



Interaction 2D/3D en réalité augmentée sur table

Carole Plasson

► To cite this version:

Carole Plasson. Interaction 2D/3D en réalité augmentée sur table. Vision par ordinateur et reconnaissance de formes [cs.CV]. Université Grenoble Alpes [2020-..], 2021. Français. NNT : 2021GRALM010 . tel-03323142

HAL Id: tel-03323142

<https://theses.hal.science/tel-03323142>

Submitted on 20 Aug 2021

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Informatique

Arrêté ministériel : 25 mai 2016

Présentée par

Carole PLASSON

Thèse dirigée par **Laurence NIGAY**, Professeure des universités,
Université Grenoble Alpes
et codirigée par **Yann LAURILLAU**, Maître de conférences HDR,
Université Grenoble Alpes
et **Dominique CUNIN**, Enseignant-Chercheur,
École Supérieure d'Art et Design de Valence

préparée au sein du **Laboratoire d'Informatique de Grenoble**
dans l'**École Doctorale Mathématiques, Sciences et**
Technologies de l'Information, Informatique

Interaction 2D/3D en Réalité Augmentée sur Table

2D/3D Interaction in Tabletop Augmented Reality

Thèse soutenue publiquement le 4 mai 2021
devant le jury composé de :

Monsieur Géry CASIEZ

Professeur des universités, Université de Lille, Rapporteur

Madame Nadine COUTURE

Professeure des universités, École d'ingénieurs ESTIA, Rapporteuse

Monsieur Georges-Pierre BONNEAU

Professeur des universités, Université Grenoble Alpes, Président

Monsieur Emmanuel DUBOIS

Professeur des universités, Université Toulouse 3, Examineur

Monsieur Martin HACHET

Directeur de recherche, INRIA Bordeaux-Sud-Ouest, Examineur



Résumé

Cette thèse s'inscrit dans le domaine de l'Interaction Humain-Machine (IHM) et est dédiée à l'interaction en Réalité Augmentée (RA) sur table. Permettre de visualiser des informations virtuelles 3D liées à une table qui sert de support physique à l'interaction est un des principaux avantages de la RA sur table. Les systèmes de RA sur table sont ainsi une alternative pertinente aux logiciels souvent complexes sur ordinateurs de bureau pour créer et manipuler des modèles 3D (modélisation 3D, conception assistée par ordinateur, urbanisme, architecture). Actuellement, ces systèmes mettent en œuvre des techniques d'interaction dans l'air héritées de la RA classique et imitant les manipulations d'objets physiques. Bien qu'intuitives et directes, ces techniques sont souvent peu précises et fatigantes pour un usage prolongé. Les systèmes de RA sur table autorisent également l'interaction tactile sur la table qui est plus précise mais moins intuitive pour des tâches 3D que l'interaction dans l'air. Dans ce contexte, la question de recherche de nos travaux concerne la conception de techniques de sélection d'objets 3D en RA sur table qui soient précises, peu fatigantes et faciles à prendre en main et à utiliser.

L'espace de conception à explorer est vaste et comprend des techniques de sélection basées sur l'interaction tactile sur la table, sur l'interaction dans l'air au-dessus de la table et sur l'interaction hybride alliant ces deux espaces d'interaction que sont la table et l'espace au-dessus. L'exploration de ce vaste espace de conception a donné lieu à trois contributions. Notre première contribution est d'identifier quel dispositif de RA, entre un casque et un dispositif portable (téléphone / tablette), est le plus efficace pour la visualisation et la sélection d'objets 3D en RA sur table. Les résultats montrent de meilleures performances et retours utilisateurs avec un casque de RA. Nous nous sommes donc concentrés sur ce dispositif dans la suite de nos travaux. Notre deuxième contribution concerne la sélection d'objets distants. Nous avançons une nouvelle technique bi-manuelle *RayLens* pour améliorer la précision du pointage d'objets distants. *RayLens* combine un rayon et une lentille grossissante 2D déplaçable en 3D. Cette technique s'est révélée performante, peu fatigante et particulièrement appréciée par les utilisateurs grâce à sa simplicité de prise en main et d'utilisation. Enfin, notre troisième contribution porte sur la sélection précise d'objets à portée de main et approfondit l'utilisation de la table comme espace d'interaction en entrée. En comparant expérimentalement les trois espaces d'interaction fournis en RA sur table (espace 2D de la table, espace 3D au-dessus de la table, espace hybride), nous soulignons l'importance du support de la table pour améliorer la précision et réduire la fatigue. Nous montrons également l'efficacité des techniques hybrides qui combinent avec succès la précision de l'interaction tactile et la rapidité de l'interaction dans l'air.

Abstract

This thesis contributes to the research field of Human-Computer Interaction (HCI). The focus of the research is on tabletop Augmented Reality (AR) interaction. Allowing visualization of 3D virtual information linked to a table that acts as a physical support for interaction is one of the main advantages of tabletop AR. Tabletop AR systems thus define a relevant alternative to the often complex desktop software used to create and visualize 3D models (3D modeling, CAD, urban planning, architecture). Currently, these systems use mid-air interaction techniques inherited from classical AR that mimic the manipulation of physical objects. Although intuitive and direct, these techniques are often imprecise and tiring for long-term use. Tabletop AR systems also allow tactile interaction on the table which is more precise but less intuitive for 3D tasks than 3D mid-air interaction. In this context, the research question of our work concerns the design of 3D object selection techniques in tabletop AR that are precise, not tiring, and easy to handle and use.

The design space to be explored is vast and includes selection techniques based on tactile interaction on the table, mid-air interaction above the table, and hybrid interaction combining the two interaction spaces i.e. the table and the space above the table. The exploration of this vast design space gave rise to three contributions. Our first contribution is to identify which AR device, between a head-mounted display (HMD) and a handheld device (smartphone/tablet), is the most efficient for the visualization and selection of 3D objects. The experimental results show better performance and user feedback with an HMD. We have therefore focused on tabletop AR with an HMD in our subsequent work. Our second contribution concerns the selection of distant 3D objects. We propose a new bi-manual technique, namely *RayLens*, to improve the accuracy of distant object pointing. *RayLens* combines a ray and a 2D magnifying lens that can be moved in 3D. This technique has proven to be efficient, not tiring, and particularly appreciated by the users thanks to its ease of use and intuitiveness. Finally, our third contribution focuses on the precise selection of objects at hand and further develops the use of the table as an input space for interaction. By experimentally comparing the three interaction spaces provided in tabletop AR (2D table space, 3D space above the table, hybrid space), we highlight the importance of the table support in improving accuracy and reducing fatigue. We also show the effectiveness of hybrid techniques that successfully combine the accuracy of tactile interaction with the speed of mid-air interaction.

Remerciements

Je tiens tout d'abord à remercier mes encadrants Laurence Nigay, Yann Laurillau et Dominique Cunin. Ils ont toujours été disponibles et à l'écoute pour m'aider dans cette thèse. Ce fut un vrai plaisir de travailler avec eux.

Je remercie Nadine Couture et Géry Casiez, rapporteurs de ce manuscrit, pour leur lecture attentive et leurs remarques constructives. Je remercie également Georges-Pierre Bonneau, Emmanuel Dubois et Martin Hachet d'avoir participé à mon jury et à la soutenance. Merci au jury pour ses nombreuses questions et ses critiques très intéressantes.

Un grand merci à mes collègues et amis de l'équipe IIHM pour ces superbes années. Merci pour leur aide, leur soutien et les très bons moments passés ensemble.

Merci à ma famille d'avoir toujours cru en moi.

Enfin, un immense merci à Alex qui m'a encouragé et soutenu tout au long de cette thèse. Il a été un soutien sans faille, et un papa très présent et attentionné pour Lily, en particulier lors de cette fin de thèse.

Sommaire

Introduction	1
I Dispositifs d’affichage pour la réalité augmentée sur table	7
1 État de l’art sur les dispositifs d’affichage et interactions associées	9
1.1 Dispositifs portables : tablettes et téléphones	9
1.1.1 Cas d’utilisation	9
1.1.2 Interactions	11
1.2 Casques et lunettes de RA	13
1.2.1 Cas d’utilisation	13
1.2.2 Interactions	15
1.3 Études comparatives existantes	17
1.4 Bilan	20
2 Contribution : tablette ou casque pour la réalité augmentée sur table	21
2.1 Éléments de conception : trois techniques étudiées	22
2.2 Implémentation d’un système de RA sur table avec un casque	26
2.2.1 Scène 3D virtuelle	26
2.2.2 Suivi précis des mouvements de l’utilisatrice	27
2.2.3 Lien entre le monde virtuel et le monde physique	28
2.3 Étude expérimentale 1 : recherche et sélection d’objets visibles en environnements épars	30
2.3.1 Objectif de l’étude et hypothèses	30
2.3.2 Implémentation	31
2.3.3 Système expérimental	31
2.3.4 Tâche	32
2.3.5 Protocole	34
2.4 Étude expérimentale 1 : résultats	35
2.4.1 Résultats quantitatifs	35
2.4.2 Résultats qualitatifs	38
2.5 Étude expérimentale 1 : discussion	40
2.6 Étude expérimentale 2 : sélection d’objets 3D occultés en environnements denses	43
2.6.1 Objectif de l’étude et hypothèse	43
2.6.2 Système expérimental	44

2.6.3	Tâche	44
2.6.4	Protocole	45
2.7	Étude expérimentale 2 : résultats	47
2.7.1	Résultats quantitatifs	47
2.7.2	Résultats qualitatifs	50
2.8	Étude expérimentale 2 : discussion	51
3	Résumé de l'étude sur les dispositifs d'affichage en réalité augmentée sur table	54

II Sélection d'objets distants : amélioration du raycasting pour la sélection en environnements denses 58

4	État de l'art sur les facilitations de pointage pour le raycasting	60
4.1	Facilitations de pointage pour le raycasting	60
4.1.1	Techniques indépendantes des cibles	61
4.1.2	Techniques sensibles aux cibles	66
4.1.3	Bilan sur les facilitations de pointage existantes pour le raycasting .	72
4.2	Les lentilles en RA et RV	75
4.2.1	Tâches d'interaction permises par les lentilles	76
4.2.2	Dimension des lentilles : plates ou volumétriques ?	80
4.2.3	Déplacement des lentilles	81
4.2.4	Bilan sur l'utilisation des lentilles en RA et RV	82
5	Contribution : extension du raycasting avec une lentille grossissante	83
5.1	Éléments de conception : <i>RayLens</i> et <i>RaySlider</i>	83
5.1.1	RayLens : une lentille grossissante couplée à un rayon virtuel	84
5.1.2	RaySlider : un curseur déplaçable sur un rayon virtuel	87
5.2	Étude expérimentale : évaluation de l'efficacité de <i>RayLens</i> face au raycas- ting et à <i>RaySlider</i>	89
5.2.1	Objectif de l'étude et hypothèses	89
5.2.2	Système expérimental	90
5.2.3	Tâche	93
5.2.4	Protocole	94
5.3	Étude expérimentale : résultats	95
5.3.1	Résultats quantitatifs	96
5.3.2	Résultats qualitatifs	101
5.3.3	Analyse des temps de déplacement de la lentille et du curseur	102
5.4	Étude expérimentale : discussion	103
5.4.1	Hypothèses	103
5.4.2	Contrôle du curseur, de la lentille, et du rayon	105

5.4.3	Avantages de la technique <i>RayLens</i>	106
6	Résumé de l'étude sur les améliorations du raycasting	108

III Sélection d'objets proches : trois espaces d'interaction disponibles en réalité augmentée sur table 110

7	État de l'art sur les techniques d'interaction 2D/3D en RA sur table : focus sur les espaces d'interaction en entrée	112
7.1	Espace 3D dans l'air au-dessus de la table	114
7.2	Espace 2D sur la table	114
7.3	Espace hybride : combiner l'espace 2D sur table et l'espace 3D dans l'air	117
7.3.1	Interaction avec des appareils mobiles	117
7.3.2	Réalité augmentée : cinq classes d'utilisation de l'espace hybride	117
7.4	Études comparatives existantes	121
7.5	Bilan	126
8	Contribution : Comparaison des trois espaces d'interaction disponibles	128
8.1	Éléments de conception : trois techniques étudiées	129
8.1.1	Espace de conception	129
8.1.2	Conception des techniques	131
8.2	Étude expérimentale 1 : sélection d'objets 3D visibles ancrés à la table dans un environnement épars	133
8.2.1	Objectif de l'étude et hypothèses	133
8.2.2	Système expérimental	133
8.2.3	Tâche	134
8.2.4	Protocole	136
8.3	Étude expérimentale 1 : résultats	137
8.3.1	Résultats quantitatifs	138
8.3.2	Résultats qualitatifs	144
8.4	Étude expérimentale 2 : sélection d'objets 3D occultés ancrés à la table dans un environnement dense	146
8.4.1	Objectif de l'étude	146
8.4.2	Système expérimental	146
8.4.3	Tâche	146
8.4.4	Protocole	147
8.5	Étude expérimentale 2 : résultats	148
8.5.1	Résultats quantitatifs	148
8.5.2	Résultats qualitatifs	151
8.6	Études expérimentales 1 et 2 : discussion	152
8.6.1	Confirmation de résultats existants	152

8.6.2 Nouveaux résultats : avantages et limitations de la décomposition de la tâche	153
9 Résumé de l'étude sur les trois espaces d'interaction	156
10 Mise en contexte : vers une installation artistique exploitant les trois espaces d'interaction disponibles en RA sur table	159
 IV Conclusion et perspectives	 167
11 Contributions	168
12 Perspectives	172
12.1 Perspectives à court terme	172
12.1.1 Perspectives à court terme sur la technique <i>RayLens</i>	172
12.1.2 Interaction hybride combinant interaction tactile et interaction dans l'air (technique <i>Mixed</i>)	174
12.2 Perspectives à long terme	175
12.2.1 Unification des travaux : techniques <i>RayLens</i> et <i>Mixed</i>	175
12.2.2 Combinaison de la RA sur table et des micro-gestes	175
12.2.3 Pointage : vers de nouvelles taxonomies	176
Bibliographie	178
Annexes	197
A Chapitre I – Étude expérimentale 1 : Questionnaire intermédiaire	197
B Chapitre I – Étude expérimentale 1 : Questionnaire final	198
C Chapitre I – Étude expérimentale 2 : Questionnaire intermédiaire	199
D Chapitre I – Étude expérimentale 2 : Questionnaire final	200
E Chapitre I – Étude expérimentale 2 : Questionnaire final	202
F Chapitre II – Questionnaire Intermédiaire	204
G Chapitre II – Questionnaire final	205
H Chapitre III – Questionnaire Intermédiaire	206
I Chapitre III – Questionnaire final	207

