d’acquisition et celle de rendu est effectuée à l’aide de la bibliothèque VRPN[26] afin d’intégrer différents systèmes de réalité virtuelle. Parmi les différents modèles déformables [20], nous avons choisi l’implémentation CGAL[10] proposant de nombreuses opérations utiles sur les maillages comme le raffinement de maillage. La scène 3D est gérée en utilisant OpenGL et la visualisation est effectuée en utilisant soit les lunettes stéréo VRex sur un moniteur cathodique en environnement bureau, soit les lunettes 3D Crystal Eyes dans la salle de réalité virtuelle du CSTB afin de rendre possible l’interaction avec l’utilisateur.

Nous proposons différents modes d’interactions avec les

maillages : sélection de sommets à l'aide d'un doigt, sélection d'arêtes avec deux doigts et sélection de face par un doigt. Différentes opérations telles que le déplacement d'un sommet ou d'une face, l'extrusion, l'ajout ou la destruction d'une arête sont présentées.



FIG. 3: Utilisation de Gestaction3D dans la salle immersive du CSTB. (a) L'utilisateur voit les objets en 3D. Sur cette photo, les deux vues sont superposées. (b) L'utilisateur utilise son gant devant les webcams. (c) Sur un modèle plus complexe, des déformations sont appliquées.

4 Mouvements 3D

4.1 Détection des positions 3D

Deux types de gants ont été construits. Le premier (figure 4) possède l'avantage de ne comporter qu'une seule diode par doigt. Des occlusions peuvent arriver lors de certains mouvements. Le second type de gant (figure 5) dispose de plusieurs diodes par doigts ce qui permet une meilleure détection lors de mouvements variés. Cependant, un filtrage est nécessaire afin d'éviter des discontinuités de trajectoire (en fonction des diodes visibles).

Dans chaque image, un premier seuillage sur l'intensité permet d'éliminer les zones sombres de l'image. Seuls les pixels restants sont convertis dans les systèmes de couleur RGB ou HSV. Le seuillage des composantes RGB est suffisant en milieu sombre. En éclairage normal (de type bureau), le seuillage sur les composantes HSV est plus efficace. Les différents seuils sont à ajuster en fonction du type de gant utilisé (les valeurs de chrominance varient) et de l'éclairage ambiant. Le réglage n'est pas à refaire en cours d'expérimentation.

Les positions 3D sont déterminées à l'aide des deux positions des diodes détectées dans les images issues des deux webcams. Les



FIG. 4: Premier prototype de gant : une diode par doigt. Des occlusions peuvent avoir lieu.



FIG. 5: Second prototype de gant : cinq diodes par doigt. On réduit considérablement les occlusions des diodes.

cameras sont physiquement placées de façon à ce que leur plans focaux soient alignés afin que la géométrie épipolaire soit dans une configuration de géométrie standard (figure 7). Théoriquement, ceci est approximatif mais les calculs à effectuer sont très simples et donc très rapides.

Nous n'avons pas encore effectué une évaluation précise de la reconstruction. Actuellement, le prototype existant permet de manipuler des sommets de façon satisfaisante lors de déformation de maillage. Cependant, une évaluation plus précise de la reconstruction est importante si on s'intéresse à des déformations précises ou bien à des gestes précis pour spécifier l'orientation et la position de la scène globale ou de la caméra.

Nous avons réalisé une évaluation légèrement quantitative en utilisant un cube de type *Keenex* (figure 8). L'utilisateur doit suivre une arête avec un doigt spécifique et la trajectoire 3D est ensuite examinée. On a ainsi pu établir que la reconstruction est généralement peu précise (encore moins que le protocole expérimental).

La théorie de la Vision par Ordinateur nous enseigne que la configuration des caméras en géométrie standard peut être obtenue soit par un positionnement précis des caméras (notre positionnement est actuellement effectué manuellement), soit en rectifiant les images. L'une de ces deux pistes doit être étudiée afin d'améliorer notre reconstruction actuelle tout en conservant le traitement en temps réel des images.

