



DichotoZoom : Technique de sélection d'objets dans un environnement virtuel 3D affiché sur Smartphone

Eliane Balaa, Mathieu Raynal, Florent Cabric, Emmanuel Dubois

► To cite this version:

Eliane Balaa, Mathieu Raynal, Florent Cabric, Emmanuel Dubois. DichotoZoom : Technique de sélection d'objets dans un environnement virtuel 3D affiché sur Smartphone. 30eme conférence francophone sur l'interaction homme-machine, Oct 2018, Brest, France. 11p. hal-01899077

HAL Id: hal-01899077

<https://hal.science/hal-01899077>

Submitted on 19 Oct 2018

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.

DichotoZoom : Technique de sélection d'objets dans un environnement virtuel 3D affiché sur Smartphone

DichotoZoom : Selection technique in a 3D virtual environment displayed on Smartphone

Eliane Balaa
IRIT, Université de Toulouse
118 Route de Narbonne
31062, Toulouse cedex 9, France
eliane.balaa@irit.fr

Florent Cabric
IRIT, Université de Toulouse
118 Route de Narbonne
31062, Toulouse cedex 9, France
florent.cabric@irit.fr

Mathieu Raynal
IRIT, Université de Toulouse
118 Route de Narbonne
31062, Toulouse cedex 9, France
mathieu.raynal@irit.fr

Emmanuel Dubois
IRIT, Université de Toulouse
118 Route de Narbonne
31062, Toulouse cedex 9, France
emmanuel.dubois@irit.fr

ABSTRACT

Except for video games, the use of 3D virtual environment (3D VE) on Smartphone has not yet been popularized. This is due to lack of interaction techniques to interact with a 3D VE composed of many objects. In this article, we focus on selecting objects in dense 3D VE. We propose DichotoZoom which allows the selection of objects inside a set of already pre-selected ones, with a limited number of actions. In a first study, we compare our technique with the Circulation technique, suggested by the literature. This study shows the effectiveness of DichotoZoom when the set contain significantly numerous elements. Finally, we studied its use with various interaction modalities used on Smartphone : using physical or graphical buttons, and gestural interactions via touch screen or moving the device itself.

CCS CONCEPTS

• **Human-centered computing** → **Pointing**; *Mobile devices*;

KEYWORDS

3D interaction; Selection task; mobile device

RÉSUMÉ

Hormis pour les jeux vidéo, l'utilisation d'environnement virtuel 3D (EV 3D) sur Smartphone n'est pas encore démocratisée. Ceci

est dû au manque de techniques d'interaction proposées pour interagir avec un EV 3D composé de nombreux objets. Dans cet article, nous nous intéressons à la sélection d'objets dans des EV 3D denses. Nous proposons DichotoZoom qui permet la sélection d'objets parmi un ensemble d'objets déjà présélectionnés, en un nombre d'actions limité. Dans une première étude, nous comparons notre technique à la technique Circulation, proposée par la littérature. Celle-ci montre l'efficacité de DichotoZoom lorsque l'ensemble présente un nombre d'éléments important. Enfin, nous avons étudié son usage avec différentes modalités d'interaction utilisables sur Smartphone : utilisation de boutons physiques ou sous forme de composants graphiques, interactions gestuelles au moyen de l'écran tactile ou du Smartphone lui-même.

MOTS-CLEFS

Interaction 3D; tâche de sélection; dispositif mobile

1 INTRODUCTION

Les environnements virtuels 3D (EV 3D) sont aujourd'hui de plus en plus répandus. Au-delà des jeux vidéo [9], nous retrouvons ces environnements dans différents types d'applications métier (CAO [20], aéronautique [25], etc.) ou grand public (par exemple pour une expérience interactive en musée [3]). Cependant, malgré les récentes améliorations apportées aux performances des dispositifs mobiles, ces environnements virtuels ne sont pas autant présents sur les petits dispositifs. Les smartphones sont parfois utilisés comme dispositif d'interaction pour un EV 3D affiché à distance [2], mais l'affichage d'un tel environnement sur Smartphone reste cantonné aux jeux vidéo ou pour des applications de réalité augmentée qui font intervenir peu d'objets virtuels dans la scène [19, 21]. Ceci peut s'expliquer par le fait que l'affichage de nombreuses données dans un EV 3D affiché sur un petit écran accentue les principaux problèmes d'interaction déjà présents sur des écrans de plus grande taille [1].

D'une part, se pose le problème de densité, c'est-à-dire des objets qui sont très proches les uns des autres. Cela peut poser un problème lors de la sélection d'objets : si plusieurs objets sont proches les uns des autres, les erreurs de pointage sont plus fréquentes. Ce problème peut être accentué lorsque l'EV 3D est affiché sur un écran plus petit : la surface d'affichage étant plus petite, les espaces entre les objets sont diminués. De plus, le pointage sur écran tactile se fait généralement au doigt, la précision du pointage est alors moins élevée que lorsqu'il est réalisé à l'aide d'un pointeur.

D'autre part, il peut également y avoir un problème d'occultation lorsque l'objet à atteindre est caché, ou partiellement caché, par d'autres éléments de la scène. L'occultation pose le problème de la recherche de l'objet. Mais ceci est aussi problématique dans le cas de la sélection pour un objet partiellement caché. Dans ce cas, le pointage de l'objet est rendu plus difficile car la surface disponible de l'objet à atteindre est réduite. Dans le cas d'un affichage sur petit écran, si la densité des objets affichés est plus élevée, le risque d'occultation sera par conséquent augmenté.

Enfin, il y a la profondeur qui pose un problème des déplacements. En effet, les dispositifs d'interaction généralement utilisés sont conçus pour une utilisation en 2 dimensions et ne proposent pas suffisamment de degrés de liberté pour pouvoir interagir correctement avec les 3 dimensions de l'EV 3D. Pour faciliter les interactions avec un EV 3D, des dispositifs en entrée proposent plus de degrés de liberté [12, 23]. Cependant, ces dispositifs, conçus pour une utilisation sur PC, ne sont pas toujours utilisables avec un Smartphone. Les interactions sont donc limitées aux capacités offertes par les dispositifs (écran tactile, bouton physique, caméra) et capteurs (accéléromètre, gyroscope) intégrés au Smartphone.

Ces trois problèmes majeurs freinent donc l'utilisation des EV 3D sur Smartphone. Dans cet article, nous nous intéressons plus particulièrement à la sélection d'un objet dans un EV 3D composé d'une forte densité d'objets et affiché sur Smartphone. Pour répondre à cette problématique, nous proposons DichotoZoom, une technique de sélection qui minimise le nombre d'actions à effectuer pour atteindre l'objet souhaité.

2 SÉLECTION DANS UN EV 3D

2.1 Techniques de pointage dans un EV 3D

Les techniques de pointage dans un EV 3D peuvent être classées en deux catégories principales [5] : d'une part, les techniques basées sur le déplacement d'un pointeur dans l'EV 3D pour réaliser du pointage tel qu'on le connaît sur les interfaces 2D classiques ; et d'autre part, les techniques basées sur la manipulation d'un rayon permettant de pointer plus facilement dans la profondeur.

Le premier type de technique consiste à déplacer un curseur dans l'EV 3D pour le positionner sur l'objet souhaité (cf. Figure 1, gauche). Le cas le plus habituel consiste à avoir un curseur de petite taille [15] ou une main virtuelle [24] que l'utilisateur doit amener à la cible voulue. Pour ce faire, soit l'utilisateur interagit avec un dispositif de pointage spécifique comme une souris 3D [26] soit il peut utiliser un dispositif de suivi des mouvements de l'utilisateur, généralement au niveau de la main, pour associer son mouvement au déplacement du pointeur dans la scène [5]. Une fois le curseur dans la cible, l'utilisateur n'a plus qu'à valider son

pointage. Cependant, des études ont montré que cette technique à base de pointeur est très lente, surtout lorsque la distance à parcourir est élevée et se trouve éloignée en profondeur [6]. Une solution pour réduire ce problème consiste à utiliser un volume de sélection plutôt qu'un curseur de petite taille [32]. Le volume de pointage est alors plus important et tous les objets se trouvant à l'intérieur (ou en contact) avec le volume peuvent alors être sélectionnés.

