

Analyse de l'influence des systèmes de visualisation immersif sur l'assemblage virtuel de fragments en archéologie

Nicolas Mellado *
INRIA - Université Bordeaux 1

Patrick Reuter †
INRIA - Université Bordeaux 2

Florent Berthaut ‡
Université Bordeaux 1

RÉSUMÉ

Dans ce papier, nous présentons une analyse de l'influence des systèmes de visualisation immersifs associés à une interface d'interaction multimodale pour la résolution d'une tâche complexe. L'utilisateur est chargé de résoudre des puzzles 3D à deux pièces en utilisant une interface à deux fois 6 degrés de liberté.

Nous avons ainsi testé les performances des utilisateurs en utilisant trois modes d'affichage différents : un affichage monoscopique, un système de vision stéréoscopique et enfin un dispositif de suivi de position de la tête pour adapter le positionnement du point de vue à la posture physique associée à un affichage stéréoscopique. Notre comparaison des données est réalisée à partir du temps nécessaire aux utilisateurs pour réaliser la tâche d'assemblage.

Malgré le faible nombre d'utilisateurs testés, cette étude montre que l'apprentissage avec le mode d'affichage stéréoscopique est plus important qu'avec les deux autres modes. On observe aussi qu'après cette phase d'apprentissage effectuée, les utilisateurs travaillant avec un affichage stéréoscopique sont globalement plus performants que ceux des autres groupes.

Index Terms: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality H.5.2 [User Interfaces]: Input devices and strategies;

1 MOTIVATION

Les fragments archéologiques trouvés lors de campagnes de fouilles (figure 1) sont souvent brisés en un grand nombre de fragments et les archéologues sont confrontés à la reconstruction de "puzzles" 3D d'une grande difficulté. Scanner les fragments et les assembler virtuellement est une solution élégante et c'est même parfois la seule solution (figure 2).



FIGURE 1: Photographie des fragments issus du jeu de données proposé par Huang et al. dans [4].

Ré-assemblage automatique La recherche de solutions automatiques au problème des puzzles archéologiques est un domaine très actif, et plusieurs travaux récents proposent des systèmes présentant une résolution globale de ces puzzles.

*e-mail: mellado@labri.fr

†e-mail:preuter@labri.fr

‡e-mail:berthaut@labri.fr

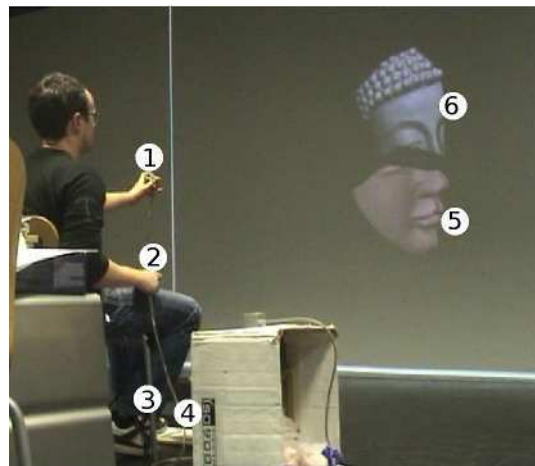


FIGURE 2: Assemblage virtuel des fragments scannés.

On peut notamment citer Brown et al. [1] qui présentent une solution pour l'assemblage de fresques fracturées. Leur système est basé sur le principe que les fragments de fresques peuvent être considérés comme planaires, et que l'on peut donc réduire le problème à la résolution d'un puzzle en 2.5D, plus simple car possédant beaucoup moins de solutions.

On peut aussi citer Huang et al. qui présentent dans [4] un système permettant de ré-assembler des fragments d'objets statuariques fracturés. C'est sur cette base que Thuswaldner et al. [7] ont pu réaliser une anastylose¹ numérique d'un temple en Turquie.

Ces trois approches sont très intéressantes en ce qui concerne l'assemblage géométrique des fragments, mais ont malgré tout un défaut commun : elles nécessitent de faire calculer au système un nombre important d'opérations déjà connues par l'archéologue. Dans la même logique, il apparaît très complexe de pouvoir transmettre au système, via un paramétrage initial, toutes les informations sémantiques que l'archéologue possède.