4.2 Modes d'interaction

Nous avons mis en place plusieurs modes d'interaction. Dans le mode *par sommet*, lorsque qu'une extrémité de doigt se trouve dans le voisinage d'un sommet, ce dernier lui est attaché et suit ses mouvements. Un mouvement brusque en retrait du sommet détache celui-ci du doigt. Dans le mode *par face*, c'est une face du polygone qui est attachée à un doigt et qui déplace son barycentre en fon-

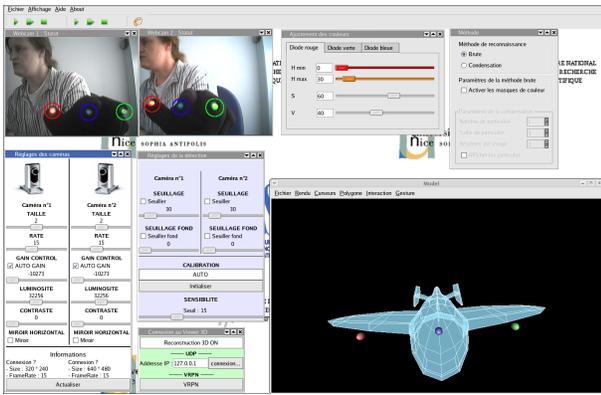


FIG. 6: Application développée pour la déformation de maillage : vue des deux webcams et les différents réglages associés. La vue en bas à droite donne la position reconstruite des extrémités des doigts dans une scène 3D comportant un avion.

tion des mouvements du doigt. Un mode de *destruction de sommet* permet de détruire un sommet ; le maillage est alors recomposé en fonction des sommets restants. Parallèlement, le mode *destruction d'arête* permet de détruite l'arête sélectionnée. Le mode *division de face* ajoute une arête entre les deux sommets sélectionnés sur une face, divisant ainsi la face en deux faces. Enfin, un mode *d'extrusion* permet de modifier la structure.

L'activation d'un mode d'interaction peut se faire par sélection d'option dans les menus de l'interface ou par les raccourcis clavier associés. Cependant, l'usage du clavier ou de la souris avec des gants est peu pratique. De plus, cela nécessite d'être proche du couple clavier/souris. C'est pourquoi des commandes gestuelles ont été introduites dans notre système.

5 Gestures

Les gestes sont des commandes gestuelles qui viennent en remplacement de clic souris ou raccourcis plus généralement utilisés dans les interfaces WIMP.

Dans [17], nous avons présenté des interviews utilisateurs concernant les gestes intuitifs que ceux-ci souhaiteraient pour différentes actions telles que la sélection d'objet (un ou plusieurs), les déplacements d'objets (rotation, translation) et les déformations d'objets. Les interviews étaient sans contraintes (choix des mains, doigts, espace de mouvement possible, ...). Pour certaines actions comme par exemple la sélection, les interviews concordent sur un geste (figure 9) tandis que pour d'autres commandes comme par exemple l'orientation, les gestes sont très variés et il n'est pas possible de dégager un résultat (figure 10).

Nous avons alors décidé de distinguer différents modes d'actions et de passer de l'un à l'autre par des commandes gestuelles ou gestes, principalement dans un plan parallèle au plan focal des caméras.

5.1 Méthode de reconnaissance

Pour des raisons de mémorisation de gestes, notamment, on va limiter le nombre de commande gestuelles. Cela nous permet également de discrétiser grossièrement les mouvements et d'augmenter ainsi la robustesse. Le principe est similaire à celui utilisé dans *Mozilla All-In-One Gestures*¹. Tout déplacement est interprété dans un plan orthogonal à la caméra et est discrétisé selon l'une des

¹<https://addons.mozilla.org/firefox/12/>

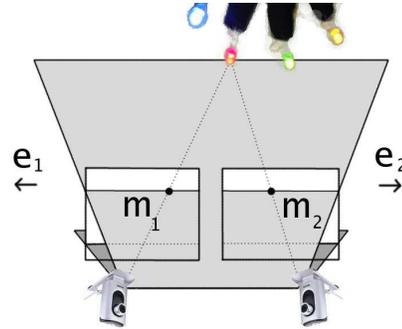


FIG. 7: Géométrie standard des caméras.

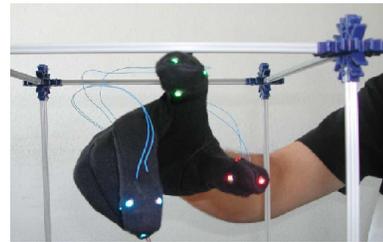


FIG. 8: Cube en Keenex de test : l'utilisateur suit les arêtes avec une extrémité de doigt.

quatre directions : haut, bas, gauche et droite. Toute suite de mouvements identiques est remplacée par un seul mouvement (voir figure 11).