Table des figures

1	Démonstration finale du projet AP2. Image issue du rapport final du projet.	6
1.1	<i>Hover Pad</i> de Seifert et al. [142]. Images extraites de [142].	10
1.2	<i>G-Raff</i> de Kim et Nam [79]. Images extraites de [79].	10
1.3	Les deux modalités d'interaction en entrée comparées par [38] pour naviguer dans des visualisations 3D virtuelles au-dessus d'une table. Images extraites de la vidéo de présentation de [38].	11
1.4	Deux types d'interaction avec une tablette pour la sélection d'un objet virtuel 3D. Images extraites de [43] (à gauche) et [9] (à droite).	12
1.5	Système de RA de Nilsson et al. [116]. Images extraites de [116].	14
1.6	Système <i>GeoGate</i> de Ssin et al. [149]. Images extraites de [149].	14
1.7	<i>DesignAR</i> de Reipschläger et Dashselt [130]. Images extraites de [130].	15
1.8	Exemples d'interactions dans l'air proposées par [138] pour l'exploration de cartes. Images extraites de [138].	16
1.9	Interactions 3D dans l'air avec un casque de RA pour le pointage d'un objet virtuel 3D. Images extraites de [81] (à gauche) et de [24] (à droite).	16
1.10	Étude comparative de Bach et al. [6]. Images extraites de [6].	17
1.11	Étude comparative de Looser et al. [95]. Images modifiées de [95].	19
2.1	Les trois techniques comparées.	22
2.2	Sélection d'un cylindre avec HMD_DT : la sphère bleue qui suit le doigt entre en collision avec l'objet, un cylindre.	23
2.3	Sélection d'un cylindre avec HMD_DT : la sphère bleue qui suit le doigt entre en collision avec un cylindre. Une barre de progression est affichée pour indiquer le temps restant avant la fin de la validation du cylindre.	25
2.4	Système expérimental générique.	26
2.5	Placement des marqueurs réfléchissants.	27
2.6	Caméra OptiTrack type Flex 13.	28
2.7	Image cible détectée et suivie par le SDK Vuforia.	29
2.8	Image cible suivie par la tablette grâce au SDK Vuforia.	31
2.9	Système expérimental.	32
2.10	Scène virtuelle pour la première étude.	32
2.11	Position initiale des 3 techniques.	33
2.12	Sélection d'un cylindre avec HMD_RC : le rayon entre en collision avec un cylindre.	33
2.13	Phase de recherche – Effet d'interaction entre la zone et la technique sur le temps moyen de réalisation de la tâche.	36

2.14	Phase de sélection – Effet d’interaction entre la zone et la technique sur le temps moyen de réalisation de la tâche.	37
2.15	Distances moyennes parcourues par les mains et la tête des participants. . .	38
2.16	Stratégies observées pour la sélection avec HMD_DT.	42
2.17	Deux exemples de scènes denses pour l’étude 2.	44
2.18	Illustration de l’espacement de densité (DS)	46
2.19	Effet d’interaction entre la condition de visibilité de la cible et la technique sur le temps de réalisation de la tâche de sélection.	48
2.20	Effet d’interaction entre la taille de la cible et la technique sur le temps de réalisation de la tâche de sélection.	48
2.21	Effet d’interaction entre la condition de visibilité de la cible et la technique sur la distance parcourue par les participants.	50
2.22	Trajectoires du doigt de 3 participants	52
4.1	Extensions du raycasting de Grossman et Balakrishnan [63] pour la sélection d’objets occultés. Images extraites de [63].	62
4.2	<i>SQUAD</i> de Kopper et al. [84] et de Bacim et al. [7]. Images extraites de [84].	64
4.3	<i>Expand</i> de Cashion et al. [39]. Image extraite de [39].	65
4.4	<i>StickyRay</i> de Steinicke et al. [151]. Image extraite de [151].	66
4.5	<i>3DBubble</i> de Vanacken et al. [163]. Images extraites de [163].	67
4.6	<i>RayCursor</i> de Baloup et al. [15]. Images extraites de [15].	68
4.7	<i>SenseShapes</i> de Olwal et al. [117]. Images extraites de [117].	69
4.8	<i>Bubble Ray</i> de Lu et al. [97]. Images extraites de [97].	71
4.9	<i>GlyphLens</i> de Tong et al. [160]. Images extraites de [160].	76
4.10	Lentilles pour la visualisation et l’exploration de données [94]. Images extraites de [94].	77
4.11	Lentilles pour des tâches de sélection.	78
4.12	Lentille avec effet miroir de Lee et Park [90]. Image extraite de [90].	79
4.13	<i>Lens frustum..</i> Images extraites de [167].	80
4.14	Déplacement d’une lentille distante à l’aide d’un proxy à portée de main de l’utilisatrice [83]. Images extraites de [83].	81
5.1	Illustration de la lentille grossissante et de son champ de vision.	85
5.2	Gestion de l’occultation de la lentille : mise en transparence des objets placés devant la lentille.	86
5.3	Deux gestions différentes de l’occultation du curseur.	87
5.4	Système expérimental.	90
5.5	Bouton Arduino placé sur le bâton physique et permettant de valider une sélection facilement avec le pouce.	91
5.6	RaySlider : slider utilisé pour déplacer le curseur le long du rayon	91

5.7	Le boîtier, contenant le nécessaire pour le fonctionnement du bouton du bâton, est attaché au poignet d'une utilisatrice.	92
5.8	Exemple d'une scène 3D créée pour l'étude expérimentale.	93
5.9	Temps moyen de réalisation des tâches de sélection pour chacune des techniques.	96
5.10	Effet d'interaction entre la taille de la cible et la technique sur le temps moyen de réalisation des tâches de sélection.	97
5.11	Effet d'interaction entre l'espacement de densité et la technique sur le temps moyen de réalisation des tâches de sélection.	98
5.12	Effet d'interaction entre la taille de la cible et la technique sur le nombre moyen d'erreurs.	99
5.13	Effet d'interaction entre l'espacement de densité et la technique sur le nombre d'erreurs.	100
5.14	[RTLX] Charge de travail totale entre 0 et 100.	101
5.15	Temps moyen passé à déplacer la lentille et le curseur (en bleu) sur le temps total passé à sélectionner les cibles. En rouge, le temps moyen pendant lequel la lentille et le curseur ne sont pas déplacés.	103
7.1	<i>Balloon Selection</i> de Benko et Feiner [19]. Images extraites de [19].	115
7.2	<i>Toucheo</i> de Hachet et al. [67]. Image extraite de [67].	115
7.3	<i>Triangle Cursor</i> de Strothoff et al. [155]. Image extraite de [155].	116
7.4	Exemples d'application de la modalité <i>Air+Touch</i> de Chen et al. [42]. Image extraite de [42].	117
7.5	Environnement immersif hybride de Benko et al. [20]. Images extraites de [20].	118
7.6	<i>PaperLens</i> de Spindler et Dachsel [146]. Images extraites de [146].	119
7.7	Interaction hybride de Xiao et al. [177]. Images extraites de [177].	120
7.8	Interaction hybride de Menzner et al. [109]. Images extraites de [109].	121
7.9	Étude comparative de Bruder et al. [34, 35] et Colley et al. [47]. Images modifiées de [35].	123
7.10	Étude comparative de Benko et Feiner [19]. Images extraites de [19].	123
7.11	Étude comparative de Veit et al. [165]. Image extraite de [165].	124
8.1	Exemple d'objets flottants 3D étudié dans [19] : Ajout d'ombres aux objets afin de les ancrer à la table.	129
8.2	Les trois techniques comparées, classées selon i) leurs espaces d'interaction en entrée, et ii) si elles impliquent ou non une décomposition de la tâche.	130
8.3	Décomposition d'une tâche de sélection 3D en une tâche 2D (étape 1) suivie d'une tâche 1D (étape 2).	131
8.4	Réalisation d'une tâche de sélection avec chacune des trois techniques.	132

8.5	Système expérimental.	134
8.6	[Étude 1] Exemple d'une scène 3D avec une pile de boîtes.	134
8.7	Jauge affichée pour indiquer le temps restant pour valider la sélection.	135
8.8	Dimensions et formes des cibles.	136
8.9	Hauteur de la boîte dans la pile.	136
8.10	Temps total de réalisation de la tâche (T_{Total}), temps pour atteindre la cible pour la première fois ($T_{Atteindre}$), et temps pour sélectionner la base (T_{Etape1}).	139
8.11	Temps total de sélection selon la forme de la boîte.	140
8.12	Temps total selon les trois dimensions de la cible.	141
8.13	Effet d'interaction entre la hauteur de la cible et la technique sur le temps total de réalisation des tâches de sélection. Le temps moyen pour atteindre les cibles pour la première fois est indiqué en pointillé.	142
8.14	Effet d'interaction entre la hauteur et la forme de la cible et la technique sur le nombre moyen d'erreurs avant validation de la cible.	143
8.15	[RTLX] Charge de travail totale entre 0 et 100.	144
8.16	[Étude 2] Exemple d'une scène 3D densément peuplée avec occultation de la boîte cible.	146
8.17	Formes des cibles pour la seconde étude.	147
8.18	Temps total de réalisation de la tâche (T_{Total}), temps pour atteindre la cible pour la première fois ($T_{Atteindre}$) et temps pour sélectionner la base (T_{Etape1}).	149
8.19	Effet d'interaction entre la forme de la cible et la technique sur le temps total de réalisation des tâches de sélection.	150
8.20	Effet d'interaction entre la forme de la cible et la technique sur le nombre d'erreurs moyen de sélection.	151
8.21	[RTLX] Charge de travail totale entre 0 et 100.	152
10.1	Déroulement du jeu.	160
10.2	Animation de l'élément sélectionné.	161
10.3	Niveau complété avec succès.	161
10.4	Technique <i>DT</i>	162
10.5	Technique <i>Mixed</i>	162
10.6	Technique <i>Balloon</i>	163
10.7	Niveau libre.	163
10.8	Storyboard complet du jeu interactif inspiré du livre Flatland.	165
10.9	Storyboard complet du jeu interactif inspiré du livre Flatland (suite).	166
12.1	Scène non abstraite et évaluation de <i>RayLens</i> dans un contexte grand public avec des utilisateurs aux profils variés.	172
12.2	Rotations de la lentille.	174

12.3 Espace hybride combinant l'espace 2D de la table et l'interaction par micro- gestes.	176
--	-----

Introduction

Motivation

Cette thèse s’inscrit dans le domaine de l’Interaction Humain-Machine (IHM) et est dédiée à l’interaction en Réalité Augmentée (RA) sur table.

De nombreux domaines tels que la modélisation 3D, la conception assistée par ordinateur (CAO) ou l’urbanisme utilisent des logiciels pour visualiser et interagir avec des informations 3D sur des ordinateurs de bureau. La prise en main et l’utilisation de ces logiciels sont complexes : la complexité réside dans la manipulation d’information 3D selon des vues en 2D (élément tridimensionnel sur un support bidimensionnel). Permettant de visualiser des éléments 3D de manière directe, simple et intuitive sur un support tridimensionnel, la RA apparaît comme une alternative efficace à l’ordinateur de bureau pour ce type de tâche 3D.

De plus, contrairement à l’immersion totale offerte en réalité virtuelle, la RA permet à une utilisatrice de visualiser et d’interagir avec des informations 2D et 3D virtuelles tout en gardant un lien direct avec le monde réel. Elle est ainsi particulièrement adéquate dans des milieux professionnels, en autorisant par exemple la prise de notes sur papier ainsi que la collaboration et les échanges même impromptus entre collègues. La RA permet également d’augmenter l’environnement de travail. Par exemple, des écrans virtuels peuvent être combinés à un écran physique [119]. Ainsi, les avantages potentiels d’un environnement de travail hybride, combinés à la possibilité de visualiser des informations virtuelles 3D directement (support tridimensionnel), nous poussent à croire que dans un futur proche des dispositifs de RA pourront être utilisés toute la journée au travail [18, 119, 125]. Or l’avènement de ce futur ne sera possible que par la conception de techniques d’interaction facilitant la prise en main ainsi que l’usage efficace et prolongé d’environnements de travail hybrides en RA.

Poursuivant cette vision, la *RA sur table*, qui permet la combinaison d’un support physique et d’informations virtuelles affichées au-dessus, est particulièrement pertinente. En effet, au bureau nous visualiserons et interagirons certainement plus régulièrement avec des objets virtuels au-dessus d’une table de réunion ou de notre bureau, qu’avec des objets flottant dans une pièce. La RA sur table permet de profiter de la surface tangible offerte par ce contexte de bureau. Par ailleurs, cette

surface n'a pas besoin d'être interactive : des études récentes [177] montrent qu'il est possible avec les casques de RA actuels de réaliser des interactions tactiles sur des surfaces non-interactives sans matériel supplémentaire (en utilisant les caméras déjà intégrées au casque). La RA sur table a trois principaux avantages :

1. Le support physique de la table peut réduire la fatigue causée par l'interaction dans l'air [19, 165].
2. La table apporte également un repère physique pour aider à la perception de la scène 3D [37, 72].
3. Il est possible de lier des informations 3D virtuelles et des informations 2D sur la table. Ceci est particulièrement pertinent dans des domaines tels que la cartographie [24, 172] ou l'urbanisme [142, 162] pour combiner par exemple un plan 2D (physique ou virtuel) d'une ville et le modèle 3D associé.

Nos travaux se concentrent donc sur les systèmes de RA sur table et nous étudions les techniques d'interaction avec des informations 3D pour de tels systèmes.

Question de recherche

Si nous considérons les nombreux domaines applicatifs de la RA sur table, comme la modélisation 3D, la CAO ou l'urbanisme, les critères pour l'interaction incluent la précision pour manipuler les objets virtuels 3D tout en étant peu fatigante pour un usage prolongé. De plus pour une large adoption de la RA sur table, les techniques d'interaction doivent aussi être faciles à prendre en main et à utiliser par des non-experts.

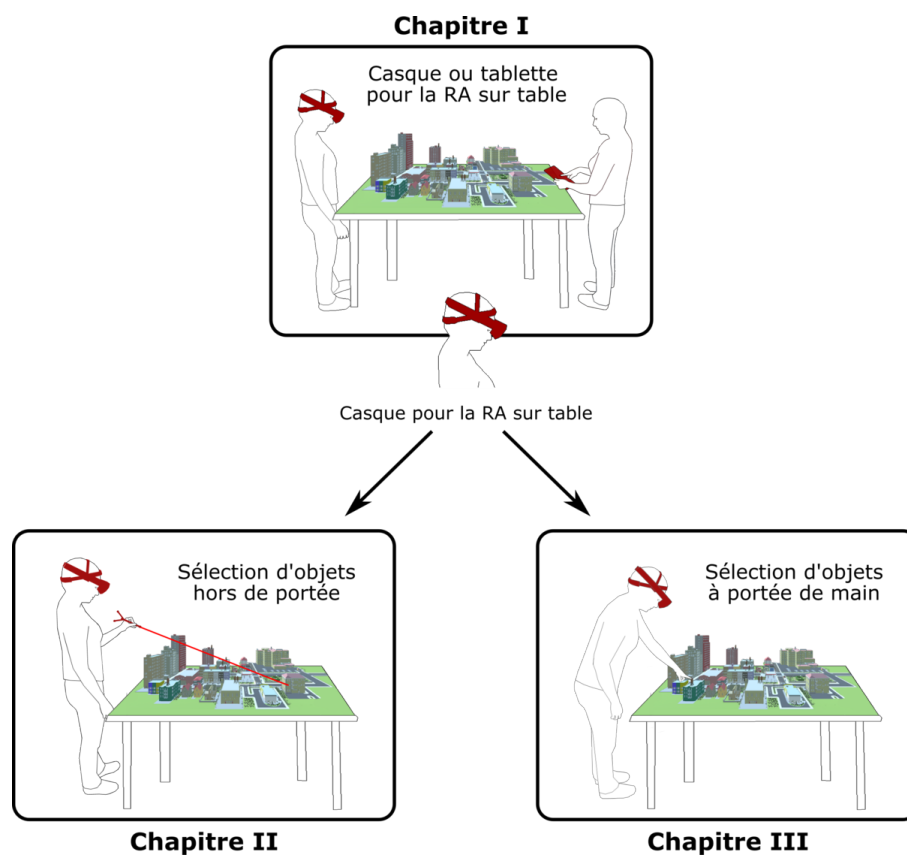
Pour répondre à ces critères, la RA sur table ouvre un vaste ensemble de possibilités de techniques d'interaction à explorer : des techniques d'interaction dans l'air au-dessus de la table, des techniques d'interaction tactile sur la table, et des techniques d'interaction hybride alliant les deux espaces d'interaction que sont la table et l'espace 3D au-dessus de la table. Les systèmes de RA sur table peuvent ainsi mettre en œuvre des techniques d'interaction dans l'air héritées de la RA classique et imitant la manipulation d'objets physiques. Bien qu'intuitives et directes, ces techniques sont souvent peu précises et fatigantes pour un usage prolongé. Les systèmes de RA sur table autorisent également l'interaction tactile sur la table qui est plus précise mais moins intuitive et directe pour des tâches 3D que l'interaction dans l'air.