Ré-assemblage manuel C'est en suivant ces considérations que Reuter et al. [6] présentent un système de ré-assemblage manuel de fragment virtuel : ArcheoTUI. Les interfaces utilisateurs classiques telles que la souris et le clavier s'avèrent vite peu performantes pour manipuler en même temps deux fois 6 degrés de liberté. La solution proposée s'inspire des travaux de Hinckley [3] où des objets réels sont utilisés en tant qu'interface pour la visualisation en neurochirurgie. L'idée est de proposer à l'archéologue un dispositif de réalité virtuelle lui permettant de manipuler les données numériques comme s'il s'agissait des fragments réels. Avec ArcheoTUI, l'utilisateur a dans chaque main deux interacteurs et tous les mouvements (translations et rotations) qu'il effectue avec ceux-ci sont immédiatement reproduits à l'écran sur les deux frag-

1. Procédure archéologique de reconstruction, le plus souvent de bâtiments, en utilisant le maximum de matériel trouvé et le minimum de matériaux extérieurs.

ments numériques associés. Il en résulte une interaction efficace de la part de l'utilisateur pour orienter et mettre en position les pièces 3D. C'est pour cette raison que nous avons utilisé le système d'assemblage manuel de fragment ArcheoTUI pour réaliser notre étude.

En plus d'une interface de manipulation adaptée, l'assemblage de fragments virtuels nécessite une bonne perception de l'environnement 3D afin de percevoir de manière correcte les positions relatives des différents fragments les uns par rapport aux autres. Si l'on considère une tâche d'assemblage répétitive dont l'utilisateur connaît à l'avance la situation finale, l'efficacité avec laquelle il peut réaliser l'opération dépend de la manière dont il perçoit les positions des objets : plus la perception est bonne, plus l'utilisateur à une image mentale valide de l'environnement 3D et peut ainsi plus efficacement déplacer les objets pour les assembler correctement.

2 DISPOSITIF D'ASSEMBLAGE

Le dispositif de test était composé d'un ordinateur Dell© XPS (Processeur Intel© Core 2 duo) utilisant comme système d'exploitation Linux XUbuntu 9.04, les calculs graphiques étant réalisés par une carte graphique NVIDIA© GeForce 9800 GTX. La programmation est basée sur un graphe de scène OpenSG, qui permet de changer facilement le mode d'affichage entre monoscopique et stéréoscopique. Les images étaient visualisées via un système de projection sur un écran de 3x3 mètres à une distance de 2 mètres de l'utilisateur. L'étude a été réalisée avec le système d'assemblage de fragment ArcheoTUI.

2.1 Interface de manipulation multi-modale

Lorsqu'il assemble deux fragments l'utilisateur doit manipuler en un même temps deux fois 6 degrés de liberté (DOF). ArcheoTUI est une interface tangible dont la partie tangible est constituée de deux blocs de bois. Une interface tangible selon Ishii [5] augmente le monde réel en couplant l'information numérique à des objets de notre environnement quotidien. Dans notre contexte les formes à assembler sont très complexes ce qui a pour conséquence pour l'utilisateur, pour ne pas avoir de surcharge cognitive, de ne pouvoir manipuler que deux objets à la fois. Le nombre de blocs à un instant t est donc réduit à deux interacteurs, ainsi une technique d'interaction bi-manuelle basée sur le cadre conceptuel de Guiard [2] est proposée. Un système d'interaction à deux pédales au pied est aussi utilisé pour associer/dissocier le mouvement des interacteurs aux mouvements des fragments numériques.

L'idée clé du système ArcheoTUI est d'utiliser deux interacteurs comme à la fois une représentation physique et un moyen de contrôler les fragments archéologiques numérisés. L'interface utilisateur est visible sur les figures 2 et 3. Dans chaque main, un interacteur (les items 1 et 2) peut être librement manipulé dans l'espace 3D. A chaque interacteur (les items 1 et 2) correspond une pédale (les items 3 et 4). C'est seulement lorsque les pédales sont enfoncées que les mouvements de translation et de rotation des interacteurs correspondants sont associés à la visualisation sur l'écran des fragments numériques (les items 5 et 6). Ainsi, les utilisateurs ont une sorte de retour haptique passif lorsqu'ils manipulent les interacteurs. Une fois que la pédale est relâchée, le mouvement de l'interacteur est dissocié de celui du fragment numérique. Ainsi, alors que la position de fragment numérique est fixée, l'utilisateur peut à loisir se remettre dans une position plus confortable.