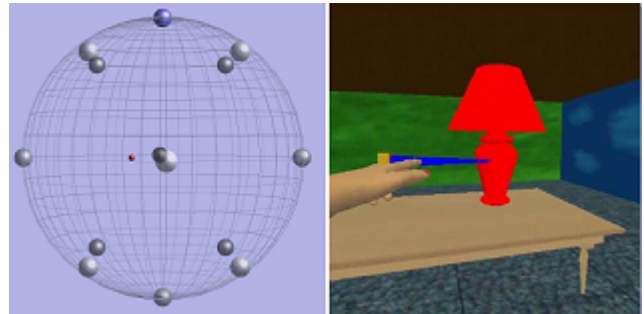


FIGURE 1: Exemple de pointeur à gauche, et de rayon à droite

La seconde catégorie de techniques concerne l'utilisation d'un rayon pour pointer l'objet désiré dans l'EV 3D (cf. Figure 1, droite). Ce type de technique est communément appelé le lancer de rayon [17]. L'utilisateur contrôle alors une extrémité du rayon et l'oriente dans la scène. Le rayon part alors à l'infini dans la direction souhaitée par l'utilisateur. Cependant, la technique à base de lancer de rayon présente un inconvénient majeur : le fait de projeter un rayon sur toute la profondeur peut amener à sélectionner plusieurs objets en même temps. Pointer un objet partiellement ou totalement occulté par un autre devient alors complexe. Pour remédier à ce problème d'occultation, « Flexible Pointer » [11] permet de courber le rayon lancé par l'utilisateur et ainsi contourner les objets qui se trouvent devant l'objet souhaité. Cependant, cette technique nécessite l'utilisation de deux dispositifs de pointage en simultanée. De plus, cette technique se trouve confrontée au même problème que le lancer de rayon traditionnel où plusieurs objets peuvent toujours être traversés par le rayon. Enfin, comme pour les pointeurs, le lancer de rayon classique souffre d'un problème de précision. En effet, il est compliqué de pointer avec précision un objet qui se trouve loin du point de vue de l'utilisateur. C'est pourquoi, des rayons volumiques ont été proposés, comme l'utilisation d'un cylindre dont l'axe central représente le rayon principal manipulé par l'utilisateur [10], ou encore « Spotlight » qui propose un cône [18] comme volume de sélection. De la même manière que pour les sphères, tous les éléments qui sont en contact avec le volume de sélection sont alors sélectionnés.

2.2 Sélection d'un sous-ensemble d'objets

Certains types d'EV 3D tels que les environnements constitués de nuages de particules structurées [30] ou non [14] peuvent contenir énormément d'éléments. Dans ce cas, plutôt que de sélectionner directement un élément en particulier, des techniques de sélection de sous-ensemble d'objets sont utilisées. Pour sélectionner le sous-ensemble voulu, Naito et al. [22] utilisent une sphère dont l'utilisateur peut augmenter ou diminuer la taille pour sélectionner

plus ou moins d'éléments. Pour effectuer une sélection plus précise du sous-ensemble de point, Hamelin et Dubois [14] proposent d'entourer à la main le sous ensemble de point. D'autres nuages de points sont structurés. Dans ce cas, il est plus facile de sélectionner un sous-ensemble en fonction de sa densité [30] ou du contexte [31]. Pour ce faire, des outils sur tablette sont proposés par Yu et al. [30, 31] pour entourer sur le plan 2D les éléments souhaités. Puis l'algorithme sélectionne les particules en fonction de la densité ou du contexte du nuage de particules.

2.3 Sélection parmi un ensemble de cibles présélectionnées

Comme nous venons de le voir, pour faciliter la sélection dans un EV 3D, beaucoup de techniques se basent sur l'utilisation d'un volume de sélection. De ce fait, pour pouvoir finaliser le pointage d'un élément en particulier, une phase de sélection parmi ces éléments présélectionnés est nécessaire. Différentes stratégies ont été proposées pour désambigüiser ces multiples sélections.

Wyss et al. [29] ont par exemple proposé d'utiliser deux rayons pour sélectionner l'objet voulu. L'objet sélectionné est alors l'objet se trouvant à l'intersection des deux rayons pointés dans l'EV 3D. Cependant, cette technique nécessite d'utiliser deux dispositifs de pointage. De plus, pour que cette technique puisse être réellement efficace, il faut que l'utilisateur puisse avoir deux points de vue de l'EV 3D suffisamment éloignés pour que l'intersection entre les 2 rayons ait réellement un sens et ne pointe qu'un seul et unique objet en commun.

D'autres travaux proposent l'utilisation d'un composant supplémentaire, généralement un menu [10, 18] pour sélectionner l'élément souhaité parmi les différents objets présélectionnés. Tous les objets présélectionnés sont affichés sous la forme d'un item dans un menu, et l'utilisateur n'a plus qu'à choisir l'item qu'il souhaite. Cependant, cette technique comporte deux inconvénients majeurs : d'une part, comment faire le lien entre un item textuel [10] et les objets de la scène ? D'autre part, ce système ne peut fonctionner qu'avec peu d'objets. En effet, dès que le nombre d'objets présélectionnés devient trop important, le menu prend trop de place à l'écran et occulte complètement l'EV 3D.

Une autre manière de faciliter le pointage des éléments du sous-ensemble présélectionné consiste à réorganiser les éléments à l'écran comme le proposent les techniques SQUAD [16] ou Expand [8]. Ces techniques disposent sur un plan 2D l'ensemble des objets présélectionnés. Ceci a l'avantage de faciliter le pointage des éléments et notamment sur un écran tactile. En revanche, si le nombre d'éléments à afficher est trop grand, ces derniers ne pourront pas tous être affichés en même temps à l'écran. D'autre part, le réagencement des éléments peut perturber l'utilisateur qui doit à nouveau rechercher l'élément qu'il souhaite dans ce nouvel espace d'affichage. D'autre part, le fait de réorganiser les éléments fait perdre leur position d'origine aux éléments et notamment leur positionnement les uns par rapport aux autres. Il est important de conserver ce positionnement si les éléments sont dépendants les uns des autres.

Enfin le dernier type de stratégies proposé pour désambigüiser la sélection consiste à utiliser une interaction supplémentaire directement avec le dispositif de pointage utilisé pour la manipulation du rayon ou du volume de sélection. La plus courante consiste à

parcourir les objets présélectionnés, soit en passant directement de l'un à l'autre [10] soit en manipulant un pointeur supplémentaire sur le rayon [13]. Dans ce dernier cas, les problèmes de précision de pointage évoqués pour les techniques à base de pointeur seront identiques : dès que les objets seront éloignés en profondeur et qu'un grand nombre d'objets seront sur le rayon, il sera alors difficile de pointer précisément l'objet souhaité. Dans le cas de la technique Circulation [10], le curseur saute d'un objet à un autre directement en appuyant sur un bouton du dispositif. Ce déplacement discret plutôt que continu permet ainsi de limiter ces problèmes de précision.

2.4 Problématique du pointage dans un EV 3D sur Smartphone

La majorité des travaux concernant les EV 3D se sont focalisés sur un affichage de ce dernier soit sur des surfaces d'affichage traditionnelles (écrans d'ordinateur, ou projections sur grandes surfaces horizontales ou verticales, par exemple) ; soit dans des environnements plus immersifs et appropriés à de la 3D [13]. L'affichage et l'interaction avec un EV 3D sur Smartphone est plus difficile car la surface d'affichage proposée par l'écran du dispositif est beaucoup plus petite. De ce fait, deux solutions sont possibles pour afficher des EV 3D complexes sur ces petits dispositifs : soit n'afficher qu'une partie de l'EV 3D pour conserver une taille d'objets assez grosse ; soit afficher l'ensemble de l'EV 3D, en réduisant la taille des objets.

Dans le premier cas de figure, si l'EV 3D n'est pas affiché entièrement, l'utilisateur a une moins bonne perception de la scène dans son ensemble. De plus, l'objet à atteindre n'est alors pas forcément visible à l'écran et cela nécessite donc une phase de navigation préalable au pointage pour rechercher l'objet désiré.

L'autre solution consiste à afficher l'ensemble de l'EV 3D. Dans ce cas, chaque objet a alors une taille plus petite. D'autre part, la taille réduite de ces écrans rend aussi difficile l'utilisation d'un pointeur : ce dernier serait de trop petite taille et nécessiterait une trop grande précision. Il semble donc plus approprié d'utiliser un volume pour sélectionner plusieurs objets, puis sélectionner l'objet parmi ceux pré-sélectionnés dans le volume. Cependant, les techniques de sélection proposées ne semblent pas adaptées à une utilisation sur Smartphone. L'utilisation d'un menu pour une sélection parmi plusieurs objets risquerait de masquer l'ensemble de l'EV3D. La meilleure solution sur petit écran apparaît donc être de conserver une vision des différentes cibles présélectionnées, puis d'affiner la sélection parmi celle-ci.