Face à ce vaste espace de possibilités, nous nous concentrons sur les techniques d'interaction pour la tâche fondamentale de sélection d'un objet virtuel 3D dans

une scène virtuelle 3D. La question de recherche de nos travaux concerne alors la conception de techniques de sélection d'objets 3D en RA sur table qui soient précises, peu fatigantes et faciles à prendre en main et à utiliser pour profiter au mieux des avantages de la RA sur table.

Contributions

Notre exploration de l'espace des possibilités des techniques d'interaction en RA sur table a donné lieu aux trois études suivantes :



Étude sur les dispositifs d'affichage pour la RA sur table.

En RA sur table, l'affichage des informations virtuelles 3D est réalisé soit par un dispositif portable (téléphone / tablette), soit par un casque de RA. Notre première contribution est d'identifier lequel de ces dispositifs de RA est le plus efficace pour la visualisation et la sélection d'objets 3D en RA sur table. Les résultats de deux études expérimentales montrent de meilleurs performances et retours utilisateurs avec un casque de RA. Nous nous sommes donc concentrés sur ce dispositif dans la suite de nos travaux.

Étude sur la sélection précise d’objets distants : amélioration du raycasting pour la sélection en environnements denses.

Notre deuxième contribution concerne la sélection d’objets distants. Dans les environnements virtuels 3D, les techniques de sélection d’objets distants sont principalement basées sur la métaphore du pointeur virtuel. Cependant, une implémentation standard de cette métaphore (c’est-à-dire un simple rayon) connaît de nombreuses limitations telles que la sélection de petits d’objets ou la sélection en environnements denses où des objets peuvent être occultés. Nous avançons alors une nouvelle technique bi-manuelle appelée *RayLens* pour améliorer le pointage par raycasting. La technique *RayLens* est composée de deux principaux composants virtuels : un rayon et une lentille 2D grossissante déplaçable en 3D. Cette technique s’est révélée performante, peu fatigante et particulièrement appréciée par les participants grâce à sa simplicité de prise en main et d’utilisation.

Étude sur la sélection d’objets proches : trois espaces d’interaction disponibles en RA sur table.

Enfin, notre troisième contribution porte sur la sélection précise d’objets à portée de main et approfondit l’utilisation de la table comme espace d’interaction en entrée. Nous présentons deux études expérimentales comparant trois techniques de sélection 3D, chacune interagissant dans un des 3 espaces disponibles en RA sur table : une technique d’interaction tactile, une technique d’interaction directe dans l’air et une technique combinant l’interaction tactile et l’interaction dans l’air. Les résultats de ces études soulignent l’importance de la table pour améliorer la précision et réduire la fatigue. Nous montrons également l’efficacité de techniques hybrides qui combinent avec succès la précision de l’interaction tactile et la rapidité de l’interaction dans l’air.

Structure du document

Ce manuscrit est structuré selon nos trois études. Nous dédions un chapitre par étude.

- **Chapitre I.** Dispositifs d’affichage pour la RA sur table.
- **Chapitre II.** Sélection d’objets distants : amélioration du raycasting pour la sélection en environnements denses.
- **Chapitre III.** Sélection d’objets proches : trois espaces d’interaction disponibles en RA sur table.

Chaque chapitre suit la même structure globale suivante : une première section présente l'état de l'art en lien avec l'étude, une seconde section présente notre contribution sur cette étude, et une troisième section conclut cette étude. Chacune de nos contributions comporte trois facettes : une facette conceptuelle, une facette réalisation logicielle et une facette expérimentale. Face au vaste espace de possibilités décrit précédemment, la facette conceptuelle de nos travaux vise l'établissement d'espaces de conception qui structurent les éléments de conception des techniques d'interaction ; et permet d'explorer de façon systématique l'espace des possibilités. Nous soulignons aussi la facette ingénierie de l'interaction humain-machine de nos contributions. Les techniques d'interaction conçues doivent être développées de façon robuste et avec une latence minimale pour permettre des évaluations expérimentales valides et impliquent le développement de logiciels complexes et distribués sollicitant plusieurs technologies. Enfin, la facette expérimentale de nos travaux est importante, avec au total cinq études expérimentales réalisées. Nous remercions les nombreux participants pour leur temps et leur implication lors de ces expérimentations.

Enfin, nous concluons ce manuscrit en résumant nos contributions. Nous présentons également des perspectives de recherche à court terme et à long terme pour l'interaction en RA sur table.

Contexte

Les travaux décrits dans ce manuscrit s'inscrivent dans le cadre d'une collaboration avec l'École Supérieure d'Art et Design (ESAD) de Valence et dans le projet *ANR AP2 Autour du plan 2D*¹. Les travaux de thèse ont été financés par l'ESAD pendant un an et par l'ANR AP2 pendant deux ans. La collaboration avec l'ESAD s'est concrétisée par le co-encadrement de la thèse par Dominique Cunin, enseignant-chercheur à l'ESAD.

Le projet AP2, qui s'est achevé en septembre 2019 avait pour objectif de développer de nouvelles façons de visualiser, interagir et plus globalement collaborer autour de l'information numérique présentée sur une surface horizontale ou verticale. Les partenaires du projet AP2 étaient la société IMMERSION à Bordeaux (coordinatrice du projet) et le laboratoire IRIT à Toulouse (équipe ELIPSE en Interaction Humain-Machine). AP2 a étudié une situation d'interaction très courante dans de nombreux domaines applicatifs : un espace 2D sur une surface horizontale ou verti-

1. <https://ra2019.anr.fr/fr/une-agence-de-financement-sur-projet/ap2>

cale dédié au référentiel commun des utilisateurs, le plan 2D, et de multiples espaces interactifs 2D/3D manipulés par les utilisateurs pour explorer et annoter le monde numérique associé au plan 2D. Le projet a abordé ces problématiques en considérant le besoin d’une interaction fluide des utilisateurs avec un environnement défini par un accroissement des données numériques et des dispositifs connectés. AP2 a traité donc spécifiquement la problématique de recherche d’interaction multimodale multi-dispositifs et collaborative, tout en visant l’intégration des technologies nécessaires aux échanges d’information et à la reconnaissance des différents dispositifs dans deux cas d’usage que sont la revue de projet architectural et la consultation publique de données 2D et 3D associées à un territoire (un campus par exemple).

Dans le cadre du projet AP2, les travaux décrits dans ce manuscrit identifient tout d’abord le dispositif de RA à privilégier pour l’affichage 3D dans ce type de systèmes. Nos travaux apportent également une meilleure connaissance des avantages et limitations des espaces d’interaction offerts par la combinaison d’une surface horizontale tangible et d’informations virtuelles 3D. Enfin, nous proposons une nouvelle technique, appelée *RayLens*, permettant d’explorer et de sélectionner des objets dans une scène 3D dense. L’interaction intuitive qu’elle fournit, ainsi que ses bonnes performances en environnements denses, laissent envisager l’utilisation de cette technique d’interaction dans des applications publiques telles que les projets architecturaux et la consultation publique d’un plan 2D/3D de campus comme proposés dans le projet AP2. Nous soulignons néanmoins que la technique *RayLens* n’a pas été intégrée dans les deux contextes applicatifs de AP2, car elle a été conçue après la fin du projet. La démonstration finale du projet AP2 schématisée par la Figure 1 a souligné l’apport du support physique de la table et a impliqué la technique *EXHI-bit*² du LIG comme interface tangible déformable permettant de positionner un bâtiment de façon directe sur la carte.

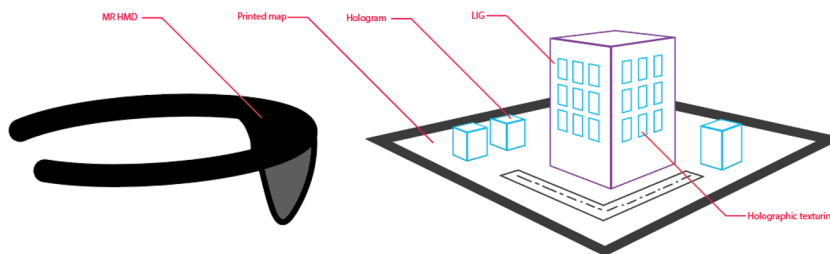


FIGURE 1 – Démonstration finale du projet AP2. Image issue du rapport final du projet AP2.

2. <http://iihm.imag.fr/publication/OMN17a/>

Chapitre I

Dispositifs d’affichage pour la réalité augmentée sur table

Introduction

L'un des principaux intérêts de la réalité augmentée (RA) sur table est de permettre de lier une vue 2D sur la table avec une vue augmentée (généralement 3D). Après étude des systèmes existants, nous faisons deux constats. Tout d'abord, nous soulignons deux types de dispositifs couramment utilisés pour l'affichage de la vue augmentée : les dispositifs portables tels que les tablettes ou les téléphones, et les casques de RA. De plus, l'utilisation d'un dispositif plutôt qu'un autre change considérablement la manière dont l'utilisatrice va appréhender et interagir avec l'information.

Ces deux constats motivent les travaux décrits dans ce premier chapitre³ dédié à l'interaction soit avec un dispositif portable soit avec un casque en RA sur table. La Section 1 dresse un état de l'art des systèmes de RA sur table structuré selon ces deux dispositifs tandis que la Section 2 présente deux études expérimentales comparant des techniques d'interaction avec une tablette et avec un casque.

Sommaire

1	État de l'art sur les dispositifs d'affichage et interactions associées	9
1.1	Dispositifs portables : tablettes et téléphones	9
1.2	Casques et lunettes de RA	13
1.3	Études comparatives existantes	17
1.4	Bilan	20
2	Contribution : tablette ou casque pour la réalité augmentée sur table	21
2.1	Éléments de conception : trois techniques étudiées	22
2.2	Implémentation d'un système de RA sur table avec un casque	26
2.3	Étude expérimentale 1 : recherche et sélection d'objets visibles en environnements épars	30
2.4	Étude expérimentale 1 : résultats	35
2.5	Étude expérimentale 1 : discussion	40
2.6	Étude expérimentale 2 : sélection d'objets 3D occultés en environnements denses	43
2.7	Étude expérimentale 2 : résultats	47
2.8	Étude expérimentale 2 : discussion	51
3	Résumé de l'étude sur les dispositifs d'affichage en réalité augmentée sur table	54

3. Les travaux présentés dans ce chapitre ont été en partie publiés à la conférence ACM-ISS 2019, the 2019 ACM International Conference on Interactive Surfaces and Spaces. Carole Plasson, Dominique Cunin, Yann Laurillau, & Laurence Nigay. Tabletop AR with HMD and Tablet : A Comparative Study for 3D Selection. <https://doi.org/10.1145/3343055.3360760>.

1. État de l’art sur les dispositifs d’affichage et interactions associées

Dans cette section, nous présentons des exemples de réalité augmentée (RA) sur table en focalisant sur le contexte d’utilisation du système et sur le dispositif qui prend en charge l’augmentation. L’objectif de cette section est à la fois 1) de souligner l’apport de la RA sur table à des contextes applicatifs variés afin de motiver l’importance de nos travaux, et 2) de mettre en exergue deux classes de RA sur table, l’une basée sur un dispositif portable et l’autre sur un casque de RA. Nous présentons d’abord les systèmes exploitant un dispositif portable puis ceux exploitant un casque de RA. Pour terminer, nous décrivons deux études qui ont été menées pour comparer ces dispositifs d’affichage.

1.1 Dispositifs portables : tablettes et téléphones

1.1.1 Cas d’utilisation

Les dispositifs portables tels que les tablettes et les téléphones sont souvent utilisés en RA. Comme une “fenêtre sur le monde virtuel”, ils permettent de percevoir l’espace virtuel seulement à travers leur écran 2D. Selon la métaphore de la *lentille magique* [22], ces dispositifs peuvent offrir à l’utilisatrice une vue alternative de l’environnement ou ajouter de nouvelles informations.

Ce type de RA est régulièrement utilisé dans le contexte d’architecture d’intérieur et de cartes augmentées en urbanisme et en géographie. Le système de Olwal et Henrysson [118] augmente le plan d’un appartement pour visualiser en 3D les décorations intérieures des pièces. L’utilisatrice “balaie” le plan 2D affiché sur la table avec son téléphone pour visualiser la décoration d’une pièce sur l’écran du téléphone. De nombreux autres systèmes augmentent une carte 2D avec du contenu 3D virtuel tel que des bâtiments. Un exemple est le système *MetaDESK* de Ullmer et Ishii [162] qui permet de coupler le plan 2D d’une ville affiché sur une table avec une vue 3D perspective des bâtiments sur une tablette. Plus récemment, Wiehr et al. proposent *ARTopos* [172] un système collaboratif pour les activités extérieures telles que le vélo ou la randonnée en haute montagne. Leur système combine une carte

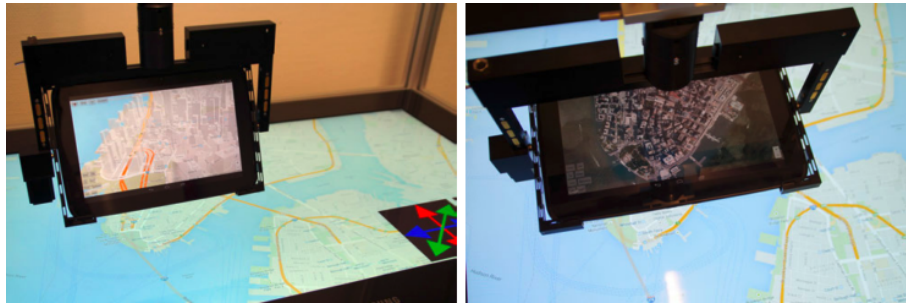


FIGURE 1.1 – *Hover Pad* de Seifert et al. [142] : combinaison d'une table interactive et d'une tablette avec placement autonome. Images extraites de [142].