2.2 Systèmes de visualisation

L'étude consistait à tester l'importance des systèmes de visualisation dans une tâche de ré-assemblage de fragments archéologiques. Nous avons donc testé trois modes de visualisation aux propriétés immersives différentes, tous sur grand écran

- Mode A : Projection monoscopique,

- Mode B : Projection stéréoscopique utilisant la technologie Infitec® pour le filtrage des paires stéréopar des lunettes passives,
- Mode C : Projection stéréoscopique couplée à un système de modification du point de vue en fonction de la position de la tête de l'utilisateur (Head Tracking). Le système exploite deux caméras permettant de détecter des 5 marqueurs fixés sur les lunettes stéréoscopiques.

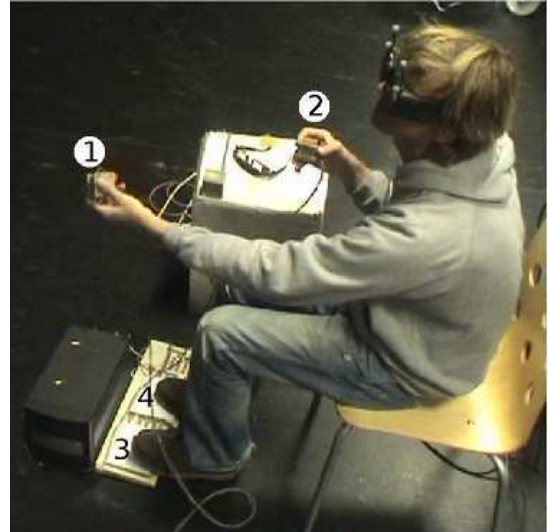


FIGURE 3: Assemblage virtuel avec l'affichage stéréoscopique et le Head Tracking du groupe C.

Le groupe utilisant le mode C pouvait modifier le point de vue dans l'environnement virtuel. Cependant, tous les utilisateurs étaient assis et ne disposaient donc pas d'une amplitude de mouvement de plus d'un mètre autour de leur position de référence. En effet, ArcheoTUI est une interface mixte bras/jambes qui nécessite une utilisation en étant assis.

3 ÉTUDE UTILISATEUR

3.1 Groupes de participants

Nous avons recruté douze participants, 10 hommes et 2 femmes, que nous avons aléatoirement divisé en trois groupes, A, B et C en fonction du mode d'affichage. Tous étaient des utilisateurs réguliers des systèmes informatiques et habitués à la visualisation d'espaces en trois dimensions.

3.2 Procédure

Les participants étaient testés individuellement. La procédure de passage était similaire pour les trois groupes et seul le système de visualisation variait. Chaque participant commençait l'étude par deux minutes d'essai de l'interface. Cette étape permet de réduire l'influence de la prise en main de l'interface (synchronisation main/pieds notamment) sur les résultats de l'expérience. Les fragments utilisés pour cette étape étaient différents de ceux utilisés pour l'étude.

Au terme de ces deux minutes, le participant devait résoudre successivement douze tests d'assemblage. Chaque essai était réalisé avec les deux mêmes fragments, mais dans des dispositions initiales différentes. Lorsque les deux pièces étaient bien positionnées, un témoin visuel en informait l'utilisateur, qui devait maintenir la position des fragments au moins 0,7 secondes. Un temps maximal de deux minutes était alloué pour réaliser la tâche, limite à partir de laquelle une nouvelle configuration initiale était proposée à l'utilisateur.

Pour l'entraînement comme pour l'étude, l'utilisateur connaissait la manière dont s'agencent les pièces, l'encadrant ayant préalablement montré une image décrivant la configuration finale à atteindre.



FIGURE 4: Position d'assemblage à retrouver par les participants.