Dans cet article, nous avons choisi de nous intéresser plus particulièrement à la sélection d'un élément dans un ensemble d'objets présélectionnés. Pour éviter les problèmes de précision du pointage avec un pointeur, et d'occlusion avec l'affichage d'une liste ou d'un menu, nous avons choisi de nous focaliser sur la solution consistant à passer le focus de sélection d'un objet à un autre. La technique Circulation [8] permet de passer d'un objet à un autre au moyen d'un appui sur un bouton du dispositif de pointage. Le premier objet sélectionné dans le sous-ensemble déjà présélectionné est celui se trouvant le moins éloigné en profondeur. A chaque appui sur le bouton, le système sélectionne l'objet suivant dans la profondeur. Par conséquent, le temps de sélection peut croître rapidement si l'objet à sélectionner est situé après un nombre important d'objets

dans la profondeur : l'utilisateur devra passer un à un tous les objets le précédant, ce qui ralentira la sélection de l'objet souhaité.

3 DICHOTOZOOM

3.1 Principe

La technique DichotoZoom que nous proposons ne concerne dans la tâche de pointage d'une cible que la seconde étape : la sélection d'une cible parmi N déjà présélectionnées. Cette technique peut donc être associée à n'importe quelle solution de pointage qui permet de sélectionner plusieurs cibles en même temps (lancer de rayon ou volume de sélection).

DichotoZoom reprend le principe de la dichotomie. Le but est de diviser par deux le nombre d'objets présélectionnés à chaque fois que l'utilisateur effectue une action de sélection. A chaque étape, l'ensemble des éléments présélectionnés est divisé en deux sous-ensembles de cibles de taille égale. Ces sous-ensembles sont constitués en fonction de la profondeur de chaque objet par rapport au point de vue de l'utilisateur (cf. Figure 2). Les $N/2$ éléments les plus proches du point de vue de l'utilisateur constituent un premier sous-ensemble (objets en violet sur la figure 2), tandis que les $N/2$ plus éloignés composent le second sous-ensemble (objets en jaune sur la figure 2). L'utilisateur choisit alors le sous-ensemble dans lequel se situe sa cible. L'EV 3D n'affiche plus que le sous-ensemble choisi et réajuste le point de vue de l'utilisateur si c'est le sous ensemble le plus éloigné qui a été choisi. L'opération est alors renouvelée sur ce nouveau sous-ensemble jusqu'à ce que l'objet souhaité se trouve au premier plan de l'EV 3D, et que l'utilisateur puisse le sélectionner sans contraintes.

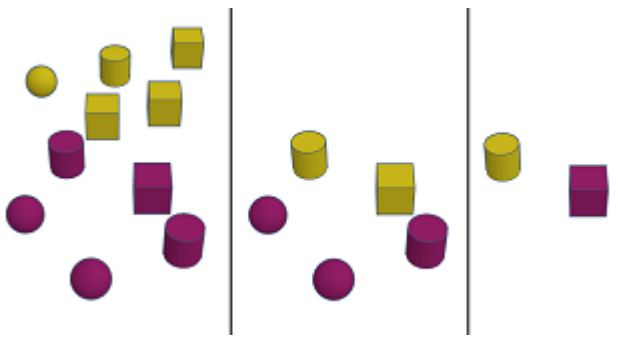


FIGURE 2: Exemple de sélection dans un ensemble d'éléments présélectionnés : la cible à atteindre est le cube représenté en violet sur l'image de gauche

3.2 Apport théorique

La technique Circulation nécessite de passer d'objet en objet pour atteindre celui désiré. Ainsi si l'utilisateur souhaite atteindre un objet qui se trouve à la N position dans un ensemble d'objets, l'utilisateur est obligé de réaliser N actions avant d'arriver à l'objet en question. Donc, plus l'objet se trouve dans un ensemble contenant une grande quantité d'objets et à une position éloignée en profondeur par rapport au point de vue de l'utilisateur, et plus l'utilisateur aura à réaliser un grand nombre d'actions pour atteindre cet objet.

Pour un ensemble de N objets, le nombre d'actions à réaliser pour atteindre l'objet désiré sera alors compris entre 1 et N selon sa position dans l'ensemble.

Dans le cas de DichotoZoom, le nombre de cibles présent dans l'ensemble est divisé par deux après chaque action de l'utilisateur. Par conséquent, l'utilisateur aura à réaliser au maximum $\log_2(N)$ actions pour atteindre la cible souhaitée dans un ensemble de N éléments présélectionnés. Donc plus le nombre d'objets présents dans l'ensemble est important et plus DichotoZoom devrait être performant pour atteindre la cible voulue par rapport à la technique Circulation.

4 ÉTUDE PRÉLIMINAIRE : EFFICACITÉ DE LA TECHNIQUE DICHOTOZOOM

4.1 Objectifs de l'étude et hypothèses

4.1.1 Objectifs. L'objectif principal de cette étude est d'évaluer la pertinence de l'utilisation d'un algorithme de dichotomie pour la sélection d'un objet. Avant même d'étudier la meilleure interaction possible avec cet algorithme, notre but est de vérifier qu'une sélection au moyen d'un algorithme de dichotomie ne perturbe pas l'utilisateur par rapport à une sélection classique dite « pas-à-pas ».

Nous envisageons deux risques principaux de perturbations pour l'utilisateur : d'une part, est-ce que cet algorithme ne va pas produire plus d'erreurs de sélection de la part de l'utilisateur qu'une sélection classique¹ ? D'autre part, est-ce que la recherche de la cible dans le nouveau sous-ensemble, après chaque action, ne va pas nécessiter un temps de recherche visuelle trop long qui serait préjudiciable pour les utilisateurs ? Ce sont donc ces différents facteurs liés à la perception de l'utilisateur que nous souhaitons analyser avant d'étudier les différentes modalités d'interaction que nous pourrions utiliser pour interagir avec cet algorithme.

4.1.2 Hypothèses. Du fait de diviser par deux à chaque action, le nombre de cibles sélectionnables, la technique DichotoZoom permet de réduire considérablement le nombre d'actions que les utilisateurs ont à réaliser pour atteindre une cible précise.

Cependant, la technique DichotoZoom reconfigure l'espace après chaque action. La cible à atteindre peut changer de zone après chaque action de l'utilisateur. Celui-ci est donc obligé de regarder après chaque action dans quelle zone l'élément souhaité est situé. Par conséquent, l'utilisateur ne pourra pas anticiper ses actions comme il peut le faire avec la technique Circulation où il peut prévoir à peu près le nombre d'actions à réaliser avant de commencer sa tâche de sélection. Nous supposons donc que le temps nécessaire pour réaliser une action sera plus important avec DichotoZoom, du fait du temps de perception des nouvelles zones (H1).

Enfin, malgré la perte de temps dans la recherche de cible et la programmation de l'action suivante, nous pensons que l'utilisateur gagnera du temps sur la réalisation de la tâche grâce à la diminution conséquente du nombre d'actions à réaliser (H2).

1. Nous entendons ici par erreur, les erreurs dues à une mauvaise interprétation du contexte par l'utilisateur, et non pas celles dues à une mauvaise précision du dispositif

4.2 Choix d'implémentation

Pour faciliter la perception des deux sous-ensembles définis par DichotoZoom, nous avons choisi de colorer les éléments du sous-ensemble le plus proche à moitié en violet, et ceux du sous-ensemble le plus éloigné du point de vue de l'utilisateur à moitié en jaune (cf. Figure 3). Nous avons fait ce choix pour notre étude afin de distinguer facilement les deux sous-ensembles. Dans le cas d'EV 3D concrets, où la couleur de chaque objet peut avoir son importance, d'autres solutions peuvent être envisagées pour distinguer les deux sous-ensembles. Par exemple, il serait possible de colorer seulement le contour des objets, ou mettre un plan 2D entre les deux sous-ensembles pour mieux distinguer la limite entre les deux.

D'autre part, nous faisons le choix dans cette première implémentation de DichotoZoom de ne pas afficher les objets qui ne sont pas sélectionnés. Ainsi, après chaque action, le nombre d'objets affichés est réduit de moitié. Cependant pour un cas réel d'usage, nous pourrions envisager d'autres solutions où tous les objets restent visibles pour conserver une vue globale de l'EV 3D. Par exemple, tous les éléments qui ne font plus partie du sous-ensemble pourraient être affichés en transparence pour que l'utilisateur puisse toujours les percevoir et les situer par rapport aux autres objets qui restent sélectionnables.