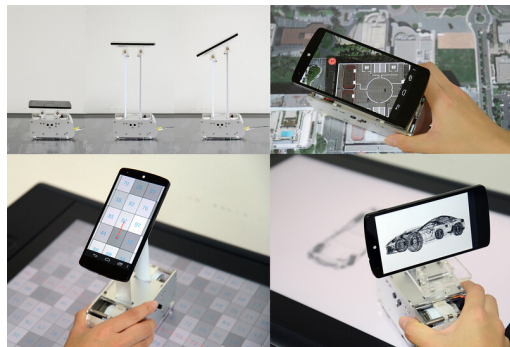


FIGURE 1.2 – *G-Raff* de Kim et Nam [79] : bloc tangible mécanique comme support d'un téléphone au-dessus d'une table interactive. Images extraites de [79].

topographique 2D sur la table et un modèle 3D du terrain affiché sur les téléphones des utilisateurs. Ces derniers peuvent alors visualiser les montagnes et planifier les chemins qu'ils prendront.

Selon le même principe de cartes augmentées, Seifert et al. [142] proposent *Hover Pad* un système composé d'une grande table interactive et d'une tablette attachée à un bras articulé. Ce système est présenté à la Figure 1.1. La tablette peut se déplacer soit de manière manuelle (l'utilisatrice la tient et la déplace) soit de manière autonome grâce à son bras articulé. Ce déplacement autonome permet plusieurs scénarios d'usage. Par exemple, l'utilisatrice fait un toucher direct sur la table et la tablette vient se positionner automatiquement au-dessus de ce point. Pour explorer la carte, l'utilisatrice peut aussi spécifier une rue dans une barre de recherche sur la table interactive et la tablette se déplace au-dessus de cette rue (plan 2D) et affiche les bâtiments qui la composent (vue 3D perspective). Seifert et al. [142] proposent également d'autres applications pour le système *Hover Pad* telles que l'exploration de données médicales volumétriques ou les jeux en réalité mixte.

Un autre exemple qui n'implique pas de toujours tenir le dispositif portable est *G-Raff* [79] de Kim et Nam. Ce système mécanique, présenté à la Figure 1.2, per-

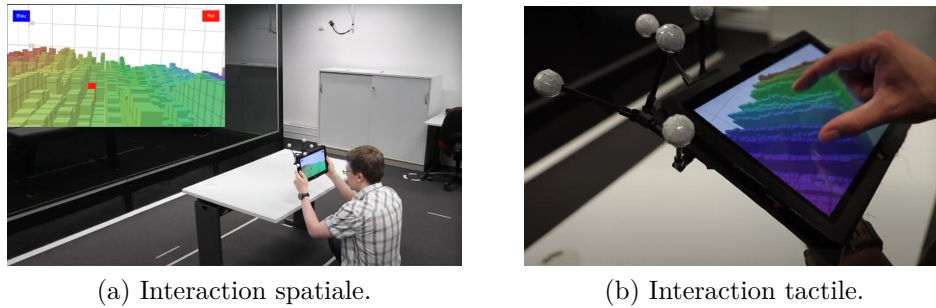


FIGURE 1.3 – Les deux modalités d’interaction en entrée comparées par [38] pour naviguer dans des visualisations 3D virtuelles au-dessus d’une table. Images extraites de la vidéo de présentation de [38].

met de déplacer et d’orienter un téléphone sur une table, sans avoir à le tenir. De plus l’interaction tangible que cette technique offre permet de bénéficier des avantages de l’interaction tangible comme souligné dans [3]. En plus de l’exploration de cartes où ils combinent un plan 2D sur la table et une vue satellite sur le téléphone, Kim et Nam proposent d’autres applications comme l’exploration de modèle 3D. Plus général que des cartes et plans (plan d’un appartement, d’une ville, d’un jeu, etc.), nous trouvons également des cas d’utilisation en visualisation d’information. Par exemple, Chen et al. ont développé *MARVisT* [43], un outil en RA sur tablette permettant à des utilisateurs non-experts de créer facilement des visualisations avec glyphes combinées à des objets du monde réel (y compris une carte papier sur une table).

1.1.2 Interactions

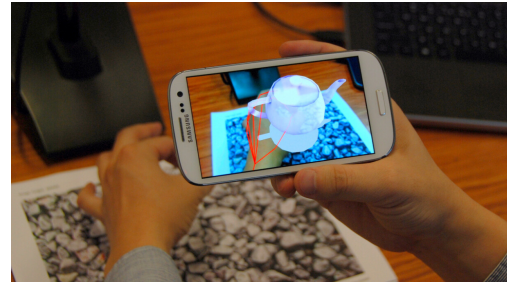
Navigation dans la scène virtuelle

Comme illustré par les exemples précédents, le dispositif portable est très souvent utilisé comme lentille magique pour présenter une autre vue de l’information présente sur la table (comme une vue satellite ou une vue 3D de bâtiments). Une des tâches principales permises par ces systèmes est donc l’exploration de l’espace d’information. Büschel et al. [38] ont comparé deux modalités d’interaction pour naviguer dans une visualisation 3D virtuelle au-dessus d’une table. Contrairement aux exemples précédents, la table sert uniquement de repère physique et non d’affichage d’une vue 2D. Les deux modalités étudiées sont les suivantes :

- L’interaction spatiale (Figure 1.3a), c’est-à-dire des mouvements physiques du dispositif portable dans l’espace autour et au-dessus de la table.



(a) Interaction tactile sur l'écran de la tablette.



(b) Interaction gestuelle derrière la tablette.

FIGURE 1.4 – Deux types d'interaction avec une tablette pour la sélection d'un objet virtuel 3D. Images extraites de [43] (à gauche) et [9] (à droite).

- L'interaction tactile sur l'écran du dispositif portable (Figure 1.3b).

Pour l'interaction spatiale, une caméra virtuelle permettant de voir la scène 3D virtuelle suit la position de la tablette. L'utilisatrice doit donc déplacer la tablette pour changer de point de vue. Pour zoomer sur une partie des données, l'utilisatrice rapproche simplement le dispositif vers la table.

Pour l'interaction tactile sur l'écran, l'utilisatrice n'a plus besoin de déplacer la tablette, la caméra est déplaçable par des interactions au toucher sur l'écran (gestes de glissement). Au-delà des déplacements de la caméra, des gestes sur l'écran permettent aussi de zoomer. Malgré une fatigue physique évidemment plus importante avec l'interaction spatiale, l'étude comparative a montré que pour une tâche de navigation, l'interaction spatiale offre une exploration des données plus intuitive et plus confortable. Les auteurs recommandent également l'utilisation de repères physiques (une table par exemple) qui fournissent des points physiques de référence à l'utilisatrice durant l'interaction.

Interactions avec la scène virtuelle

En plus de la navigation dans la scène virtuelle, l'utilisatrice peut également interagir avec les objets virtuels : annotation, sélection, manipulation, etc. Pour cela, plusieurs techniques d'interaction sont utilisées : les interactions tactiles sur l'écran du dispositif et les gestes dans l'air derrière le dispositif. Ces deux types d'interaction sont illustrés à la Figure 1.4.

Interactions tactiles sur l'écran. En plus de l'affichage de l'augmentation, le dispositif portable peut être utilisé comme support pour des interactions tactiles en entrée. Kim et Nam [79] proposent des interactions couplées entre le contenu de la table et le téléphone. L'utilisatrice peut par exemple copier/coller et manipuler des objets

affichés sur la table par des interactions sur le téléphone. Pour cela, l'utilisatrice place le téléphone au-dessus d'objets numériques affichés sur la table. Ces objets sont ensuite copiés sur le téléphone via une interaction au toucher sur l'écran. Une fois ces objets visibles sur le dispositif, l'utilisatrice les réorganise, les déplace les uns par rapport aux autres par des gestes de glissement sur l'écran. Les objets présents sur l'écran peuvent être déplacés n'importe où sur la table en déplaçant le téléphone. Pour terminer, les objets sont replacés sur la table (à la position du téléphone) par une interaction au toucher direct sur l'écran.

Des tâches de sélection ou d'annotation du contenu augmenté sont également réalisées sur le dispositif portable par des interactions tactiles [21, 43, 79, 172]. Avec *ARTopos* [172], l'utilisatrice touche simplement l'écran du téléphone pour annoter l'augmentation 3D d'une carte topographique 2D (placement de points d'intérêt, traçage d'un chemin). Ainsi l'interaction avec le contenu 3D virtuel se fait donc sur la projection 2D du modèle 3D affichée sur l'écran du dispositif portable.

Interactions gestuelles dans l'air derrière le dispositif. Plusieurs études [9, 44, 74, 80] explorent la manipulation directe de modèles 3D, y compris pour des tâches complexes, grâce au suivi des mouvements de la main ou des doigts effectués derrière le dispositif portable (au niveau de la position de l'objet 3D).

Dans une étude [73] comparant l'interaction tactile et l'interaction gestuelle avec un téléphone pour la sélection d'objets 3D virtuels flottant dans l'air, Hürst et Van Wezel montrent que : (1) l'interaction gestuelle atteint un niveau élevé d'engagement des utilisateurs et obtient des retours qualitatifs positifs, mais (2) l'interaction tactile sur l'écran du téléphone est plus rapide. Ils ont conclu que ses mauvaises performances rend l'interaction gestuelle pour la manipulation directe d'objets virtuels "irréalisable dans des applications concrètes". Ainsi, avec un dispositif portable, des techniques d'interaction tactile et gestuelle peuvent être utilisées pour manipuler des objets virtuels. Cependant, pour la sélection d'objets virtuels, il est préférable d'utiliser l'interaction tactile sur l'écran comme illustrée à la Figure 1.4a.

1.2 Casques et lunettes de RA

1.2.1 Cas d'utilisation

Bien que les casques de RA soient souvent utilisés pour afficher des objets 3D flottant dans l'air, plusieurs études sur les systèmes d'information géographique

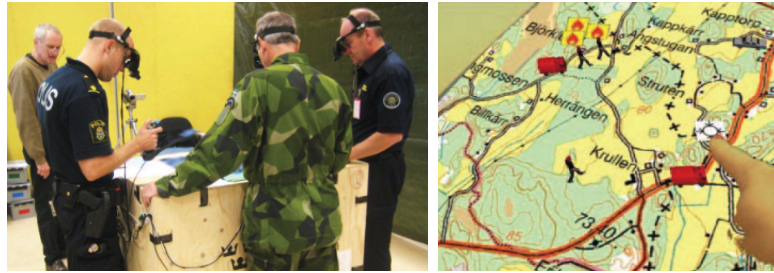


FIGURE 1.5 – Système de RA de Nilsson et al. [116] pour la planification de tâches en gestion de crises. Images extraites de [116].

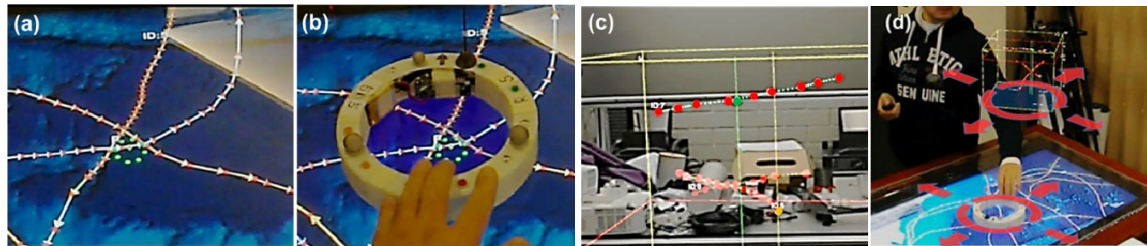


FIGURE 1.6 – Système *GeoGate* de Ssin et al. [149] pour la visualisation de données maritimes. Images extraites de [149].

3D utilisent un casque pour augmenter une carte 2D [25, 70]. Dans le domaine de la topographie, Bobrich [24] présente des *cartes augmentées* combinant une carte papier 2D et un modèle de terrain 3D projeté par-dessus. Pour des services d'aide à l'orientation de voyageurs, le système de Kim et al. [81] augmente des bâtiments 3D au-dessus d'une carte 2D affichée sur une table. Nilsson et al. [116] proposent un système de RA sur table (carte 2D papier + informations virtuelles) pour faciliter la planification de tâches en gestion de crises et la collaboration entre les acteurs des différents domaines d'intervention. Pour la gestion d'un feu de forêt par exemple, trois domaines d'intervention doivent collaborer : les services médicaux, la police et l'armée. Ce système de RA, présenté à la Figure 1.5 permet la collaboration de trois participants, un par domaine d'intervention. Tous sont placés autour d'une table afin de visualiser le plan 2D de la zone qui est en feu. Chacun est également équipé d'un casque de RA pour visualiser des informations virtuelles telles que l'emplacement des groupes médicaux, de police et militaires. Chacun peut déplacer ses troupes et aussi visualiser celles des autres. Ils peuvent aussi ajouter des points d'intérêt virtuels.

Ssin et al. présentent le système *GeoGate* [149] pour visualiser des données maritimes (Figure 1.6). Une carte 2D est affichée sur la table et l'utilisatrice place un support tangible en forme de cercle pour sélectionner une zone d'intérêt sur la carte. Un casque HoloLens est utilisé pour afficher les informations géo-temporelles corres-

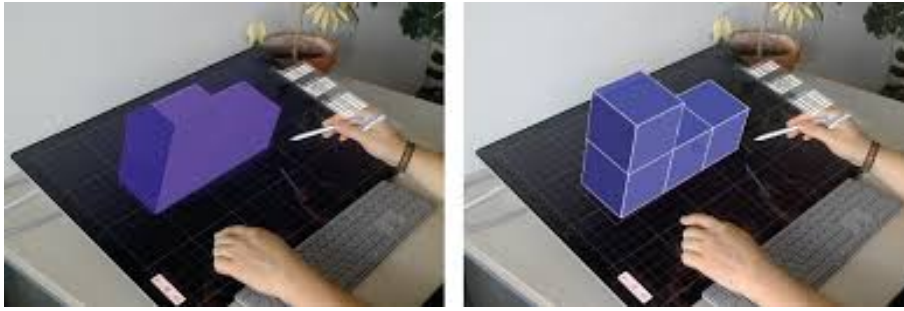


FIGURE 1.7 – *DesignAR* de Reipschläger et Dashselt [130]. Images extraites de [130].

pondant à la zone d'intérêt : ces informations sont représentées sous la forme d'un "cube espace-temps" [59] (récemment repris en réalité virtuelle sur table dans [54]).

Les systèmes utilisant un casque de RA permettent d'afficher un modèle en 3D, et non une projection 2D du modèle 3D comme avec les dispositifs portables (voir Section 1.1). Ils permettent également d'interagir avec les mains (qui ne portent pas de dispositif d'affichage). Ainsi, au-delà des cartes augmentées, ces systèmes peuvent aussi être utiles en modélisation 3D. Par exemple, Dedual et al [49] combinent une surface multi-touch pour l'interaction en entrée et un casque de RA pour l'affichage 3D. Un objet 3D virtuel affiché au-dessus de la table peut ainsi être manipulé (rotation, changement d'échelle, etc.) par des interactions au toucher sur la table.