Apprentissage de l'interface multi-modale Les participants devaient avoir une très faible expérience du système d'assemblage ArcheoTUI. L'interface de manipulation en elle-même nécessite une période d'utilisation pour être prise en main et être exploitée convenablement. Comme nous souhaitons comparer les systèmes d'affichage immersif dans le cadre d'un assemblage de fragments archéologiques et non l'influence de ces systèmes sur l'utilisation de l'interface ArcheoTUI, il était nécessaire que l'utilisateur puisse s'entraîner à manipuler des objets avant de commencer les tests.

Le temps nécessaire à l'utilisateur pour trouver la position valide était enregistré pour chaque essai d'assemblage. Cette information est nécessaire pour comparer l'évolution des capacités de perception des positions des fragments virtuels au fur et à mesure des essais. En effet, l'assemblage de fragments virtuels nécessite une bonne perception 3D pour concevoir une image mentale des positions proches de celles de l'environnement virtuel. Si cette perception est réduite, l'utilisateur peut se tromper et ainsi mal placer les fragments, et ainsi mettre plus de temps pour accomplir la tâche.

3.3 Résultats numériques

Nos mesures ont été réalisées sur les douze essais consécutifs des utilisateurs. Les figures 5, 6 et 7 représentent, respectivement pour les groupes A, B et C, le temps passé par chaque utilisateur pour réaliser les tâches d'assemblage, ainsi que la performance moyenne. On peut voir sur ces figures que la courbe de progression moyenne décroît plus vite et plus régulièrement pour le groupe B que pour les groupes A et C.

Pour vérifier cette tendance, nous avons utilisé le test de Wilcoxon, qui permet de comparer deux ensembles de données dont la distribution n'est pas forcément homogène. On considère que les six premiers essais représentent la phase d'apprentissage (nommée par la suite phase 1), et que les six suivants sont plus représentatifs d'une phase d'utilisation (nommée par la suite phase 2). Nous avons ainsi comparé deux à deux les différences de performances entre ces deux phases pour les trois groupes des modes A, B et C.

Premier Test : Comparaison des groupes A et B La première hypothèse nulle que nous avons testé est : *la progression est plus nette pour le groupe B que pour le groupe A*. La figure 8 présente l'écart en secondes entre la moyenne de la phase 1 et de la phase 2 selon Wilcoxon pour les utilisateurs des groupes A et B. Le rang de chaque utilisateur correspond à sa position si l'on trie les données selon cet écart par ordre croissant sachant qu'un rang plus grand signifie une plus grande amélioration. On peut ainsi voir que le groupe B obtient un rang cumulé de 25 bien plus important que

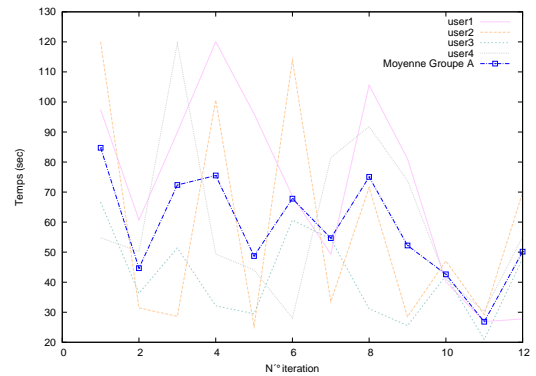


FIGURE 5: Relevé des temps pour le groupe A.

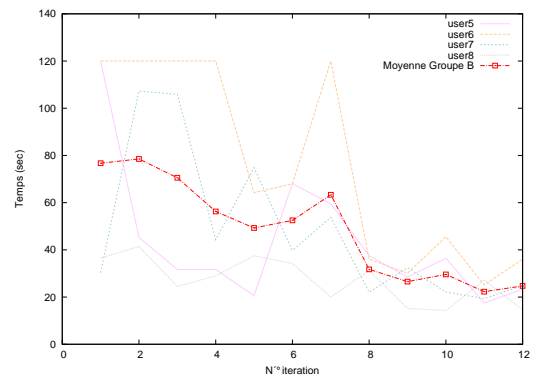


FIGURE 6: Relevé des temps pour le groupe B.