Dans le cas de *Mozilla All-In-One Gestures*, la reconnaissance de gestes est commandée par le clic droit de la souris. La reconnaissance s'arrête dès que le bouton est relâché. Dans le cas de l'utilisation du laser, une zone de l'espace visible par la caméra est utilisée pour la détection des gestes. La reconnaissance est alors activée par la présence ou non du laser. Pour les gants en 3D, il est difficile de spécifier une zone de l'espace réservée à la détection de gestes : on veut pouvoir utiliser tout l'espace pour les autres actions et on veut pouvoir déclencher une commande depuis la position courante de la main. On a alors choisi d'amorcer une gesture en joignant pouce et majeur et de la terminer en écartant ces doigts. La distance seuil entre ces doigts est paramétrable en fonction du gant utilisé. Cette distance est bien-sûr calculée en 3D. Étant un acte volontaire de l'utilisateur, des occlusions sont peu probables.

5.2 Utilisation des gestes

Pour relier les gestes à des applications, nous avons choisi de relier des gestes à des raccourcis clavier. Les applications clientes sont des applications fonctionnelles sans les gestes mais dont les actions sont commandées classiquement par une interface WIMP (menus, items, boutons et, alternativement, raccourcis claviers).

Ainsi, il nous est possible de relier notre module de reconnais-



FIG. 9: Sélection : tout le monde est d'accord !



FIG. 10: Cas de la rotation de la caméra : différents utilisateurs, différents gestes !

sance de gestes à une applications 3D ou 2D. Nous donnons des exemples dans la section suivante.

6 Applications visées

Nous avons visé à la fois les applications en deux dimensions et les applications nécessitant des interactions en trois dimensions.

Nous nous sommes intéressés principalement à deux domaines d'applications.

6.1 Application utilisant des interactions 2D

Nous avons utilisé notre module de reconnaissance de gestes 2D pour une application simple de gestion de photographies. La visualisation de grandes collections de photos a déjà été étudiée par de nombreux auteurs, notamment [1, 6, 21]. Nous ne nous sommes

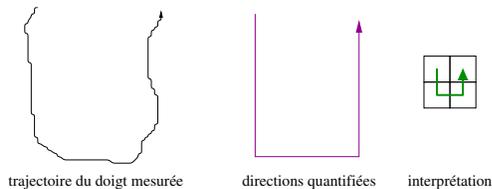


FIG. 11: Reconnaissance de geste : les directions sont quantifiées selon les quatre directions : haut, bas, gauche et droite, puis les redondances éliminées.

pas intéressés à cette partie mais plutôt aux interactions possibles selon différentes modes de visualisation. Notre application permet d'interagir directement avec l'application en utilisant un pointeur laser, outil largement utilisé lors de présentation. Nous présentons ci-dessous les gestes utilisés pour lesquels les zones de début et de fin de geste ont été prises en considération. Des évaluations utilisateurs précises sont à prévoir.

image suivante

image précédente

zoom +

zoom -

rangement dans un répertoire

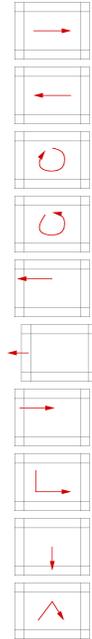
visualisation d'un dossier favori

chargement des images d'un dossier

entrer dans le dossier sélectionné

suppression d'un dossier

changement de mode de visualisation



6.2 Application en Synthèse d'Images.

Il existe de nombreux logiciels de modélisation 3D. Ces logiciels présentent généralement des interfaces avec plusieurs projections entre lesquelles il faut naviguer afin d'exprimer une forme volumique 3D. Un premier plus serait une vue 3D, chose aisée avec les *displays* actuels (lunettes stéréoscopiques et écrans 3D). Mais l'utilisateur interagit toujours avec la souris et le clavier. Cela peut être efficace pour une navigation dans une scène mais est moins adapté au modelage d'une forme.

On se propose d'étudier les interactions possibles entre un périphérique 3D et un maillage. Pour cela, un jeu de gestes a été déterminé pour les commandes suivantes :

mode vertex		détruit vertex	
mode face		extrusion	
détruit arête		divise face	
orientation		translation	

Une évaluation utilisateurs de ces gestes et surtout de l'utilisabilité des gestes combinées aux mouvements 3D est nécessaire. Il conviendra d'étudier également d'autres alternatives, notamment concernant le nombre de doigts utilisés, le nombre de mains ainsi que d'autres modalités.