Plus récemment, Reipschläger et al. présentent *DesignAR* [130, 131], un poste de travail hybride permettant l'interaction au toucher direct sur une surface 2D horizontale pour créer et modifier des objets 3D affichés dans le casque de RA. Par exemple, une utilisatrice peut dessiner avec un stylo les contours d'une forme sur la table (un carré) et un modèle 3D de son dessin (un cube) est alors augmenté au-dessus. L'utilisatrice peut également modifier une forme 3D à l'aide du stylo : comme le montre la Figure 1.7, l'utilisatrice passe à une vue perspective 2D du modèle (Figure 1.7 - gauche), interagit par interaction au toucher sur la table, puis retourne à une vue 3D du modèle pour voir le résultat en RA (Figure 1.7 - droite). Dans ces deux derniers systèmes, la surface horizontale (table, écran) est utilisée comme dispositif d'interaction en entrée permettant une interaction plus précise que l'interaction dans l'air.

1.2.2 Interactions

Pour l'exploration de cartes, Satriadi et al. [138] proposent par exemple des techniques d'interaction à main levée pour changer d'échelle ou se déplacer dans une carte 2D, comme illustré à la Figure 1.8.

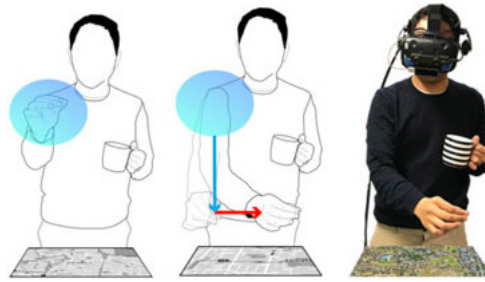
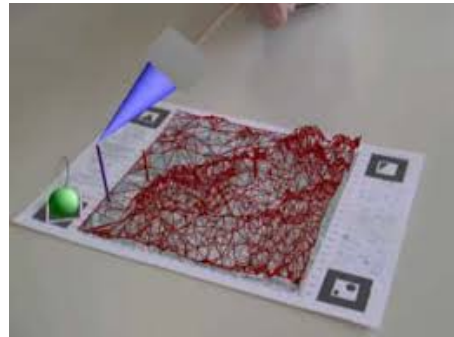


FIGURE 1.8 – Exemples d’interactions dans l’air proposées par [138] pour l’exploration de cartes. Un geste de pincement suivi d’un mouvement de la main dans l’air permet à la fois de changer d’échelle (flèche bleue) et de se déplacer dans la carte (flèche rouge). Images extraites de [138].



(a) Interaction par toucher directement sur l’objet d’intérêt dans l’air.



(b) Interaction par raycasting avec un bâton physique.

FIGURE 1.9 – Interactions 3D dans l’air avec un casque de RA pour le pointage d’un objet virtuel 3D. Images extraites de [81] (à gauche) et de [24] (à droite).

Pour pointer et sélectionner du contenu augmenté, en particulier des objets virtuels 3D, les interactions basées sur des gestes dans l’air comme illustrées à la Figure 1.9 sont couramment mises en œuvre [127]. Aussi désignée comme métaphore de la main virtuelle, l’utilisatrice touche et manipule directement les objets virtuels 3D. Cela peut être fait à main levée (en s’appuyant sur des gestes du doigt ou de la main [36, 81] comme à la Figure 1.9a) ou en tenant un objet physique comme dispositif de pointage (par exemple un stylo) [24, 60, 70]. Bobrich [24] présente un système de cartes augmentées qui permet d’interagir avec des objets 3D virtuels avec un petit bâton : l’utilisatrice tient un bâton physique prolongé par un pointeur virtuel (Figure 1.9b). Ce type d’interaction repose sur la métaphore du pointeur virtuel. L’interaction tactile sur la table définit aussi une autre modalité d’entrée pour les systèmes de RA sur table interactive mais reste principalement utilisée pour des tâches longues et complexes comme la modification et la manipulation d’objets 3D (voir les systèmes de modélisation 3D présentés dans la Section 1.2.1).

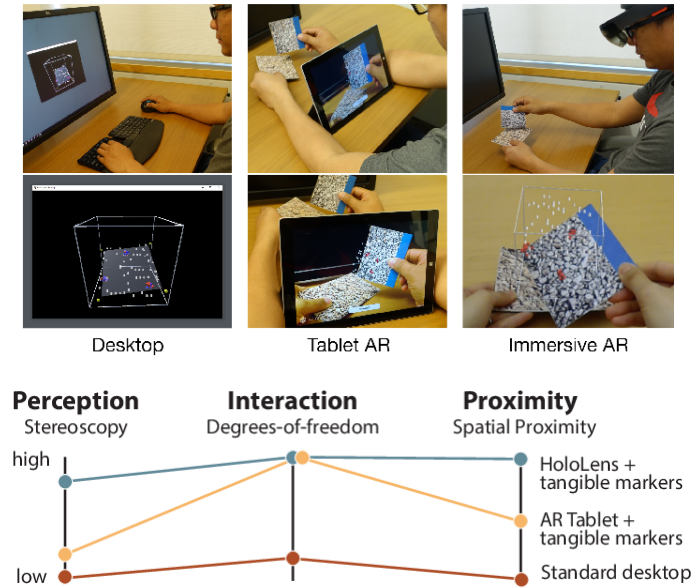


FIGURE 1.10 – Étude comparative de Bach et al. [6]. Images extraites de [6].

1.3 Études comparatives existantes

Comme nous l'avons vu précédemment, deux types de dispositifs sont utilisés pour la RA sur table : les téléphones/tablettes et les casques. Nous constatons que très peu d'études comparatives ont été menées pour comparer ces deux types de dispositifs. Nous relevons ainsi que deux études de la littérature qui comparent soit les deux types de dispositifs soit les métaphores d'interaction associées à ces deux dispositifs :

- La première étude [6] compare spécifiquement une technique d'interaction utilisant une tablette avec une technique d'interaction utilisant un casque de RA. L'étude n'est pas spécifique à la RA sur table mais s'applique à la RA en général.
- La seconde étude [95] oppose les métaphores de la main virtuelle et du pointeur virtuel à celle de la lentille magique dans le cas de la RA sur table. Les métaphores de la main virtuelle et du pointeur virtuel (Figure 1.9) sont souvent appliquées dans le cas d'un casque de RA (Section 1.2) tandis que la métaphore de la lentille magique est fréquemment utilisée avec un téléphone/tablette (Section 1.1). Il convient néanmoins de noter que dans cette étude les trois techniques utilisent un casque.

Dans une étude comparative, Bach et al. [6] ont considéré trois dispositifs : un poste de travail classique, et deux appareils de RA, une tablette et un casque *Holo-*

loLens 1 [110] pour l'exploration de nuages de points 3D (Figure 1.10). Contrairement au poste de travail où l'interaction repose classiquement sur le clavier et la souris, les deux techniques de RA proposent des interactions à 6 degrés de liberté où l'utilisatrice tient des marqueurs tangibles pour déplacer et faire pivoter les hologrammes. Ces environnements immersifs tangibles appelés TUIs (*tangible user interfaces*), offrent à l'utilisatrice des interactions naturelles et intuitives car elles sont proches des interactions que nous faisons dans le monde réel. Les auteurs classent ces trois techniques suivant trois aspects :

- Stéréoscopie
- Degrés de liberté
- Proximité spatiale entre l'espace d'interaction et les objets

Les trois techniques comparées lors de l'étude sont décrites selon ces trois aspects à la Figure 1.10. Nous constatons que les différences entre les deux systèmes de RA sont la perception de la scène (vue stéréoscopique pour le casque, vue monoscopique pour la tablette) et la proximité entre l'interaction (espace moteur) et la perception des objets (espace visuel). Avec la technique utilisant la tablette, l'utilisatrice interagit derrière l'écran alors que les données sont affichées sur l'écran. Au contraire, avec le casque l'utilisatrice interagit directement sur les hologrammes (les deux espaces moteur et visuel étant confondus).

Lors de l'expérimentation, les participants sont invités à effectuer des tâches d'interaction classiques en visualisation d'information :

- Une tâche d'estimation de la distance entre des paires de points (Quelle est la paire de points la moins espacée entre les rouges et les jaunes?).
- Une tâche d'estimation du nombre de clusters.
- Une tâche de placement de plan de coupe passant par trois points spécifiques.
- Une tâche de sélection.

Pour la tâche de sélection en RA (tablette et casque), l'utilisatrice tient dans une de ses mains un marqueur tangible permettant de modifier la position et l'orientation du nuage de points. L'utilisatrice doit d'abord placer ce marqueur sous le nuage de points : la visualisation est alors modifiée selon l'orientation et la position de ce marqueur (interaction simple et directe via 6 DOFs). Pour sélectionner un point dans le nuage : (1) avec la tablette, l'utilisatrice sélectionne l'objet par interaction au toucher direct sur l'écran de la tablette, (2) avec le casque, l'utilisatrice pointe directement l'objet d'intérêt avec un autre marqueur.

Les résultats montrent que l'interaction classique avec un poste de travail est

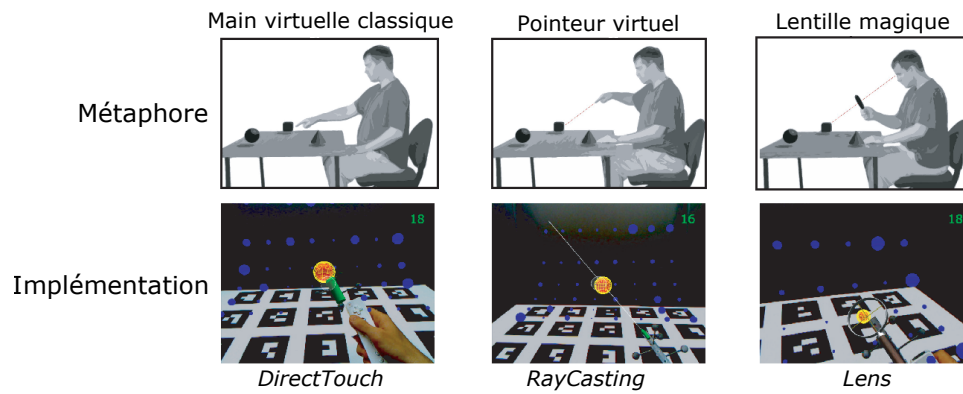


FIGURE 1.11 – Étude comparative de Looser et al. [95]. Images modifiées de [95].

performante pour toutes les tâches étudiées et est généralement plus performante que les deux autres techniques. La technique d'interaction avec la tablette est en moyenne la technique la moins précise et la plus lente de l'étude. L'interaction avec le casque est au moins aussi précise mais plus lente que l'interaction classique avec le poste de travail. Selon les auteurs, ceci pourrait être dû à la phase d'apprentissage importante observée pour la technique avec le casque.

De plus, la technique avec le casque offre de bonnes performances pour les tâches qui nécessitent une perception spatiale et des interactions avec un haut degré de liberté telles que les tâches de sélection et de plan de coupe. Pour la tâche de sélection, les utilisateurs sont plus précis en utilisant la technique avec le casque. Pour la tâche de plan de coupe, la technique avec le casque est plus rapide que les deux autres techniques (tablette et poste de travail). Ces bons résultats de la technique avec le casque peuvent être dus à sa manière intuitive d'interagir avec l'hologramme (rotation, déplacement des hologrammes). Les participants interagissent comme ils le feraient avec des objets physiques (6 DOF).

L'importance de l'aspect *proximité* pourrait expliquer les mauvais résultats de la technique avec la tablette, car l'affichage des objets est distant de la zone d'interaction. De plus, la technique avec la tablette implique une interaction 3D sur des objets perçus en 2D (projection 2D sur l'écran). Cette "incompatibilité" pourrait aussi complexifier les tâches réalisées avec cette technique.

Dans le contexte de la RA sur table, Looser et al. [95] comparent trois techniques de sélection d'objets virtuels 3D (Figure 1.11) : une technique de main virtuelle *DirectTouch*, une technique de pointeur virtuel *RayCasting* et une technique de lentille magique *Lens*. Les trois techniques de sélection sont basées sur un casque sans vue stéréoscopique. La lentille est attachée à une Wiimote tenue par l'utilisatrice. L'utili-

satrice regarde à travers cette lentille virtuelle et la place sur les objets à sélectionner. Cette technique implémente un lancer de rayon pour la sélection d'objets : une sélection est effectuée lorsque le rayon, partant des yeux de l'utilisatrice et passant à travers la lentille, entre en collision avec un objet virtuel. Cette interaction est proche d'une interaction avec tablette. Cependant, la technique *Lens* limite la sélection au centre de la lentille, contrairement à la tablette qui permet en général de sélectionner les objets présents n'importe où sur l'écran par interaction au toucher. Pour cette étude, les participants devaient sélectionner des sphères affichées (en environnements épars et denses) au-dessus d'une table. Les résultats montrent tout d'abord que *DirectTouch* est la technique la plus précise. Cependant, la technique de la lentille (*Lens*) est plus rapide et est préférée aux deux autres techniques. Les auteurs expliquent que, même si toutes les techniques utilisent un casque, les techniques *DirectTouch* et *RayCasting* ont été beaucoup plus impactées par le manque de vision stéréoscopique, biaisant l'étude expérimentale en avantageant la technique de la lentille magique. En effet, *Lens*, de la même manière que la tablette, réduit la tâche de sélection 3D à une tâche de sélection 2D uniquement. L'absence de stéréoscopie pour cette technique est donc moins impactante que pour les deux autres. Enfin, les auteurs n'ont pas trouvé d'impact de la densité de l'environnement sur les performances des trois techniques.

1.4 Bilan

Pour des tâches simples comme la sélection d'objets virtuels 3D, les techniques d'interaction pour la RA sur table diffèrent en termes de modalité de sortie (tablette/téléphone, casque) et de modalité d'entrée (interaction au toucher direct sur l'écran 2D, interaction en 3D dans l'air). En effet, pour les dispositifs portables (tablette/téléphone), les tâches sont généralement réalisées par une interaction tactile sur l'écran. Avec le casque, l'interaction dans l'air semble être la modalité d'entrée la plus couramment utilisée et la plus directe pour interagir avec des objets 3D car elle intègre complètement les 3 dimensions de l'espace 3D.

Comme nous l'avons souligné dans cet état de l'art, très peu d'études ont été menées afin de comparer ces dispositifs d'affichage pour la RA sur table. Compte tenu des dispositifs actuellement disponibles pour la RA, une question demeure : quel dispositif utiliser entre un dispositif portable offrant une interaction tactile et un casque offrant une interaction 3D dans l'air ? La section suivante (Section 2) est consacrée à cette question.