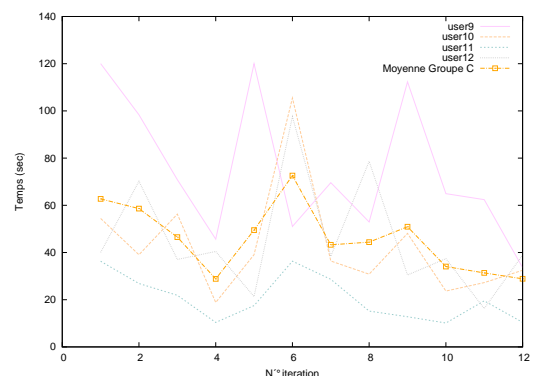


FIGURE 7: Relevé des temps pour le groupe C.

11 pour le groupe A, ce qui confirme l'hypothèse testée. La valeur de confiance est convenable avec une valeur de $p = 0,0286$.

Second Test : groupe C Si le mode B semble avoir un effet positif sur la courbe de progression, les résultats des utilisateurs du mode C semblent plus nuancés. Nous avons donc comparé les modes A et C pour tester la présence d'un écart statistique, avec l'hypothèse nulle suivante : *la progression est plus nette pour le groupe C que pour le groupe A*. Dans ce contexte, les rangs cumulés des deux groupes sont proches, avec 16 pour le groupe A et 20 pour le groupe C. La valeur importante de l'indice de confiance $p = 0,3429$ ne permet pas de valider cette hypothèse.

Groupe	Utilisateur	Diff Phase1-Phase2	Rang
A	User1	-7,96	1
	User2	19,73	2
	User3	33,36	3
	User4	37,91	5
B	User5	36,13	4
	User6	39,69	6
	User7	52,16	7
	User8	56,60	8

FIGURE 8: Données du test de Wilcoxon pour la comparaison des groupes A et C.

Troisième Test Étant donné le faible écart statistique entre les groupes A et C, il apparaît cohérente que l'écart entre les groupe B et C soit similaire à l'écart mesuré pour la première hypothèse (groupe A et B). Le test de Wilcoxon confirme cela en donnant un indice de confiance de $p = 0,0286$ similaire avec celui calculé pour le premier test.

3.4 Observations

Nous avons recueilli ces données à la fois en observant le comportement des utilisateurs pendant l'étude et en leur faisant remplir un questionnaire sur leurs impressions en fin de test.

Groupe A (Monoscopique) Le problème le plus récurrent rencontré par les utilisateurs du groupe A est lié à l'absence de perception de la profondeur. Tous ont été confrontés plusieurs fois à un mauvais placement dû à un fragment translaté selon l'axe orthogonal au plan de visualisation (écran). S'ils ont généralement compensé cette difficulté en modifiant le point de vue, tous ont relevé cette difficulté en fin de test.

Groupe B (Stéréoscopique) Le commentaire revenant sur tous les questionnaires de ce groupe est que la stéréoscopie aide à la perception des profondeurs et des volumes. Pendant les tests après quelques essais, tous les utilisateurs arrivaient à placer naturellement les deux fragments sur un même plan parallèle à l'écran, et n'avaient ensuite qu'à gérer le rapprochement et l'orientation des pièces. Cependant comme un utilisateur le précise, la perception de la profondeur est plus utile dans la phase de rapprochement que dans la phase d'assemblage proche, car les variations de distance sont plus importantes.

Groupe C (Stéréoscopique et Head Tracking) Lors du test, les utilisateurs du mode C étaient la plupart du temps statiques, et ne cherchaient donc pas à bouger leur tête pour voir depuis un autre point de vue. La majorité des utilisateurs ont d'ailleurs précisé ne pas sentir l'effet du Head Tracking pendant le test. Un utilisateur a indiqué qu'il aurait eu plus tendance à utiliser cette technique s'il avait pu être debout et se déplacer autour de l'objet virtuel.

Il est difficile de dire si l'écart de performance observé entre le mode B et le mode C est dû à une gêne causée par le Head Tracking ou à une différence entre les capacités des utilisateurs. En effet, si ces résultats peuvent donner une tendance sur les avantages d'un mode, ils restent néanmoins très liés aux performances initiales individuelles, à cause d'un nombre de participant à l'étude trop peu élevé. L'autre nuance à prendre en compte est la taille des fragments virtuels dans l'espace 3D et les possibilités de déplacement des utilisateurs. Si l'on veut laisser la possibilité aux participants d'observer l'objet sous plusieurs points de vue, il est indispensable d'en adapter la taille en fonction du champs de mouvement.