6.3 Application en segmentation d'images 3D

Il est clairement établi qu'il n'existe pas de méthode de segmentation d'images universelle et totalement automatique. Les méthodes, ou plutôt leurs paramètres doivent être adaptés au cas par cas, ou plutôt pour des familles d'images : c'est une calibration des méthodes au type d'images. Des algorithmes présentant des avantages quant à la représentation du résultat sont sensibles à l'initialisation (l'éternel problème des minima locaux et du minimum global). D'autre part, les résultats sont rarement suffisamment précis pour des applications médicales et certains praticiens retouchent les segmentations à la main afin d'améliorer le résultat. Le passage aux images 3D alourdit considérablement ces post-traitements (retouches coupes par coupes), sans parler des images 4D (généralement des séquences d'images 3D).

Les méthodes actuelles d'intervention sur les segmentations sont donc très fastidieuses, coûteuses en temps et très peu ergonomiques : il s'agit le plus souvent de retraiter les images volumiques coupe par coupe et pixel par pixel. Nous proposons d'améliorer ces interactions en les effectuant directement en trois dimensions avec des méthodes de visualisation et d'interactions adaptées. Ceci constitue une des perspectives de cette étude.

7 Conclusion et Perspectives

Ce prototype réalisé et effectif nous permet d'interagir en 3D avec des objets et des maillages selon des modes simples. Des améliorations sont à prévoir notamment en ce qui concerne la précision de la reconstruction 3D nécessaire. Il est maintenant nécessaire d'effectuer de réelles études d'ergonomie afin de déterminer si notre approche combinant commandes gestuelles et gestes 3D convient ou bien s'il est préférable d'utiliser une autre main ou bien encore une autre modalité (parole, ...).

Acknowledgements

Nous remercions les étudiants qui ont participé aux interviews utilisateurs ainsi toutes les personnes qui ont encouragé ce travail à l'École Polytechnique Universitaire de Nice-Sophia Antipolis, au laboratoire I3S et au CSTB Sophia Antipolis.