2. Contribution : tablette ou casque pour la réalité augmentée sur table

Dans la section précédente (Section 1), nous avons présenté des exemples de systèmes de réalité augmentée (RA) sur table tels que les systèmes d'information géographique [24, 172], les systèmes de planification urbaine [142, 162] ou les systèmes de conception d'intérieur [118]. Ces systèmes permettent une double représentation du plan en combinant une représentation 2D affichée sur la table avec une représentation 3D (par exemple des paysages, des bâtiments ou des pièces d'un appartement). La structure de la section précédente (Section 1) souligne néanmoins une différence notable entre ces systèmes : ceux utilisant un dispositif portable pour visualiser l'augmentation 3D et ceux reposant sur un casque de RA. L'affichage de l'augmentation 3D et l'interaction avec les objets virtuels 3D sont donc différents :

- *Les tablettes et les téléphones.* Ils offrent à l'utilisatrice une projection 2D sur l'écran du modèle virtuel 3D (une vue comme une “fenêtre sur le monde virtuel”). Ce type de RA implique généralement des interactions au toucher direct sur l'écran du dispositif portable.
- *Les casques de RA.* Ils permettent un affichage 3D du modèle virtuel 3D et impliquent généralement des interactions 3D dans l'air pour interagir avec les objets virtuels.

Ces deux dispositifs et techniques d'interaction associées sont couramment utilisés. La question est alors de savoir si un dispositif est à privilégier pour la RA sur table. Par exemple, pour des systèmes d'information géographique, quel dispositif faut-il utiliser pour visualiser et interagir avec les modèles de terrain : un téléphone [172] ou un casque [24] ? Quel dispositif faut-il considérer pour explorer et interagir avec une ville augmentée en 3D : une tablette comme dans *MetaDesk* [162] et *Hover Pad* [142] ou un casque comme dans [81] ? Pour répondre à ces questions nous avons mené une étude comparative que nous décrivons dans cette section. Les résultats de cette étude ont aussi permis d'orienter nos travaux suivants (Chapitres II et III).

Ainsi dans cette section, nous examinons les avantages et les limitations des dispositifs portables et des casques pour les tâches fondamentales de recherche et de sélection d'objets virtuels dans une scène 3D. Pour cela, nous avons implémenté

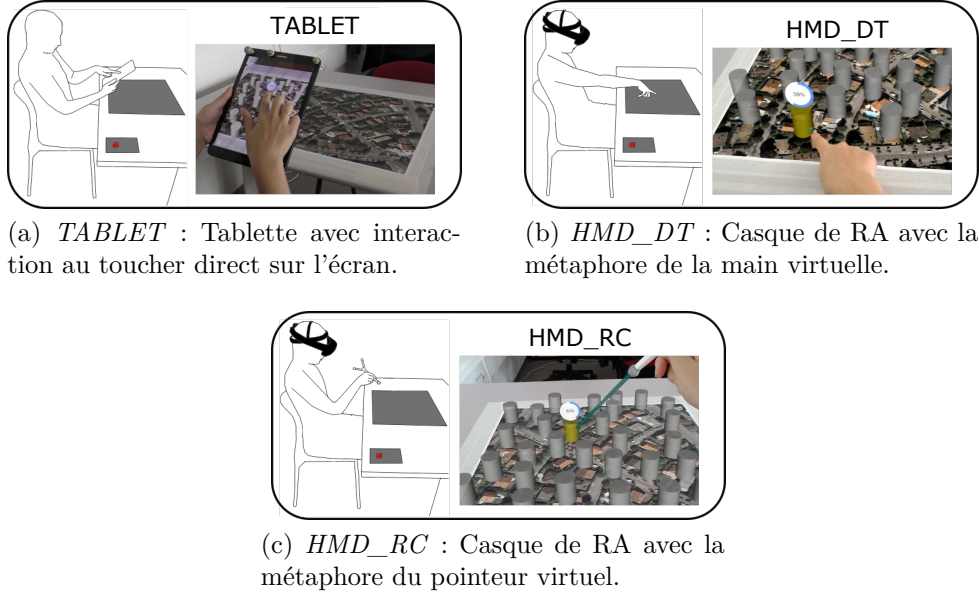


FIGURE 2.1 – Les trois techniques comparées.

trois techniques, présentées en Section 2.1. Les trois techniques implémentent les interactions couramment utilisées avec les deux types de dispositifs. Au sein du système expérimental de RA sur table présenté en Section 2.2, nous avons réalisé deux études expérimentales comparant ces trois techniques. La première compare ces techniques pour des tâches de recherche et de sélection d'objets visibles en environnements épars. Le protocole ainsi que les résultats obtenus sont présentés dans les Sections 2.3 à 2.5. La seconde étude expérimentale permet de comparer ces mêmes techniques dans des environnements denses où les objets cibles sont occultés. Le protocole et les résultats obtenus sont détaillés dans les Sections 2.6 à 2.8.

2.1 Éléments de conception : trois techniques étudiées

Les trois techniques de sélection étudiées, l'une avec une tablette et les deux autres avec un casque, sont présentées à la Figure 2.1. Pour la condition avec tablette, la sélection se fait par interaction au toucher direct sur l'écran. Cette technique offre une interaction à distance de l'objet à sélectionner car elle implémente la métaphore du raycasting : un rayon, partant du point de collision entre le doigt de l'utilisatrice et l'écran, est envoyé dans la scène 3D virtuelle. Bien que l'objet soit distant, l'utilisatrice interagit directement avec la projection 2D de l'objet distant sur l'écran. L'utilisation de ce proxy affiché sur l'écran permet de réduire la distance

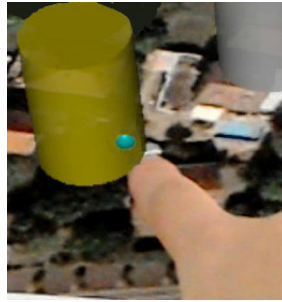


FIGURE 2.2 – Sélection d’un cylindre avec HMD_DT : la sphère bleue qui suit le doigt entre en collision avec l’objet, un cylindre.

entre l’objet et l’utilisatrice.

Pour la condition avec un casque, les deux techniques considérées correspondent aux deux métaphores d’interaction 3D dans l’air [5] : la main virtuelle (“grasping metaphor” [89]) et le pointeur virtuel (“pointing metaphor” [89]). La main virtuelle standard implémente un mapping 1-1 entre la main réelle de l’utilisatrice et la main virtuelle. Avec une telle technique l’utilisatrice pointe donc directement l’objet qu’elle veut sélectionner. Ceci définit une manière intuitive de sélectionner un objet car elle simule la façon dont nous manipulons les objets dans le monde réel [31]. Cette technique est généralement utilisée pour la sélection d’objets proches. Pour des objets distants, la technique la plus couramment utilisée est le pointeur virtuel standard, également appelé *raycasting*. L’utilisatrice contrôle un rayon provenant de sa main ou de l’objet physique tenu et dont l’orientation est définie par l’orientation de la main ou de l’objet tenu.

Les trois techniques sont conçues pour permettre à l’utilisatrice d’interagir avec une scène 3D virtuelle projetée sur une table physique. L’utilisatrice peut rechercher et sélectionner un objet cible dans un ensemble d’objets virtuels 3D placés sur la table. Les objets 3D virtuels considérés dans l’étude sont abstraits : ce sont des cylindres (Figure 2.1). Dans ce qui suit, nous détaillons les trois techniques d’interaction :

- **TABLET.** L’utilisatrice tient une tablette pour visualiser la scène 3D virtuelle au-dessus de la table et doit toucher la projection 2D d’un objet sur l’écran pour le sélectionner (Figure 2.1a).
- **HMD_DT.** L’utilisatrice porte un casque de RA *HoloLens 1* pour visualiser la scène 3D virtuelle. Pour sélectionner un objet virtuel, elle le “touche” directement avec son index (Figure 2.1b). Comme illustré à la Figure 2.2, une sphère bleue virtuelle suit le bout du doigt de l’utilisatrice. La sélection se

produit lorsque cette sphère entre en collision avec un objet.

- **HMD_RC.** L'utilisatrice porte également un casque de RA *HoloLens 1* pour visualiser la scène. Pour la sélection, elle tient un bâton physique dans sa main dominante. En utilisant la métaphore du raycasting, un rayon virtuel est affiché comme une extension du bâton (Figure 2.1c). La sélection se produit lorsque le rayon entre en collision avec un objet. La largeur du rayon est égale au diamètre de la sphère au bout du doigt dans HMD_DT.

Nous avons choisi d'utiliser un bâton plutôt que le doigt seul car l'interaction semblait moins ambiguë avec un objet tangible dans la main. En effet les techniques de pointage par raycasting avec le doigt peuvent être classées en deux groupes [4] : les techniques *eye-rooted* où le rayon part des yeux et passe par le bout du doigt, et les techniques *hand-rooted* où le rayon prolonge le doigt de l'utilisatrice. Cependant avec le bâton, comme dans l'étude comparative de Looser et al. [95] (Figure 1.11), la solution est sans ambiguïté, le rayon prolongeant l'objet physique.

Étude pilote – L'importance d'un retour visuel virtuel

Une étude préliminaire avec 6 participants a été réalisée afin de tester l'utilisabilité des trois techniques. Les résultats de cette étude ont tout d'abord permis de valider l'implémentation des techniques TABLET et HMD_RC. De plus pour cette étude pilote, la technique HMD_DT conçue initialement était légèrement différente. En effet, dans la première version de HMD_DT testée dans l'étude pilote, la sphère bleue qui suit le doigt de l'utilisatrice et qui sert de "collisionneur" n'était pas affichée (sphère invisible). Les retours des participants ont montré d'importants problèmes de perception de la profondeur et une sous-estimation de la distance de la cible avec la première version de HMD_DT. En effet, en RA, il est souvent difficile d'appréhender les distances et les positions d'objets physiques par rapport à des objets virtuels. Ce type de problèmes a été identifié dans plusieurs études [85, 122, 157]. Dans le cas de la technique HMD_DT, nous supposons plusieurs causes extraites de [85] : la distorsion de la perception de profondeur, le manque d'occultation et la latence du système.

Le problème de distorsion de la profondeur et le fait qu'un objet virtuel n'est jamais vraiment occulté par un objet physique sont les causes principales. En effet, avec la technique HMD_DT, l'utilisatrice doit placer son doigt dans l'objet virtuel pour le sélectionner. Cependant, lorsque le doigt arrive devant l'objet virtuel, le doigt ne l'occulte pas car l'objet est affiché au niveau des yeux de l'utilisatrice et non au-



FIGURE 2.3 – Sélection d'un cylindre avec HMD_DT : la sphère bleue qui suit le doigt entre en collision avec un cylindre. Une barre de progression est affichée pour indiquer le temps restant avant la fin de la validation du cylindre.

dessus de la table. Ainsi, même si l'objet semble être au milieu de la table, sa main est toujours derrière l'affichage, et l'objet n'est donc jamais occulté. Pour résoudre ce problème, la technique HMD_DT a été modifiée pour afficher une sphère virtuelle sur le bout du doigt de l'utilisatrice (Figure 2.3). Par conséquent, la sphère et la cible étant toutes deux virtuelles, l'utilisatrice peut clairement observer la collision et l'occultation de la sphère lorsqu'elle entre dans la cible.

Le fait d'afficher la sphère virtuelle au bout du doigt atténue sans le supprimer le problème de la latence du système de RA. En effet, même si le doigt de l'utilisatrice est suivi précisément dans l'espace (suivi à 120 Hz par un système OptiTrack), la sphère qui suit son doigt se déplace avec une légère latence (induite par l'envoi de la position du doigt à l'HoloLens et à la mise à jour de la position de la sphère). Néanmoins cette latence est directement observable par le retour visuel sous la forme d'une sphère. De plus au moment où la sphère entre dans un objet, nous affichons une barre de progression au-dessus de cet objet afin de notifier à l'utilisatrice la détection d'une collision et d'indiquer le temps restant pour valider la sélection (Figure 2.3). Le retour visuel sous la forme sphère et d'une barre de progression rend ainsi observable l'état courant du système au cours d'une sélection. Ainsi, la latence n'est pas réduite mais l'état courant du système est observable à tout moment par l'utilisatrice.

La sphère suivant le bout du doigt a donc été affichée pour les deux études expérimentales présentées dans les sections suivantes. Contrairement à cette étude pilote, aucun participant n'a indiqué de problème de perception de la profondeur durant ces deux études.

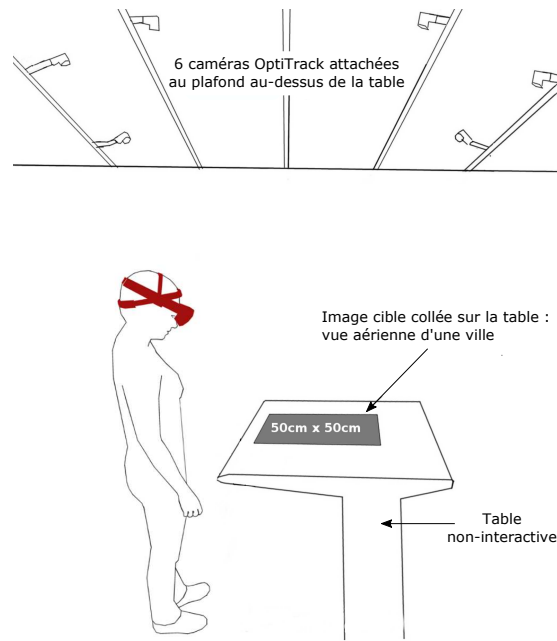


FIGURE 2.4 – Système expérimental générique.

2.2 Implémentation d'un système de RA sur table avec un casque

Dans cette section, nous présentons le système expérimental de RA sur table avec un casque et son implémentation. Le but de ce système est de permettre des interactions 3D dans l'air pour réaliser des pointages précis d'objets virtuels 3D ancrés à une table. Cette implémentation est commune à toutes nos études expérimentales (Chapitres I, II et III). Les différents éléments composant le système (caméras, table, image cible) sont présentés à la Figure 2.4.

2.2.1 Scène 3D virtuelle

Pour afficher la scène 3D virtuelle nous utilisons un casque de RA Microsoft HoloLens 1 [110]. L'appareil HoloLens est composé de deux écrans transparents haute définition 1278×720 . Son champ de vision est de $30 \times 17,5$ degrés, sa fréquence d'images est de 60 images par seconde et il pèse environ 579 g. La scène composée des objets à sélectionner est créée par le moteur de jeu vidéo *Unity*¹. La gestion des collisions entre objets pour la sélection avec le casque (collision entre la cible et soit le rayon, soit la sphère qui suit le doigt par exemple) est également gérée par Unity.

1. <https://unity.com/>

2.2.2 Suivi précis des mouvements de l'utilisatrice

Les techniques avec le casque impliquent des gestes dans l'air pour interagir avec les objets virtuels. Avec l'HoloLens 1, il est possible de détecter certains gestes de la main mais le suivi est généralement peu précis et relativement lent. Pour un suivi précis, nous avons opté pour un système *OptiTrack*² combinant des marqueurs réfléchissants et des caméras infrarouges capables de suivre ces marqueurs dans l'espace. Le suivi d'un objet physique est donc réalisé en fixant simplement un ou plusieurs marqueurs sur l'objet d'intérêt.