Remarques diverses La majorité des remarques ont été proposées par les utilisateurs du mode A, plus en difficulté donc plus tentés de trouver des alternatives et des améliorations au système. Le premier point qui revient plusieurs fois est l'ajout d'éléments de repère dans la scène 3D, comme par exemple un plan de travail, associé à un calcul d'ombrage permettant de déduire les profondeurs

des objets. L'autre tendance est l'ajout de fenêtres de visualisation pour plusieurs points de vue, permettant à l'utilisateur de positionner relativement l'un à l'autre les deux fragments selon tous les axes.

La seconde remarque est plus d'ordre comportementale : les utilisateurs ont tendance à éloigner l'objet de l'écran dans l'espace 3D, alors que cela a pour effet négatif de réduire la précision des actions de l'utilisateur. Comme on peut le voir à la figure 2, les fragments utilisés sont affichés à une taille importante. Un utilisateur a analysé sa propre tendance à éloigner les objets virtuels car il souhaitait qu'ils apparaissent entiers dans son champ de vision, pour éviter d'avoir à bouger la tête.

4 CONCLUSION ET PERSPECTIVES

Cette première étude utilisateur permet d'identifier l'utilité de l'affichage stéréoscopique pour résoudre une tâche d'assemblage de fragments archéologiques. On peut cependant noter deux faits importants qui peuvent nuancer ce résultat. En premier lieu, l'affichage monoscopique peut être amélioré pour permettre de transmettre des informations de profondeur : points de vue multiples, ombres portées, etc. Cela permettrait d'avoir des informations de profondeur quelque soit le mode d'affichage utilisé et ainsi identifier de manière plus juste les avantages de affichages immersifs. En second lieu, il est important de noter que cette étude doit être vue comme un préliminaire à une étude plus vaste, comportant notamment un nombre de sujet beaucoup plus important, afin de masquer les performances individuelles. Il peut aussi être intéressant de réaliser une sélection des utilisateurs par un test préliminaire en fonction de leurs capacités naturelles de perception dans l'espace et éventuellement comparer leurs réactions face au système.

ACKNOWLEDGEMENTS

Les auteurs souhaitent remercier Pascal Guitton, Sebastian Knoedel et toutes les personnes ayant participé à l'étude. Merci à Anne Gégout-Petit pour son aide à la conception du test de Wilcoxon. Merci à l'Université technologique de Vienne pour le modèle 3D utilisé. Cette étude a été réalisée dans le cadre du projet SeARCH soutenu par l'Agence Nationale de Recherche ANR-09-CORD-019-01

RÉFÉRENCES

- [1] B. J. Brown, C. Toler-Franklin, D. Nehab, M. Burns, D. Dobkin, A. Vlachopoulos, C. Dumas, S. Rusinkiewicz, and T. Weyrich. A system for high-volume acquisition and matching of fresco fragments : Reassembling Thera wall paintings. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 27(3), Aug. 2008.
- [2] Y. Guiard. Asymmetric division of labor in human skilled bimanual action : The kinematic chain as a model. *Journal of Motor Behavior*, pages 486–517, 1987.
- [3] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. Passive real-world interface props for neurosurgical visualization. In *CHI '94 : Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 452–458, New York, NY, USA, 1994. ACM.
- [4] Q.-X. Huang, S. Flöry, N. Gelfand, M. Hofer, and H. Pottmann. Reassembling fractured objects by geometric matching. *ACM Trans. Graphics*, 25(3) :569–578, 2006.
- [5] H. Ishii and B. Ullmer. Tangible bits : towards seamless interfaces between people, bits and atoms. In *CHI '97*, pages 234–241, New York, NY, USA, 1997. ACM.
- [6] P. Reuter, G. Rivière, N. Couture, N. Sorraing, L. Espinasse, and R. Vergnien. A tangible user interface for the virtual reassembly of fractured archaeological objects. In *Proceedings of VAST 2007*. Eurographics, 2007.
- [7] B. Thuswaldner, S. Flöry, R. Kalasek, M. Hofer, Q.-X. Huang, and H. Thür. Digital anastylosis of the octagon in ephesos. *J. Comput. Cult. Herit.*, 2(1) :1–27, 2009.