Enfin, un utilisateur n'a besoin que de 3 actions distinctes pour faire fonctionner DichotoZoom : deux actions pour sélectionner les sous-ensembles proche ou éloigné, puis une action pour annuler la dernière sélection. Pour cette première étude, nous avons donc choisi d'utiliser les boutons du Smartphone pour effectuer ces trois actions. Nous avons fait ce choix pour éviter un apprentissage de gestes ou des problèmes de détection ou de reconnaissance que nous pourrions avoir avec des gestes sur l'écran ou avec le Smartphone. L'utilisateur sélectionne donc la zone la plus proche ou la plus éloignée de son point de vue respectivement avec les boutons de volume « UP » et « DOWN ». Il peut annuler la dernière action en appuyant sur le bouton « Retour arrière ». Enfin, la validation d'une sélection se fait en appuyant sur l'écran tactile. La cible la moins éloignée en profondeur est alors sélectionnée.

4.3 Matériel et méthode

4.3.1 Participants. Douze personnes (10 hommes et 2 femmes) ont participé à cette étude. Ils avaient en moyenne 20 ans (écart-type de 1 an), 11 parmi eux étaient droitiers. Tous étaient étudiants à l'université et avaient l'habitude d'utiliser un Smartphone.

4.3.2 Dispositif. L'expérimentation s'est déroulée avec un Samsung Galaxy SII doté d'un processeur Dual-Core 1.2 GHz, le tout fonctionnant sous Android 2.3.4. L'écran tactile a une taille de 4.3 pouces de diagonale, avec une résolution de 480×800 pixels. L'environnement virtuel 3D a été réalisé en Open GL. Lors de l'expérimentation, les participants étaient assis à un bureau. Ils pouvaient tenir le smartphone comme ils le souhaitaient.

4.3.3 Choix de la tâche et de la technique de base. Comme nous avons pu le voir précédemment, la sélection d'une cible dans un EV 3D avec une forte densité d'objets se décompose généralement en 3 sous-tâches : la navigation dans l'environnement, la sélection d'un sous-ensemble d'objets au moyen d'un rayon ou d'un volume, puis la sélection de l'objet souhaité dans ce sous-ensemble.

L'algorithme de dichotomie que nous proposons n'intervient que dans la troisième sous-tâche pour la sélection d'un objet dans un sous-ensemble contenant encore de nombreux objets.

Des techniques telles que le menu [10] ou SQUAD [16] fonctionnent avec ces trois mêmes sous-tâches (cf. section 2.3). Cependant dans [10] et [16] les objets présélectionnés sont respectivement réorganisés sur un menu linéaire ou sur un plan 2D pour être plus facilement sélectionnables. Dans notre cas d'étude, nous souhaitons que les objets conservent leur place d'origine dans l'environnement pour que l'utilisateur ne soit pas perturbé au moment de cette dernière tâche de sélection. C'est pourquoi nous avons finalement choisi la technique Circulation [10] comme technique de base avec laquelle nous comparons DichotoZoom. Comme les deux premières sous-tâches (navigation et sélection du sous-ensemble d'objets) peuvent être réalisées de la même manière avec les deux techniques, nous avons ciblé notre étude exclusivement sur la dernière sous-tâche. Les participants n'avaient pas à faire une sélection initiale avec un rayon ou un volume pour déterminer le sous-ensemble. La tâche que nous demandions aux participants était de sélectionner une cible parmi un ensemble d'objets. Cet ensemble correspondrait, dans la tâche complète, au sous-ensemble d'objets sélectionnés. Pour manipuler la technique Circulation, les interactions se font aussi au moyen des boutons physiques : l'utilisateur peut passer d'une cible à une autre en appuyant sur le bouton de volume « UP », et peut revenir à la sphère précédente en appuyant sur le bouton « Retour arrière ». Enfin, de la même manière que pour DichotoZoom, l'utilisateur appuie sur l'écran pour valider sa sélection.

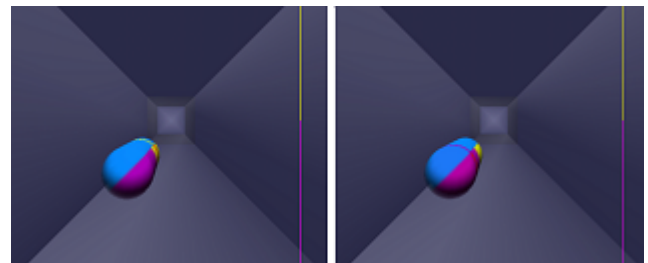


FIGURE 3: Ensemble d'éléments présélectionnés : la cible à atteindre est dans la zone la plus éloignée (à gauche) ; et dans la zone la plus proche (à droite)

4.3.4 Procédure. La tâche demandée aux utilisateurs consistait à sélectionner une cible sphérique parmi 32 affichées dans un EV 3D. La consigne était d'utiliser la technique de sélection proposée jusqu'à avoir la cible souhaitée au premier plan de l'EV 3D.

Pour simuler la profondeur et l'occlusion partielle, les 32 sphères sont toutes placées à des profondeurs Z différentes (plus d'une fois le diamètre des sphères), mais avec des coordonnées dans le plan (X, Y) assez proches les unes des autres (la distance sur le plan 2D entre les centres de deux cibles successives n'excède pas la valeur du rayon des cibles). Ainsi, nous simulons l'occultation et la densité des sphères présentes dans l'ensemble. Pour que les participants puissent percevoir l'ensemble des sphères, celles-ci étaient affichées en semi-transparence. Le niveau de transparence a été défini empiriquement préalablement à l'expérimentation.

Comme mentionné précédemment, la validation d'une sélection se faisait par un appui sur l'écran. L'évaluation ne portant pas sur la précision du pointage au moment de la validation, nous considérons la tâche terminée dès lors que l'utilisateur appuie sur l'écran, même si la cible n'est pas correctement pointée et nous prenons comme cible validée celle se trouvant en première position sur l'EV 3D. Par conséquent, nous comptons une erreur si la cible pointée présente au premier plan n'est pas la cible demandée. En revanche, nous ne comptons pas d'erreur si le pointage n'est pas effectué correctement sur la cible la plus proche mais que celle-ci est bien la cible voulue.

4.3.5 Design. Chaque participant a réalisé deux sessions : une avec chaque technique (DichotoZoom et Circulation). L'ordre de passage des deux techniques a été contrebalancé entre les utilisateurs. Dans chaque session, les participants ont réalisé trois exercices constitués chacun de 32 tâches de sélection. Les participants ont pu se familiariser pendant 5 minutes avant chaque session avec la technique qu'ils allaient utiliser. L'ordre des tâches était aléatoire. Pour chaque tâche d'un même exercice, la position de la cible à valider était différente. Toutes les positions de l'EV 3D étaient donc utiliser une fois à chaque exercice. Ainsi, nous avons obtenu 2304 tâches de sélection (12 participants \times 2 sessions \times 3 exercices \times 32 tâches).

4.4 Résultats

Pour analyser nos résultats, nous nous appuyons sur des techniques d'estimations basées sur des amplitudes d'effet et des intervalles de confiance à 95 %, en accord avec les récentes recommandations de l'APA [27]. Les scripts utilisés sont ceux utilisés par [4, 28] et disponibles en ligne²

4.4.1 Temps d'exécution de la tâche. Le temps d'exécution de la tâche est calculé à partir du moment où les 32 sphères de l'EV 3D sont affichées à l'écran, jusqu'au moment où l'utilisateur appuie sur l'écran pour valider sa sélection (appui sur la sphère désirée). Le temps moyen pour effectuer cette tâche de sélection est de 8111 ms (CI [7735; 8637]) avec la technique Circulation contre 7569 ms (CI [7219; 7943]) avec la technique DichotoZoom (cf. figure 4). La technique DichotoZoom permet donc aux participant de réduire leur temps de sélection de 5,93 % (CI [-0,72; 11,98]) par rapport à la technique Circulation. Le calcul du ratio DichotoZoom sur Circulation nous donne un ratio de 0,94 (CI [0,88; 1,01]). La technique DichotoZoom permet donc aux utilisateurs de réduire leur temps de sélection d'environ 6 %. Cependant, l'intervalle de confiance dépasse légèrement 1. Par conséquent, la différence de temps entre les deux techniques n'est pas très conséquente.