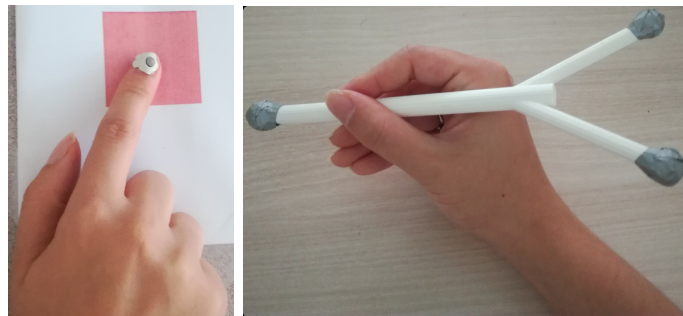


FIGURE 2.5 – Placement des marqueurs réfléchissants.

Pour une technique de sélection au toucher sur l'objet virtuel (Chapitres I et III), un marqueur est fixé sur l'index de l'utilisatrice afin de connaître à tout instant la position du bout du doigt. La Figure 2.5-gauche montre le placement de ce marqueur.

Pour une technique de sélection par raycasting (Chapitres I et II), la position du bâton (utilisé comme pointeur) mais également son orientation doivent être connues. Pour cela, un seul marqueur n'est pas suffisant. Nous utilisons alors trois marqueurs comme le montre la Figure 2.5-droite. Le bâton a été imprimé avec une imprimante 3D, il est très léger (environ 10 g) et mesure 21 cm de long.

Le suivi des marqueurs est fait par 6 caméras USB *Flex 13* (Figure 2.6), toutes attachées au plafond afin d'avoir une vue globale de l'environnement tout en limitant l'occultation des marqueurs. Ces caméras, dont la résolution est de 1280×1024 , permettent un suivi des marqueurs à 120 images par seconde. Le logiciel *Motive* collecte et recoupe les données envoyées par chacune des caméras afin de déduire la position 3D de chacun des marqueurs.

2. <https://optitrack.com/>

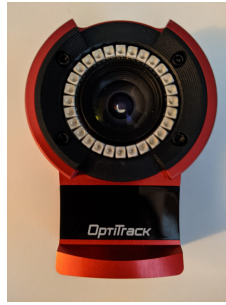


FIGURE 2.6 – Caméra OptiTrack type Flex 13.

2.2.3 Lien entre le monde virtuel et le monde physique

L’objectif est que l’utilisatrice perçoive la scène virtuelle comme si elle était posée sur une table physique. De plus l’utilisatrice doit pouvoir interagir avec cette scène, comme sélectionner un objet en plaçant le doigt à l’intérieur de celui-ci. Il convient donc de créer un lien entre la scène virtuelle (les objets à sélectionner) et le monde physique (la table, le doigt de l’utilisatrice, le bâton) pour que les objets soient tous représentés dans le même espace. Une fois ce lien créé, la correspondance entre le monde virtuel et le monde physique permet par exemple la sélection d’un objet en plaçant son doigt dedans et aussi l’extension du bâton physique par un rayon virtuel.

Pour créer ce lien, nous utilisons tout d’abord la plateforme *Vuforia*, couramment utilisée pour le développement d’applications en RA. Le kit de développement (SDK) *Vuforia* utilisé dans Unity offre des fonctionnalités de vision par ordinateur comme la reconnaissance et le suivi d’images physiques par les dispositifs de RA tels que les tablettes et les casques. Ces images physiques sont appelées *images cibles* dans *Vuforia*. Ainsi, grâce à *Vuforia* il est possible d’afficher des objets virtuels au-dessus d’une image cible³, même si l’image est en mouvement. Tant que cette image est visible par la caméra du dispositif de RA et est détectée par *Vuforia*, les objets virtuels resteront affichés au-dessus. Nous plaçons donc sur la table une image cible afin d’afficher la scène virtuelle directement au-dessus. Comme le montre la Figure 2.7, l’image collée sur la table et suivie par *Vuforia* représente une vue satellite imprimée d’une ville. Cette image est schématisée par une zone grise sur la table à la Figure 2.4.

Les performances de l’HoloLens 1 ne permettent cependant pas d’utiliser *Vuforia* en continu. Pour optimiser les performances, nous utilisons alors le concept “d’ancrage spatiale” (*world anchor* dans la documentation) proposé dans la boîte à outils MR-Toolkit pour Unity, développée par Microsoft. Cette ancre représente un point phy-

3. <https://library.vuforia.com/features/images/image-targets.html>



FIGURE 2.7 – Image cible détectée et suivie par le SDK Vuforia.

sique dans l'espace d'une pièce et est utilisée en interne par l'HoloLens pour la sauvegarde persistante des positions des hologrammes ou des menus dans une pièce. Les ancres spatiales sont importantes au sein de la plateforme HoloLens : elles sont suivies au cours du temps voire ajustées si besoin pour maintenir leurs positions initiales, quels que soient les mouvements de l'utilisatrice. Ainsi, pour éviter d'utiliser Vuforia tout au long des études expérimentales, une ancre spatiale est créée et placée dans le coin en bas à gauche de l'image cible dès que celle-ci est détectée, puis Vuforia est arrêté. Le suivi de l'image n'est donc plus nécessaire car la scène virtuelle est placée en fonction de cette ancre spatiale : l'origine (position $(0, 0, 0)$) du monde virtuel est la position de l'ancre et par extension le coin en bas à gauche de l'image cible. Enfin, nous initialisons le système OptiTrack pour que l'origine des marqueurs soit également le coin de l'image (en plaçant l'équerre de calibrage à cet endroit lors du calibrage des caméras). Une fois ces étapes réalisées, le lien entre le monde virtuel et le monde physique est créé et tous les objets, qu'ils soient physiques ou virtuels, sont dans le même espace (l'espace dont l'origine est le coin inférieur gauche de l'image). La position de chacun des marqueurs peut alors être envoyée au casque via des sockets. Dans le cas du raycasting, et connaissant la position des marqueurs du bâton dans l'espace (obtenue par les caméras), il est maintenant simple d'étendre ce bâton avec un rayon virtuel. Le rayon est alors utilisé comme collisionneur pour sélectionner des objets virtuels.

2.3 Étude expérimentale 1 : recherche et sélection d'objets visibles en environnements éparés

2.3.1 Objectif de l'étude et hypothèses

Basé sur le système expérimental ci-dessus, cette étude expérimentale a pour objectif de comparer les trois techniques d'interaction présentées en Section 2.1 pour des tâches de recherche et de sélection. L'étude est conçue pour distinguer deux phases dans la tâche : la phase de recherche et la phase de sélection. En effet, la perception de la scène et la manière de la parcourir diffèrent selon que l'utilisatrice se sert d'une tablette (TABLET) ou porte un casque (HMD_DT et HMD_RC). De plus, par rapport à la technique HMD_DT impliquant un toucher direct de l'objet virtuel 3D, les techniques HMD_RC et TABLET basées sur un rayon permettent à une utilisatrice d'interagir à distance avec les objets virtuels placés sur la table : leurs performances pour la sélection des cibles distantes devraient donc être différentes de celles de HMD_DT. En conséquence, le premier objectif de cette étude est de comparer les trois techniques pour chacune des deux phases en plaçant des cibles plus ou moins loin de l'utilisatrice. De plus, comme notre étude présente des similitudes avec celle de Looser et al. [95] (voir Section 1.3 du Chapitre I), le deuxième objectif est de comparer nos résultats avec les leurs. Pour cette étude comparative, nous formulons trois hypothèses :

- *H1* : Les techniques avec le casque sont plus rapides que celle basée sur la tablette pour trouver des cibles virtuelles. Le casque ne nécessite que des mouvements de tête pour rechercher une cible, contrairement à la tablette qui implique aussi des mouvements des bras.
- *H2* : Pour des cibles éloignées, les techniques permettant d'interagir à distance (HMD_RC et TABLET) sont plus rapides et moins fatigantes que la technique HMD_DT d'interaction au toucher direct sur la cible.
- *H3* : La technique basée sur la tablette est préférée par les participants et est plus rapide que celles avec le casque pour sélectionner des cibles virtuelles. Cette hypothèse est formulée sur la base des résultats obtenus par la technique *Lens* par rapport aux techniques *DirectTouch* et *RayCasting* dans l'étude de Looser et al. [95].



FIGURE 2.8 – Image cible suivie par la tablette grâce au SDK Vuforia.

2.3.2 Implémentation

L'implémentation des deux techniques utilisant le casque (HMD_DT et HMD_RC) est décrite dans la Section 2.2. Nous détaillons ici l'implémentation de l'interaction au toucher avec la tablette. Pour la technique TABLET, nous utilisons une tablette Samsung Galaxy Tab S3. Son écran de 9,7" a une résolution de 2048×1536 . La tablette pèse 429 g et est une des plus légères du marché au moment de l'étude. Comme pour les techniques avec le casque, la scène virtuelle ainsi que les interactions au toucher direct sur l'écran sont gérées par Unity. L'image cible présente sur la table est détectée et suivie par Vuforia et permet à l'utilisatrice d'observer les objets à sélectionner comme s'ils étaient placés sur la table (Figure 2.8). Vuforia étant spécifiquement développé et optimisé pour les appareils portables tels que les téléphones et les tablettes (contrairement au casque de RA), la technique TABLET repose donc seulement sur Vuforia et n'utilise pas d'ancrage spatiale (concept développé par Microsoft pour l'HoloLens).

2.3.3 Système expérimental

Le système expérimental de cette étude, présenté à la Figure 2.9, est basé sur le système de RA sur table décrit en Section 2.2. Les participants sont assis devant la table physique (qui est non-interactive). Comme expliqué en Section 2.2, l'image suivie par Vuforia (zone grise sur la table à la Figure 2.9) désigne la zone où sont placés tous les objets virtuels. Un carré rouge de $4 \text{ cm} \times 4 \text{ cm}$ est collé à droite de l'image, à une distance de 25 cm. Il représente la position initiale des tâches de recherche et de pointage. La position précise de la chaise est indiquée par un marquage au sol. Cette position de la chaise permet à une utilisatrice d'atteindre avec sa main n'importe quelles cibles placées sur l'image. Cependant, elle doit se pencher en avant pour atteindre celles au bout de l'image (les cibles les plus distantes).

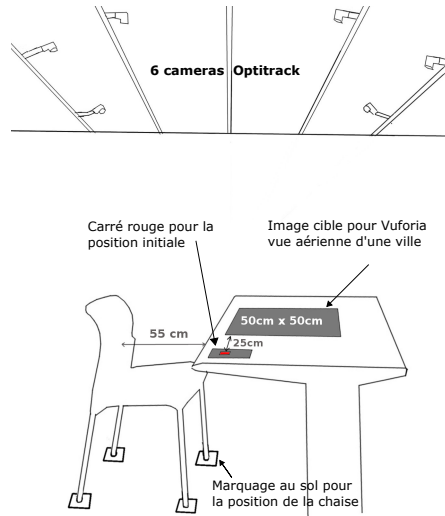


FIGURE 2.9 – Système expérimental.

2.3.4 Tâche

Sachant que la taille de la cible a un impact sur les performances de la tâche de pointage [55, 98], nous avons décidé d'utiliser des cylindres comme objets à sélectionner : leur largeur étant toujours la même quelles que soient les lignes d'approche [61]. La scène est composée de 30 cylindres 3D (1 cylindre cible rouge et 29 cylindres distracteurs gris) comme illustré à la Figure 2.10a. Les cylindres mesurent 5 cm de hauteur avec un diamètre de 3,5 cm. La distance ayant également un impact sur les performances avec des mouvements différents selon les techniques pour les tâches de recherche et de sélection, la scène est divisée en cinq zones (invisibles) représentées à la Figure 2.10b. Les cylindres sont répartis de manière égale entre les zones. Le placement est aléatoire dans la zone, mais est fait pour éviter que les cylindres ne se chevauchent.



(a) Exemple d'une scène virtuelle.



(b) Découpage de la scène en 5 zones (invisibles).

FIGURE 2.10 – Scène virtuelle pour la première étude.

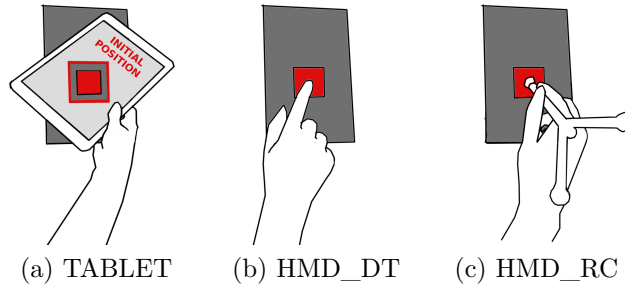


FIGURE 2.11 – Position initiale des 3 techniques.

La tâche d'exécution étant séparée en 2 phases, les participants doivent revenir à une position initiale, illustrée à la Figure 2.11, avant chaque phase. Cette position consiste à regarder et à placer le doigt, le bout du bâton ou le centre de la tablette sur le carré rouge placé sur la table à droite de la scène (voir le système expérimental à la Figure 2.9). Lors de la première phase, les participants doivent trouver la cible le plus rapidement possible et dire "OK" pour indiquer qu'ils l'ont trouvée. Environ 90% de la scène est visible si l'utilisatrice se place au fond de la chaise (pour TABLET, estimation calculée avec la tablette placée à 25 cm du corps). Les participants se replacent à nouveau en position initiale et doivent ensuite sélectionner la cible.

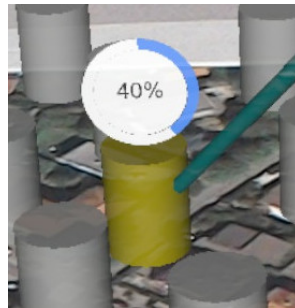


FIGURE 2.12 – Sélection d'un cylindre avec HMD_RC : le rayon entre en collision avec un cylindre. Le cylindre devient alors jaune et une barre de progression est affichée au-dessus pour indiquer le temps restant avant la fin de la validation.

Comme le montre la Figure 2.12, la couleur de la cible change pour indiquer une sélection en cours (le cylindre passe du gris au jaune). La sélection est validée si l'utilisatrice maintient cette sélection pendant 1 seconde. Durant cette seconde, une barre de progression est affichée au-dessus de la cible, comme illustré à la Figure 2.12. Cette barre est réinitialisée si l'utilisatrice perd la cible avant la fin de la validation.

2.3.5 Protocole

L'étude est conçue comme une étude expérimentale intra-sujet avec deux variables indépendantes : la technique d'interaction utilisée (TECH) et la zone dans laquelle se trouve la cible (ZONE). Pour cette étude, un ensemble de 15 scènes 3D ont été générées avant l'étude permettant trois tâches de recherche et de sélection à l'intérieur de chacune des 5 zones. Elles sont partagées entre tous les participants et toutes les techniques. L'ordre des techniques est contrebalancé par un carré latin tandis que les scènes comprenant des cibles dans les différentes zones sont présentées dans un ordre aléatoire pour chacune des techniques.

Nous commençons par une brève introduction de l'étude. Les participants doivent ensuite effectuer des phases de recherche et de sélection sur les 15 terrains plus 3 terrains d'entraînement pour chacune des 3 techniques d'interaction. À la fin de chaque technique, les participants remplissent un questionnaire. L'étude se termine avec un questionnaire final et un entretien. L'étude a duré environ 40 minutes par participant.