Références

- [1] B. B. Bederson. PhotoMesa : A Zoomable Image Browser using Quantum Treemaps and Bubblemaps. In *ACM Conference on User Interface and Software Technology (UIST'01)*, pages 71–80, Orlando, Florida, Oct. 2001. ACM Press.
- [2] F. Bérard. The Magic Table : Computer-Vision Based Augmentation of a Whiteboard for Creative Meetings. In *IEEE Workshop on Projector-Camera Systems, in conj. with ICCV (PROCAM)*, 2003.
- [3] D. A. Bowman, E. Kruijff, J. J. LaVIOLA, and I. Poupyrev. *3D User Interfaces : Theory and Practice*. Addison-Wesley, 2004.
- [4] X. Cao and R. Balakrishnan. VisionWand : Interaction Techniques for Large Displays Using a Passive Wand Tracked in 3D. In *Symposium on User Interface Software and Technology (UIST)*, pages 173–182, Vancouver (Canada), 2003. ACM Press.
- [5] A. Celentano and P. Mussio, editors. *Working Conference on Advanced Visual Interfaces*, Venezia, Italy, May 2006. ACM-SIGCHI, ACM-SIGMM, SIGCHI Italy, ACM Press.
- [6] K. Cooper, O. de Bruijn, R. Spence, and M. Witkowski. A Comparison of Static and Moving Presentation Modes for Image Collections. In Celentano and Mussio [5], pages 381–388.
- [7] P. Cuisinaud, J. Soula, and D. Lingrand. Gestaction 3D. In *17ème Conférence Francophone sur l'Interaction Homme-Machine (IHM)*, pages 345–346, Toulouse (France), Sept. 2005. AFIHM, ACM Press.
- [8] K. G. Derpanis. A Review of Vision-Based Hand Gestures. internal report, Centre for Vision Research, York University (Canada), 2004.
- [9] K. G. Derpanis, R. P. Wildes, and J. K. Tsotsos. Hand Gesture Recognition within a Linguistics-Based Framework. In T. Pajdla and J. Matas, editors, *European Conference on Computer Vision*, volume LNCS 3021, pages 282–296, Prague (Czech Republic), May 2004. Springer.
- [10] A. Fabri, G.-J. Giezeman, L. Kettner, S. Schirra, and S. Schönherr. On the Design of CGAL a Computational Geometry Algorithms Library. *Software - Practice and Experience*, 11(30) :1167–1202, 2000.
- [11] O. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
- [12] T. Grossman and R. Balakrishnan. An Evaluation of Depth Perception on Volumetric Displays. In Celentano and Mussio [5], pages 193–200.
- [13] A. Jaimes and N. Sebe. Multimodal Human Computer Interaction : A Survey. In *IEEE International Workshop on Human Computer Interaction in conjunction with ICCV*, Beijing (China), Oct. 2005.
- [14] S. Kallio, J. Kela, J. Mäntyjärvi, and J. Plomp. Visualization of Hand Gestures for Pervasive Computing Environments. In Celentano and Mussio [5], pages 480–483.
- [15] C. Kirstein and H. Müller. Interaction with a Projection Screen Using a Camera-tracked Laser Pointer. In *Conference on MultiMedia Modeling*, pages 191–192. IEEE Computer Society, 1998.
- [16] J.-F. Lapointe and G. Godin. On-Screen Laser Spot Detection for Large Display Interaction. In *Workshop on Haptic Audio Environments and their Applications (HAVE 2005)*, pages 72–76, Ottawa, Ontario, Canada, Oct. 2005. IEEE.
- [17] D. Lingrand, P. Renevier, A.-M. Pinna-Déry, X. Cremaschi, S. Lion, J.-G. Rouel, D. Jeanne, P. Cuisinaud, and J. Soula. Gestaction3D : a platform for studying displacements and deformation of 3D objects using hands. In *International Conference on Computer-Aided Design of User Interfaces (CADUI)*, pages 105 – 114, Bucarest (Romania), June 2006. Springer-Kluwer.
- [18] T. B. Moeslund and L. Nørgaard. A Brief Overview of Hand Gestures used in Wearable Human Computer Interfaces. Technical Report CVMT 03-02, Laboratory of Computer Vision and Media Technology, Aalborg (Denmark), 2003.
- [19] T. B. Moeslund, M. Størring, and E. Granum. A Natural Interface to a Virtual Environment through Computer Vision-estimated Pointing Gestures. In *Int. Workshop on Gesture and Sign Language based Human-Computer Interaction (GW 2001)*, pages 59–63, London (UK), Apr. 2001.
- [20] J. Montagnat, H. Delingette, and N. Ayache. A review of deformable surfaces : topology, geometry and deformation. *Image and Vision Computing (IVC)*, 19(14) :1023–1040, Dec. 2001.
- [21] M. Porta. Browsing Large Collections of Images through Unconventional Visualization Techniques. In Celentano and Mussio [5], pages 440–444.
- [22] J. Segen and S. Kumar. Gesture VR : Vision-Based 3D Hand Interface for Spatial Interaction. In *International Conference on Multimedia (ACM Multimedia)*, pages 455–464, Bristol (UK), Sept. 1998. ACM Press.
- [23] B. Shizuki, T. Hisamatsu, S. Takahashi, and J. Tanaka. Laser Pointer Interaction Techniques using Peripheral Areas of Screens. In Celentano and Mussio [5], pages 95–98.
- [24] G. Smith, T. Salzman, and W. Stürzlinger. 3D Scene Manipulation with 2D Devices and Constraints. In *Graphics Interface*, pages 135–142, Ottawa (Ontario, Canada), June 2001.
- [25] T. Starner, B. Leibe, D. Minnen, T. Westyn, A. Hurst, and J. Weeks. The perceptive workbench : Computer-vision-based gesture tracking, object tracking, and 3D reconstruction for augmented desks. *Machine Graphics and Vision (MGV)*, 14 :59–71, 2003.
- [26] R. M. Taylor, T. Hudson, A. Seeger, H. Weber, J. Juliano, and A. Heller. VRPN : A Device-Independent, Network-Transparent VR Peripheral System. In *ACM Symposium on Virtual Reality Software & Technology (VRST)*, Banff (Canada), Nov. 2001. ACM, SIGGRAPH, and SIG-CHI, ACM Press.
- [27] G. Ye, J. J. Corso, and G. D. Hager. Gesture Recognition Using 3D Appearance and Motion Features. In *Workshop on Real-Time Vision for Human-Computer Interaction (in conj. with CVPR)*, Washington, DC (USA), June 2004. IEEE Computer Society.
- [28] Z. Zhang, R. Deriche, O. Faugeras, and Q. T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal (AIJ)*, 1-2(78) :87–119, 1994.