4.4.2 Temps d'exécution en fonction de la série. Les participants devaient réaliser 3 fois successivement le même exercice avec chaque technique. La Figure 5 présente les temps d'exécution de la tâche en fonction de l'exercice réalisé. Ainsi nous pouvons constater qu'il y a un effet d'apprentissage plus important pour la technique DichotoZoom où les participants ont effectué les tâches en 8170 ms (CI [7668; 8728]) lors du premier exercice, alors qu'ils ont effectué les mêmes tâches en 7389 ms (CI [7067; 7754]) lors du deuxième exercice. Le ratio entre ces deux premiers exercices montre que

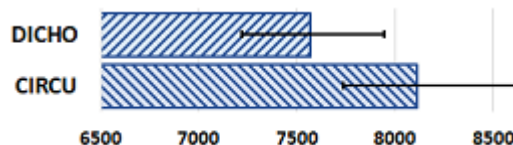


FIGURE 4: Temps d'exécution de la tâche en ms pour les techniques Circulation et DichotoZoom et intervalle de confiance à 95 %

les participants ont été 9,24 % (CI [5,50; 13,35]) plus rapide lors du second exercice par rapport au premier. De plus, nous pouvons voir que cet effet d'apprentissage s'estompe rapidement car entre les deuxième et troisième exercices, les performances des participants progressent beaucoup plus faiblement avec 7189 ms (CI [6795; 7567]) lors du troisième exercice, soit un gain moyen de 1,58 % (CI [-1,25; 3,63]) entre le deuxième et troisième exercice. D'autre part, cet effet d'apprentissage est beaucoup moins marqué pour la technique Circulation où les participants ont réalisé les tâches en 8390 ms (CI [7930; 9039]), 8099 ms (CI [7697; 8873]), et 7846 ms (CI [7443; 8193]) pour les 3 exercices consécutifs, soit des gains de 3,33 % (CI [0,57; 7,26]) et 2,63 % (CI [0,7; 7,2]) respectivement entre les exercices 1 et 2, puis 2 et 3.

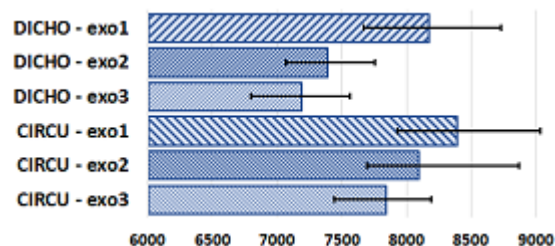


FIGURE 5: Temps d'exécution de la tâche en ms pour chaque technique (Circulation vs DichotoZoom) en fonction de l'exercice et intervalle de confiance à 95 %

4.4.3 Temps d'exécution de la tâche en fonction de la position. La Figure 5 présente l'évolution du temps de sélection en fonction de la position de la cible à atteindre dans l'ensemble des éléments proposés. Dans le cas de la technique Circulation (marqueurs bleus sur la Figure 5), nous pouvons constater que le temps de sélection est dépendant de la position dans l'ensemble : En effet, la droite de régression linéaire obtenue sur la moyenne des temps de sélection en fonction de la position dans l'ensemble des éléments donne un coefficient de corrélation très élevé ($R = 0,99$). L'équation obtenue ($y = 4031,2 + 262,8x$) nous permet de prédire qu'à chaque position supplémentaire dans l'ensemble, le temps de sélection sera impacté d'environ 263 ms supplémentaires.

En revanche, le temps de sélection d'une cible n'est pas directement lié à sa position dans l'ensemble lorsque l'on utilise la technique DichotoZoom (marqueurs rouges sur la Figure 6). En effet, la droite de régression linéaire calculée sur la moyenne des temps de sélection avec la technique DichotoZoom donne un coefficient de corrélation très faible ($R < 0,05$). Ceci peut s'expliquer par le

2. <http://www.aviz.fr/ci/>

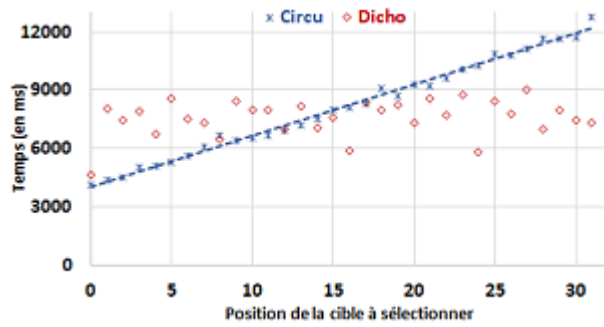


FIGURE 6: Évolution du temps moyen de sélection (en ms) en fonction de la position de la cible à atteindre

fait que la position de la cible n'est pas directement liée au nombre d'actions à effectuer pour l'atteindre : en effet, une cible très éloignée (par exemple, celle située à la 25^e position) peut être atteinte avec beaucoup moins d'actions (théoriquement 2 actions lorsque l'ensemble contient 32 éléments) qu'une autre cible beaucoup plus proche dans l'ensemble (par exemple, celle se situant à la 7^e position qui nécessite 5 actions avec un ensemble de même taille).

4.4.4 Temps d'exécution de la tâche en fonction du nombre d'actions à réaliser. Selon le nombre de cibles présélectionnées et la position de la cible à atteindre, le nombre d'actions à réaliser avec DichotoZoom pour atteindre la cible souhaitée peut varier. Le Tableau 1 résume, pour le cas de notre étude, les actions à réaliser en fonction de la position de la cible. Les caractères 'U' et 'D' indiquent un appui respectivement sur le bouton « UP » (pour sélectionner la zone la plus proche) ou sur le bouton « DOWN » (pour sélectionner la zone la plus éloignée). Nous pouvons constater que dans le cas où 32 cibles forment l'ensemble des éléments présélectionnées, l'utilisateur aura au maximum 5 actions de navigation entre les zones à réaliser pour atteindre la cible qu'il souhaite. Lorsque celle-ci se trouve au premier plan, l'utilisateur n'aura plus qu'à la pointer sur l'écran du Smartphone pour la sélectionner. Cette dernière action finale de pointage n'est pas comptée dans les actions que nous prenons en compte dans cette section.

TABLE 1: Actions de sélection à réaliser avant de pouvoir pointer la cible souhaitée.

1		9	UD	17	D	25	DD
2	UUUUD	10	UDUUD	18	DUUUD	26	DDUUD
3	UUUD	11	UDUD	19	DUUD	27	DDUD
4	UUUDD	12	UDUDD	20	DUUDD	28	DDUDD
5	UUUD	13	UDD	21	DUD	29	DDD
6	UUUDUD	14	UDDUD	22	DUDUD	30	DDUD
7	UUUD	15	UDDD	23	DUDD	31	DDDD
8	UUDDD	16	UDDDD	24	DUDDD	32	DDDDD

La Figure 7 présente l'évolution du temps de sélection en fonction du nombre d'actions à réaliser. Pour la technique DichotoZoom (en rouge sur la Figure 7), la droite de régression linéaire obtenue avec les moyennes des temps de sélection donne un coefficient de

corrélation très élevé ($R > 0.97$). Ce coefficient très proche de 1 confirme la corrélation entre le temps de sélection et le nombre d'actions de sélection à réaliser pour la technique DichotoZoom.

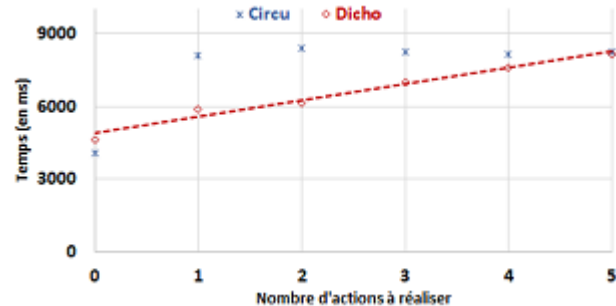


FIGURE 7: Évolution du temps moyen de sélection en fonction du nombre d'actions à réaliser

L'équation de la droite de régression linéaire pour la technique DichotoZoom ($y = 4887,8 + 675,35 x$) montre que chaque nouvelle action nécessite 675ms supplémentaires à l'utilisateur. A noter qu'avec la technique Circulation, comme il faut faire une action pour avancer d'une position, la droite de régression obtenue pour le nombre d'action est la même que celle obtenue pour les 5 premières positions (cf. Figure 6). Le coût d'une nouvelle action de sélection est donc de 265ms pour la technique Circulation.

Cette différence de temps entre les deux techniques peut s'expliquer par le fait qu'avec la technique DichotoZoom, après chaque action effectuée, les objets restants sont divisés en deux ensembles. L'utilisateur peut donc difficilement anticiper si l'objet sera dans l'ensemble proche ou éloigné du point de vue de l'utilisateur. Avec Circulation, comme la sélection avance d'objet en objet par ordre croissant dans la profondeur, il est plus facile pour l'utilisateur d'anticiper le fait que l'objet souhaité est encore éloigné et donc effectuer rapidement plusieurs actions successives.

La technique DichotoZoom apparaît donc comme plus rapide que la technique Circulation, dans le cas où le nombre d'objets présents dans l'ensemble de départ est conséquent. Plus le nombre d'objets est grand et plus l'écart sur le nombre d'actions à effectuer sera élevé et donc à l'avantage de la technique DichotoZoom.

4.4.5 Correction et erreur de sélection. Pour évaluer la précision des deux techniques, nous avons mesuré deux aspects : le taux d'annulation d'une action et le taux d'erreur de sélection.

Nous avons calculé le taux d'annulation d'une action comme étant le nombre de fois où les participants ont utilisé le bouton « retour arrière » (pour annuler la sélection précédemment effectuée) par rapport au nombre d'actions de sélection effectuées au cours de la tâche. Avec la technique Circulation, les participants ont annulé une de leur action de sélection dans 1,24 % (CI [0,76; 1,95]) des cas alors qu'ils ont réalisé cette même action dans 2,72 % (CI [0,99; 6,16]) des cas avec la technique DichotoZoom. D'autre part, nous pouvons voir sur la figure 8 que l'intervalle de confiance de la technique DichotoZoom est bien plus grand que celui de la technique Circulation. Ceci nous indique qu'il y a donc une plus

grande variabilité entre participants pour l'utilisation de l'annulation d'une action avec DichotoZoom qu'avec Circulation.

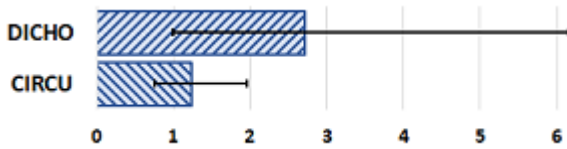


FIGURE 8: Taux d'annulation d'une action (en %) et intervalle de confiance à 95 %

Le taux d'erreur de sélection est le nombre de fois où les participants ont effectué une erreur de sélection par rapport au nombre total de tâche à réaliser. Les participants ont effectué beaucoup moins d'erreurs avec la technique Circulation qu'avec la technique DichotoZoom. Les participants n'ont fait aucune erreur avec Circulation, alors que le taux d'erreur est de 4,46 % (CI [1,06; 11,27]) avec DichotoZoom. Le taux d'erreur élevé de DichotoZoom s'explique par un taux d'erreur très élevé pour deux participants : en effet, les participants 3 et 5 ont effectué respectivement 18,48 % et 24,47 % d'erreur. Cependant, nous pouvons aussi noter que 7 des 12 participants n'ont commis aucune erreur avec DichotoZoom.

5 AVEC QUELLE MODALITÉ D'INTERACTION ?

Nous avons pu voir, au moyen de l'étude préliminaire, que la sélection au moyen d'un algorithme de dichotomie ne perturbait pas l'utilisateur et lui permettait de sélectionner plus rapidement une cible après peu d'apprentissage de la technique. Cette seconde étude vise à trouver la meilleure modalité d'interaction permettant à l'utilisateur d'interagir avec cette technique de sélection sur un Smartphone.

5.1 Modalités d'interaction étudiées

La technique DichotoZoom fonctionne de manière discrète, et nécessite deux actions distinctes de la part de l'utilisateur pour pouvoir sélectionner la zone adéquate (sélection de la zone proche ou de la zone éloignée du point de vue de l'utilisateur). Pour affiner la technique d'interaction, nous avons étudié l'utilisation de différentes modalités d'interaction contribuant à une distinction rapide entre les deux zones. Nous avons donc choisi d'étudier la sélection par actions discrètes (e.g. appuyer sur un bouton) et par action continue (e.g. un geste distinctif), et ce pour deux cas : dans une forme physique (utilisation du Smartphone lui-même) ou numérique (interaction directe sur l'écran tactile du Smartphone). Ainsi, la combinaison de ces deux caractéristiques nous a amené à étudier 4 modalités (cf. Tableau 2).

Leur fonctionnement respectif avec la technique DichotoZoom est le suivant :

- **Button** est la modalité basée sur les boutons physiques du Smartphone. Ce sont les boutons de volume qui servent à choisir entre la zone proche (bouton « DOWN ») ou éloignée (bouton « UP ») du point de vue de l'utilisateur. Le bouton

TABLE 2: Modalités d'interaction évaluées pour la sélection d'une cible au moyen de la technique DichotoZoom

		Interaction	
		Discrète	Continue
Dispositif	Physique <i>Smartphone</i>	Button	Tilt
	Logiciel <i>Écran tactile</i>	Touch	Swype

« Retour arrière » permet d'annuler la dernière sélection effectuée. Il s'agit de la même modalité que celle utilisée dans la première étude.

- **Touch** est basée sur l'utilisation de boutons logiciels affichés sur l'écran. Les boutons sont affichés sur le bord de l'écran pour ne pas perturber l'utilisation de l'EV 3D. Les boutons de sélection de la zone proche ou éloignée se situent dans les coins haut et bas du côté du pouce, de manière à pouvoir être facilement accessibles pour le pouce de la main tenant le smartphone. Le bouton d'annulation de la dernière action se situe en bas de l'écran dans le coin opposé.
- **Tilt** est la modalité gestuelle basée sur les inclinaisons réalisées directement avec le Smartphone. Les inclinaisons sont calculées en fonction d'une position neutre du Smartphone que nous avons définie comme étant celle du Smartphone tenu à l'horizontal, avec l'écran tactile orienté vers le ciel. Une action d'inclinaison consiste à passer de la position neutre à une inclinaison dans la direction voulue, puis à revenir à la zone neutre. L'utilisateur penche le Smartphone, écran vers lui pour choisir la zone qui est proche de son point de vue dans l'EV 3D ; et penche le Smartphone écran vers l'avant pour choisir la zone la plus éloignée. Pour annuler la dernière action effectuée, l'utilisateur doit pencher le Smartphone vers la gauche.
- **Swype** est la modalité gestuelle directement utilisée sur l'écran tactile. L'utilisateur effectue des gestes du haut vers le bas ou du bas vers le haut pour sélectionner respectivement la zone proche ou celle éloignée du point de vue de l'utilisateur. Un geste de la droite vers la gauche est utilisé pour annuler la dernière action réalisée.

5.2 Matériel et méthode

Cette seconde étude est basée sur le même type de tâche que la première : i.e. sélectionner une cible parmi les 32 sphères affichées. La procédure et les consignes étaient aussi identiques à la première étude. Les mêmes dispositifs technologiques ont été utilisés.

Douze personnes (6 femmes et 6 hommes) ont participé à cette étude. La moyenne d'âge était de 31,4 ans (écart-type de 4 ans) et 3 des participants étaient gauchers. Tous étaient des utilisateurs quotidiens d'un Smartphone, et aucun n'avait participé à la première étude.

Les 12 participants ont effectués une session avec chacune des 4 techniques d'interaction proposées. L'ordre de présentation des techniques a été contrebalancé avec un carré latin de manière à ce que chaque technique soit testée autant de fois à chaque position. Lors de la première étude, nous avons vu que les performances

des participants évoluaient principalement entre les deux premiers exercices et restaient quasiment stables lors du troisième exercice. C'est pourquoi, nous avons choisi pour cette étude de ne réaliser que 2 fois chaque exercice de 32 tâches. Dans chaque répétition, la position de la cible à atteindre différait d'une tâche à une autre de manière à tester l'ensemble des 32 positions de l'ensemble des objets proposés. Au total, nous avons donc récolté les résultats de 3072 tâches de sélection (12 participants \times 4 sessions \times 2 répétitions \times 32 tâches).

Après chaque session, les participants remplissaient un questionnaire SUS [7] permettant de qualifier l'utilisabilité de la technique. Enfin, à la fin des quatre sessions, il leur était demandé de classer les quatre techniques par ordre de préférence.

5.3 Résultats

Pour analyser nos résultats, nous nous appuyons sur les mêmes techniques d'estimation que celles utilisées pour l'étude 1.

5.3.1 Temps d'exécution de la tâche. Le calcul de la moyenne du temps d'exécution de la tâche pour chaque modalité montre que les participants ont été plus performant avec la modalité Touch (6968 ms, CI [6626; 7395]) qu'avec la modalité Button (7566 ms, CI [7274; 8039]) déjà utilisée lors de la première étude. En revanche, les temps moyen d'exécution avec les modalités Swype (9022 ms, CI [8566; 9488]) et Tilt (9144 ms, CI [8743; 10176]) sont bien plus élevés qu'avec les deux autres techniques. La Figure 9 présente les ratios des trois nouvelles modalités par rapport à celle utilisée lors de la première étude (Button). Il apparait que seule la modalité Touch a un ratio inférieur à 1 (*moyenne* = 0,921 et CI [0,896; 0,950]), ce qui montre que cette modalité est plus efficace que la modalité Button. En revanche, les deux autres modalités ont des ratios largement supérieurs à 1 (*Swype / Button* = 1,195 avec un CI [1,148; 1,243] et *Tilt / Button* = 1,211 avec CI [1,156; 1,3]), ce qui montre que les participants ont toujours effectué les tâches plus lentement avec ces deux modalités par rapport à la modalité Button initialement utilisée.

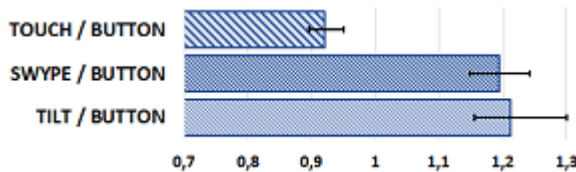


FIGURE 9: Ratio des temps de sélection et intervalles de confiance à 95 %. Les ratios représentent le temps mis avec les différentes modalités par rapport au temps mis avec la modalité Button déjà utilisée dans l'étude 1.

Cette différence de temps de sélection entre les différentes modalités est étayée par l'analyse du temps de sélection d'une zone en fonction du nombre d'actions à réaliser (cf. Figure 10). On peut constater que les droites de régression linéaire des techniques Touch et Button ont des pentes (respectivement $y = 681,81x + 4241,5$ et $y = 700,65x + 4799,5$) moins élevées que celles des techniques Tilt et Swype (respectivement $y = 1052,7x + 4872,6$ et $y = 1181,1x + 4313,3$). Une action de sélection avec une des

techniques « continue » est donc plus couteuse en temps qu'avec une technique de « discrète ».

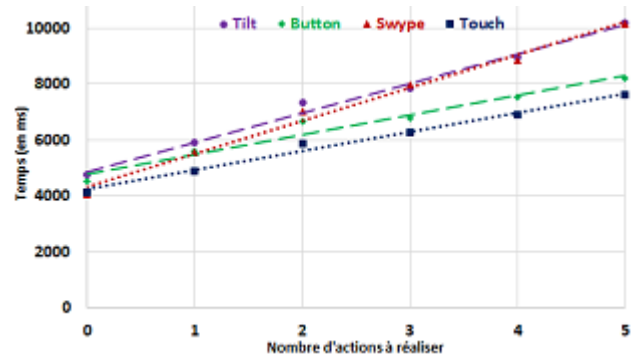


FIGURE 10: Temps de sélection d'une zone par modalité en fonction du nombre d'actions à réaliser

5.3.2 Précision. Comme pour la première étude, nous avons évalué la précision des techniques, en calculant le taux d'annulation et d'erreur de sélection. En ce qui concerne le taux d'erreur de sélection, les modalités Touch et Button s'avèrent être les plus précises avec des taux d'erreurs de seulement 0,26 % (CI [0; 0,65]) et 0,13 % (CI [0; 0,39]). Les modalités Swype avec 0,66 % (CI [0,13; 1,31]) et Tilt avec 0,57 % (CI [0,13; 1,44]) ont des taux d'erreur un peu plus élevés. Cependant, les taux d'erreurs pour chaque modalité restent assez faible (moins de 1 %). D'autre part, les intervalles de confiance de chaque modalité (cf. Figure 11) se chevauchent. On ne peut donc pas parler de différence significative entre les quatre modalités pour les erreurs.

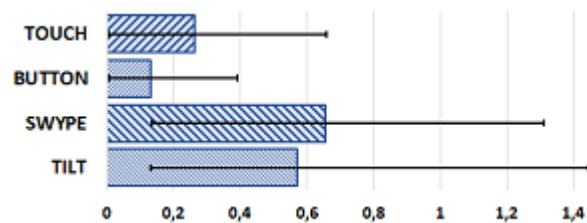


FIGURE 11: Taux d'erreurs de sélection et intervalles de confiance à 95 %.

La précision des modalités se mesure surtout avec le taux d'annulation. Dans le cas de Touch et Button, le retour arrière est assez peu utilisé (respectivement 0,54 %, CI [0,21; 1,16] et 0,29 %, CI [0,16; 0,44]). En revanche, il est beaucoup plus utilisé dans le cas des modalités Swype (2,64 %, CI [1,12; 4,78]) et Tilt (2,78 %, CI [1,64; 4,77]). Les intervalles de confiance de chaque modalité (cf. Figure 12) sont quasiment disjoints entre les modalités Touch et Button d'une part, et Swype et Tilt d'autre part. Il y a donc une différence assez marquée sur l'utilisation du retour arrière entre les deux premières modalités citées et les deux dernières.

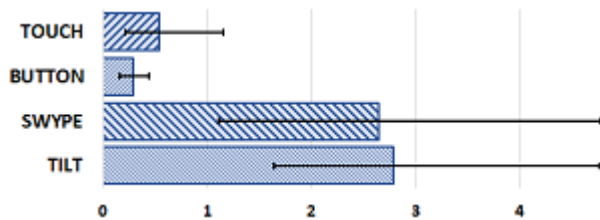


FIGURE 12: Taux d'annulation des actions et intervalles de confiance à 95 %.

5.3.3 Retours utilisateurs. Les scores obtenus au moyen du questionnaire SUS montrent une meilleure utilisabilité pour les deux techniques dites « Discrètes ». En effet, les modalités Button et Touch obtiennent des scores de 92 et peuvent être qualifiées comme « Excellente » au regard des qualificatifs employés au regard des résultats [7]. Les modalités Swype et Tilt obtiennent des scores plus faibles (respectivement 78 et 74), mais ces scores permettent quand même de les juger comme « Bonnes ». A la fin de l'expérimentation, les participants devaient classer les 4 techniques selon leur préférence. La Figure 13 présente ces classements en donnant pour chaque technique le nombre de fois où les participants les ont classées respectivement en 1^{ère}, 2^e, 3^e ou 4^e position.

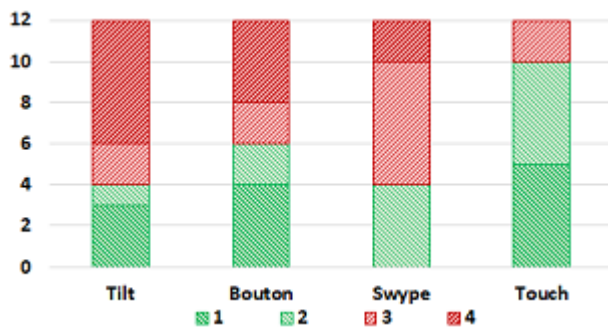


FIGURE 13: Ordre des techniques : nombre de fois où chaque technique a été positionnée en 1^{ère} (vert foncé), 2^e (vert clair), 3^e (rouge clair) et 4^e (rouge foncé) position

La préférence des participants pour les deux versions à base de boutons est cohérente avec les scores SUS obtenus : La modalité Touch ressort comme étant la plus appréciée en étant classée 5 fois en première position et autant de fois en seconde position. La modalité Button est apparue 4 fois comme la technique la plus appréciée par les participants. A l'opposé, la modalité Tilt a été classée 6 fois en dernière position, alors que la modalité Swype n'est jamais ressortie comme étant la technique préférée des participants.

Enfin, à la fin de l'expérimentation, les participants pouvaient faire des commentaires sur les différentes techniques. Ils ont expliqué leur préférence pour les techniques Touch et Button par leur simplicité d'utilisation (dit respectivement 10 et 6 fois), et que ces techniques ne nécessitaient pas de faire des efforts lors de leur utilisation (dit 2 fois pour chaque). La technique Button a cependant été jugée « ancienne » (3 fois) ou « déjà vu » (2 fois). Bien que les

techniques Swype et Tilt aient été jugées « Fun » « Moderne » (respectivement 4 et 7 fois), ces techniques sont apparues plus difficile et compliquée à utiliser (4 et 3 fois) et pouvaient nécessiter un effort de concentration (1 fois chacune) et étaient moins précise (1 fois chacune).

6 CONCLUSION

La sélection d'un objet dans un EV 3D comprenant une forte densité d'objets s'effectue souvent en deux phases : une première sélection d'objets dans une zone de l'EV 3D, puis une sélection plus fine à l'intérieur de cet ensemble d'éléments présélectionnés. La technique DichotoZoom que nous proposons permet de sélectionner rapidement un objet à l'intérieur de cet ensemble. Quel que soit le nombre d'éléments, le nombre d'actions à réaliser reste inférieur à la technique Circulation. D'autre part, cette technique nécessite simplement deux actions discrètes distinctes pour pouvoir effectuer la sélection, et peut donc être facilement manipulée au moyen de différentes modalités. De ce fait, de simples boutons (physiques ou logiciels) suffisent pour faire fonctionner rapidement cette technique. Notre seconde étude a montré que ces simples boutons étaient même la modalité d'interaction la plus efficace (en vitesse et en précision), et la préférée des utilisateurs pour utiliser la technique DichotoZoom.

A cette première étude sur les interactions, d'autres études suivront pour étudier les différents retours visuels qui pourraient être proposés aux utilisateurs pour minimiser au mieux la transition entre deux actions de l'utilisateur. En effet, lors de la sélection d'une zone, l'ensemble des objets encore sélectionnables est divisé par deux et deux nouveaux sous-ensembles sont créés. Par conséquent, il est important pour l'utilisateur de visualiser au mieux les éléments qu'il peut encore sélectionner et dans quelle zone de sélection se trouve l'objet souhaité. Nous étudierons donc différentes manières de présenter ces informations à l'utilisateur pour qu'il puisse l'utiliser dans des EV 3D concrets.

Enfin, cette technique étant simple d'utilisation et ne nécessitant pas de ressources particulières, elle pourrait être facilement utilisée sur différents dispositifs encore plus petite ou présentant peu de moyens d'interaction (par exemple, Smartwatches ou lunettes).

RÉFÉRENCES

- [1] Eliane Balaa, Mathieu Raynal, Youssef Bou Issa, and Emmanuel Dubois. 2014. Classification of Interaction Techniques in the 3D Virtual Environment on Mobile Devices. In *Proceedings, Part I, of the 6th International Conference on Virtual, Augmented and Mixed Reality: Designing and Developing Virtual and Augmented Environments - Volume 8525*. Springer-Verlag New York, Inc., New York, NY, USA, 3–13. DOI : http://dx.doi.org/10.1007/978-3-319-07458-0_1
- [2] Louis-Pierre Bergé, Emmanuel Dubois, and Mathieu Raynal. 2015. Design and Evaluation of an "Around the SmartPhone" Technique for 3D Manipulations on Distant Display. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction (SUI '15)*. ACM, New York, NY, USA, 69–78. DOI : <http://dx.doi.org/10.1145/2788940.2788941>
- [3] Louis-Pierre Bergé, Gary Perelman, Adrien Hamelin, Mathieu Raynal, Cédric Sanza, Minica Houry-Panchetti, Rémi Cabanac, and Emmanuel Dubois. 2014. Smartphone Based 3D Navigation Techniques in an Astronomical Observatory Context : Implementation and Evaluation in a Software Platform. *International Journal on Advances in Software* 7, no. 3 & 4 (décembre 2014), 551–566.
- [4] Lonni Besançon and Pierre Dragicevic. 2017. The Significant Difference between p-values and Confidence Intervals. In *29ème conférence francophone sur l'Interaction Homme-Machine, AFIHM* (Ed.). AFIHM, Poitiers, France, 10. <https://hal.inria.fr/hal-01562281> Alt.IHM.

- [5] Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics (I3D '97)*. ACM, New York, NY, USA, 35–ff. DOI: <http://dx.doi.org/10.1145/253284.253301>
- [6] Doug A Bowman, Donald B Johnson, and Larry F Hodges. 2001. Testbed evaluation of virtual environment interaction techniques. *Presence : Teleoperators & Virtual Environments* 10, 1 (2001), 75–95.
- [7] John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [8] Jeffrey Cashion, Chadwick Wingrave, and Joseph J LaViola Jr. 2012. Dense and dynamic 3d selection for game-based virtual environments. *IEEE transactions on visualization and computer graphics* 18, 4 (2012), 634–642.
- [9] Florian Daiber, Lianchao Li, and Antonio Krüger. 2012. Designing Gestures for Mobile 3D Gaming. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12)*. ACM, New York, NY, USA, Article 3, 4 pages. DOI: <http://dx.doi.org/10.1145/2406367.2406371>
- [10] N-T Dang, Hong-Ha Le Hong-Ha Le, and M Tavanti. 2003. Visualization and interaction on flight trajectory in a 3D stereoscopic environment. In *Digital Avionics Systems Conference, 2003. DASC'03. The 22nd*, Vol. 2. IEEE, 9–A. DOI: <http://dx.doi.org/10.1109/DASC.2003.1245905>
- [11] Alex Olwal Steven Feiner. 2003. The flexible pointer : An interaction technique for selection in augmented and virtual reality. In *Proc. UIST'03*. 81–82.
- [12] Bernd Froehlich, Jan Hochstrate, Verena Skuk, and Anke Huckauf. 2006. The GlobeFish and the GlobeMouse : Two New Six Degree of Freedom Input Devices for Graphics Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 191–199. DOI: <http://dx.doi.org/10.1145/1124772.1124802>
- [13] Tovi Grossman and Ravin Balakrishnan. 2006. The Design and Evaluation of Selection Techniques for 3D Volumetric Displays. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06)*. ACM, New York, NY, USA, 3–12. DOI: <http://dx.doi.org/10.1145/1166253.1166257>
- [14] Adrien Hamelin and Emmanuel Dubois. 2015. Design and Evaluation of an Interaction Technique for Volume Selection in a 3D Point Cloud. In *Proceedings of the 27th Conference on L'Interaction Homme-Machine (IHM '15)*. ACM, New York, NY, USA, Article 3, 10 pages. DOI: <http://dx.doi.org/10.1145/2820619.2820622>
- [15] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. A Survey of Design Issues in Spatial Input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology (UIST '94)*. ACM, New York, NY, USA, 213–222. DOI: <http://dx.doi.org/10.1145/192426.192501>
- [16] Regis Kopper, Felipe Bacim, and Doug A Bowman. 2011. Rapid and accurate 3D selection by progressive refinement. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*. IEEE, 67–74.
- [17] Jiandong Liang. 1993. JDCAD : A highly interactive 3D modeling system. In *Proceedings of the 3rd International Conference on CAD and Computer Graphics*. 217–222.
- [18] Jiandong Liang and Mark Green. 1993. Geometric modeling using six degrees of freedom input devices. In *3rd Int'l Conference on CAD and Computer Graphics*. Citeseer, 217–222.
- [19] Asier Marzo, Benoît Bossavit, and Martin Hachet. 2014. Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. ACM, 13–16.
- [20] Daniel Medeiros, Lucas Teixeira, Felipe Carvalho, Ismael Santos, and Alberto Raposo. 2013. A tablet-based 3D interaction tool for virtual engineering environments. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. ACM, 211–218.
- [21] Annette Mossel, Benjamin Venditti, and Hannes Kaufmann. 2013. 3DTouch and HOMER-S : intuitive manipulation techniques for one-handed handheld augmented reality. In *Proceedings of the Virtual Reality International Conference : Laval Virtual*. ACM, 12.
- [22] Masaki Naito, Buntarou Shizuki, Jiro Tanaka, and Hiroshi Hosobe. 2009. Interaction techniques using a spherical cursor for 3d targets acquisition and indicating in volumetric displays. In *13th International Conference Information Visualisation*. IEEE, 607–612. DOI: <http://dx.doi.org/10.1109/IV.2009.70>
- [23] Gary Perelman, Marcos Serrano, Mathieu Raynal, Celia Picard, Mustapha Derras, and Emmanuel Dubois. 2015. The Roly-Poly Mouse : Designing a Rolling Input Device Unifying 2D and 3D Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 327–336. DOI: <http://dx.doi.org/10.1145/2702123.2702244>
- [24] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The go-go interaction technique : non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*. ACM, 79–80.
- [25] Mathieu Raynal, Emmanuel Dubois, and Bénédicte Schmitt. 2013. Towards unification for pointing task evaluation in 3D desktop virtual environment. In *Human Factors in Computing and Informatics*. Springer, 562–580.
- [26] Bénédicte Schmitt, Mathieu Raynal, and Emmanuel Dubois. 2012. A composite approach to evaluate two interaction techniques for a 3D pointing task. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 159–160.
- [27] Gary R. VandenBos. 2009. *Publication Manual of the American Psychological Association*. American Psychological Association Sixth Edition. Washington. DC.
- [28] Wesley Willett, Bernhard Jenny, Tobias Isenberg, and Pierre Dragicevic. 2015. Lightweight relief shearing for enhanced terrain perception on interactive maps. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3563–3572.
- [29] Hans Peter Wyss, Roland Blach, and Matthias Bues. 2006. iSith-Intersection-based spatial interaction for two hands. In *3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*. IEEE, 59–61.
- [30] Lingyun Yu, Konstantinos Efstathiou, Petra Isenberg, and Tobias Isenberg. 2012. Efficient structure-aware selection techniques for 3D point cloud visualizations with 2DOF input. *IEEE Transactions on Visualization & Computer Graphics* 12 (2012), 2245–2254.
- [31] Lingyun Yu, Konstantinos Efstathiou, Petra Isenberg, and Tobias Isenberg. 2016. CAST : Effective and efficient user interaction for context-aware selection in 3D particle clouds. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 886–895.
- [32] Shumin Zhai, William Buxton, and Paul Milgram. 1994. The “Silk Cursor” : investigating transparency for 3D target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 459–464.