Nous avons recruté 12 volontaires droitiers au sein de notre laboratoire, âgés de 24 à 40 ans ($moyenne = 31,4 \pm 5,78$). Afin de minimiser l'effet «WoW», nous avons sélectionné des participants ayant déjà testé la RA, en particulier avec un casque, mais aucun d'entre eux n'était expert ni n'a participé à l'étude pilote (étude présentée à la Section 2.1).

Au regard des trois hypothèses à tester, nous considérons trois métriques objectives (temps de réalisation, précision lors de la validation, distance moyenne des mouvements lors d'une tâche de sélection) et deux métriques subjectives (fatigue physique, préférence des utilisateurs). Le temps de réalisation de la phase de recherche est le temps passé entre le départ de la position initiale et le moment où l'utilisatrice dit "OK". Le temps de sélection est le temps entre le départ de la position initiale et la première collision avec la cible. Pour évaluer la précision de la validation de sélection, nous comptons le nombre de collisions avant chaque validation. La distance moyenne des mouvements effectués par l'utilisatrice est calculée à partir des données capturées par le système OptiTrack (sauvegarde de la position de tous les marqueurs pour toutes les tâches).

Pour les mesures subjectives, nous avons conçu un questionnaire pour évaluer chaque technique et un questionnaire final pour comparer les techniques. Le questionnaire à propos de chaque technique (Annexe A) contient :

- Les questions posées dans l'étude de Looser et al. [95] afin de pouvoir directement comparer nos résultats avec les leurs. Les questions sont indiquées dans

le Tableau 2.2.

- Des questions sur le niveau de fatigue des participants pour plusieurs zones du corps [144] : les yeux, le visage, le cou/les épaules, les bras, les doigts, et le dos. Les participants évaluent leur niveau de fatigue sur une échelle de Likert en 4 points (1=“Aucune douleur”, 4=“Douleur sévère”).

Le questionnaire final (Annexe B) permet de classer les techniques, de la moins physiquement fatigante à la plus physiquement fatigante et par ordre de préférence.

2.4 Étude expérimentale 1 : résultats

Nous présentons dans cette section les résultats quantitatifs et qualitatifs de l'étude. La normalité des données est testée à l'aide du test Shapiro-Wilk. Pour les données normalement distribuées (mesures de distance des mouvements), nous utilisons un test statistique par analyse de variance (ANOVA) et des tests de Student (ou tests t) avec ajustement de Bonferroni pour les comparaisons par paires. Pour les données qui ne suivent pas une loi normale (mesures des temps de réalisation et de la précision de validation), nous utilisons un test de Friedman χ^2 non paramétrique et un test des rangs signés de Wilcoxon avec ajustement de Bonferroni pour les comparaisons par paires. Pour les analyses non-paramétriques à plusieurs variables indépendantes, nous appliquons une transformation de rang aligné (*Aligned ranks transformation-ART* [175]). Dans les graphiques suivants, nous représentons les valeurs moyennes (m) et les intervalles de confiance (IC) à 95% calculés avec la méthode bootstrap Bca (la fonction du bootstrap étant la moyenne) [51].

2.4.1 Résultats quantitatifs

Temps de réalisation moyens – Phase de recherche

Nous observons un effet statistiquement significatif de la technique TECH [$\chi^2 = 18$, $df = 2$, $p = 0,0001$] et de la zone ZONE [$\chi^2 = 17$, $df = 4$, $p = 0,002$] sur le temps de réalisation de la tâche de recherche. Les comparaisons par paires montrent qu'il est significativement plus lent ($p = 0,0015$) de rechercher une cible avec TABLET ($m = 3049$ ms, $IC = [2825, 3167]$) qu'avec HMD_RC ($m = 2186$ ms, $IC = [2053, 2347]$) ou HMD_DT ($m = 2111$ ms, $IC = [1901, 2274]$).

Effet d'interaction ZONE \times TECH. Nous observons un léger effet d'interaction entre ZONE et TECH [$F_{8,154} = 1,54$, $p = 0,065$] sur le temps de réalisation de la tâche

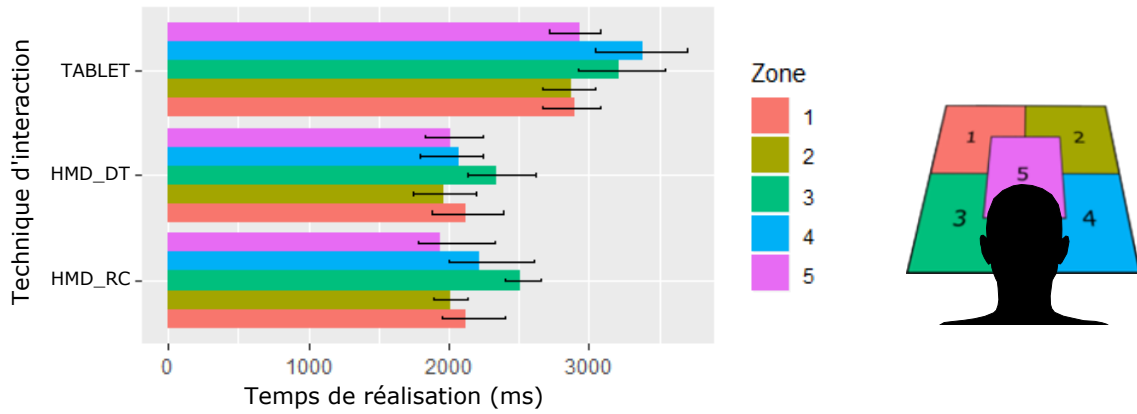


FIGURE 2.13 – Phase de recherche – Effet d’interaction entre la zone et la technique sur le temps moyen de réalisation de la tâche.

de recherche. Les résultats sont présentés à la Figure 2.13. Les comparaisons par paires montrent qu’il est significativement plus lent (pour les 5 zones, $p < 0,007$) de trouver la cible avec TABLET plutôt qu’avec HMD_DT ou HMD_RC.

Pour les deux techniques avec le casque, les participants passent plus de temps pour trouver une cible en zone 3 que dans les autres zones (Figure 2.13). Pour HMD_DT, les comparaisons par paires montrent une différence statistiquement significative entre les zones 3 et 5 ($p = 0,02$) et également une légère différence entre les zones 2 et 3 ($p = 0,06$). Pour HMD_RC, les comparaisons par paires montrent une différence statistiquement significative entre les zones 2 et 3 ($p = 0,01$) et une légère différence entre les zones 3 et 5 ($p = 0,06$).

Pour la technique TABLET, nous n’avons pas observé de différence entre les zones pour la phase de recherche.

Temps de réalisation moyens – Phase de sélection

Nous observons un effet statistiquement significatif de la technique TECH [$\chi^2 = 21$, $df = 2$, $p < 0,0001$] et de ZONE [$\chi^2 = 21$, $df = 4$, $p = 0,0003$] sur le temps de réalisation de la phase de sélection. Les comparaisons par paires montrent que les participants mettent significativement plus de temps à sélectionner une cible avec la technique TABLET ($m = 2312$ ms, $IC = [2144, 2472]$) qu’avec les techniques HMD_RC ($m = 1294$ ms, $IC = [1170, 1423]$, $p = 0,002$) et HMD_DT ($m = 1592$ ms, $IC = [1345, 1846]$, $p = 0,002$). Aucune différence statistiquement significative n’est observée entre les deux techniques utilisant le casque.

Effet d’interaction ZONE \times TECH. Nous observons un léger effet d’interaction entre ZONE et TECH [$F_{8,154} = 1,75$, $p = 0,06$] sur le temps de réalisation de la tâche de

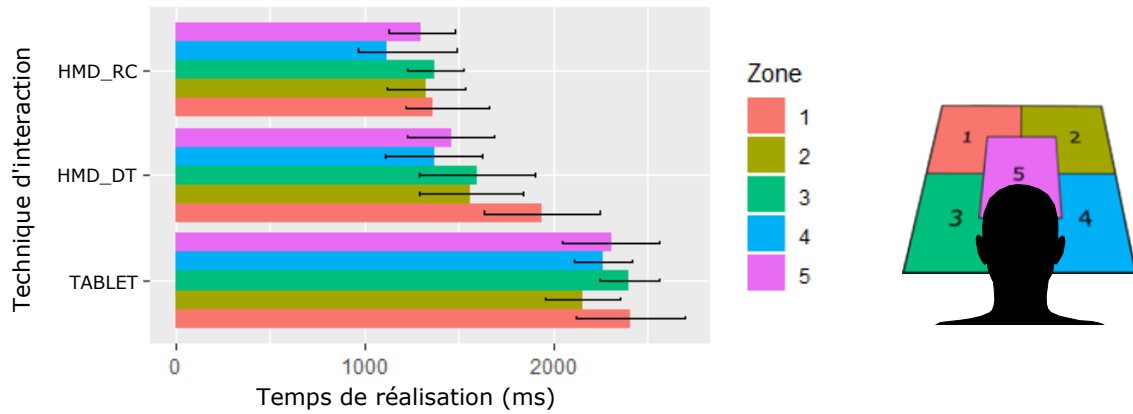


FIGURE 2.14 – Phase de sélection – Effet d’interaction entre la zone et la technique sur le temps moyen de réalisation de la tâche.

sélection. Les résultats sont présentés à la Figure 2.14. Pour toutes les zones, les comparaisons par paires montrent que HMD_RC est significativement plus rapide que TABLET (Pour les zones 1, 2, 3, et 5 : $p = 0,002$; pour la zone 4 : $p = 0,003$). Pour toutes les zones, les comparaisons par paires montrent que HMD_DT est aussi significativement plus rapide que TABLET ($p = 0,002$). Dans la Figure 2.14 nous observons cependant un chevauchement de 22% entre les intervalles de confiance de HMD_DT et TABLET pour la zone 1. La différence des temps de réalisation de la phase de sélection entre ces deux techniques est donc moindre pour la zone 1 (zone la plus éloignée de la position initiale). Enfin, nous n’avons trouvé aucune différence statistiquement significative entre les deux techniques avec le casque sauf pour les cibles de la zone 1. Pour celles-ci, HMD_DT est significativement plus lent que HMD_RC ($p = 0,015$).

Pour HMD_DT uniquement, nous observons un effet statistiquement significatif de la zone sur les temps de réalisation de la phase de sélection. Les comparaisons par paires montrent qu’il est plus long de sélectionner une cible avec HMD_DT dans la zone 1 que dans les zones 4 ($p = 0,005$) et 5 ($p = 0,013$). Nous n’avons pas observé de différence significative entre les zones pour les autres techniques.

Précision de la tâche de sélection

Le nombre d’erreurs est le nombre de collisions détectées avant la validation de la sélection de la cible. La dernière collision avec la cible avant la validation n’est pas comptée. En moyenne, les participants sont précis avec toutes les techniques avec moins d’une erreur avant validation (TABLET : $m = 0,3$ erreur, $IC = [0,16; 0,45]$; HMD_RC : $m = 0,87$ erreur, $IC = [0,57; 1,17]$; et HMD_DT : $m = 0,6$ erreur, $IC = [0,3; 1,08]$). Nous n’observons pas de différence statistiquement significative

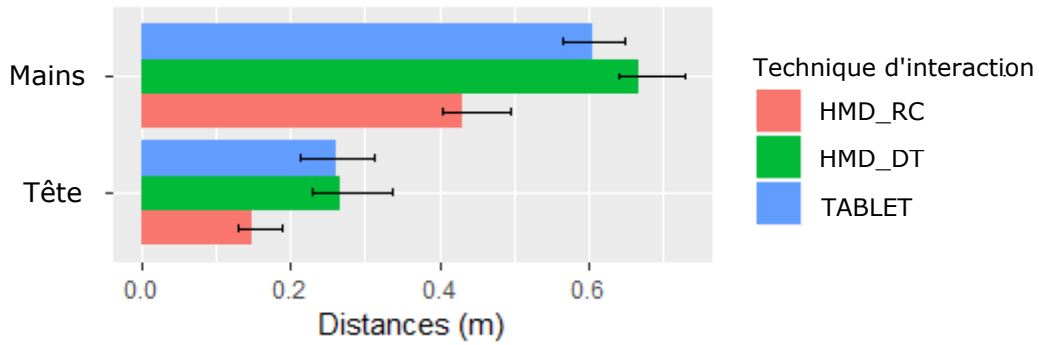


FIGURE 2.15 – Distances moyennes parcourues par les mains et la tête des participants.

entre les techniques.

Distances moyennes des mouvements lors de la phase de sélection

En plaçant des marqueurs réfléchissants additionnels au niveau de la tête des participants et sur la tablette (en plus des marqueurs nécessaires pour les techniques HMD_DT et HMD_RC) nous avons calculé les distances moyennes parcourues par les mains et la tête des participants. Les résultats obtenus sont présentés à la Figure 2.15. Nous observons tout d’abord que la technique a un effet statistiquement significatif sur les mouvements des mains [$F_{2,22} = 43,48$, $p < 0,0001$] et de la tête [$F_{2,22} = 14,49$, $p < 0,0001$]. Les comparaisons par paires montrent que les distances parcourues par les mains et la tête sont en moyenne plus courtes avec HMD_RC qu’avec HMD_DT ($p = 0,002$) ou TABLET ($p < 0,007$). Nous n’observons pas de différence statistiquement significative entre HMD_DT et TABLET.

2.4.2 Résultats qualitatifs

Fatigue

Nous reportons dans le Tableau 2.1 les résultats des questionnaires sur la fatigue induite par les techniques (questionnaire rempli à la fin de chaque technique, Annexe A). Pour les techniques avec le casque, les participants ressentent de la fatigue surtout au niveau du cou et des épaules. Pour la technique avec la tablette, la fatigue est principalement ressentie au niveau des bras : en effet 6 des 12 participants indiquent 3 sur une échelle de 4. À la fin de l’étude expérimentale (Annexe B), les participants devaient classer les techniques de la moins fatigante à la plus fatigante. Pour la technique la moins fatigante : 6 des 12 participants ont choisi HMD_RC, 3 participants ont choisi HMD_DT et les 3 derniers ont choisi TABLET. La tech-

	HMD_RC		HMD_DT		TABLET	
	<i>moy.</i>	σ	<i>moy.</i>	σ	<i>moy.</i>	σ
Yeux	1,50	0,87	1,50	0,76	1,25	0,60
Visage	1,50	0,95	1,42	0,86	1,0	0,0
Cou/ Épaules	1,83	0,90	2,08	0,76	1,58	0,76
Bras	1,67	0,85	1,5	0,65	2,25	0,83
Doigts	1,25	0,60	1,17	0,37	1,42	0,64
Dos	1,75	0,83	1,75	0,83	1,75	0,83

Tableau 2.1 – Résultats de l'évaluation de la fatigue induite par les techniques par partie du corps (moyennes et écarts types sur une échelle de Likert à 4 points, 1="aucune douleur" et 4="douleur sévère").

	HMD_RC		HMD_DT		TABLET	
<i>Questions</i>	<i>moy.</i>	σ	<i>moy.</i>	σ	<i>moy.</i>	σ
I found the selection technique easy to understand	4,25	0,72	4,50	0,50	4,25	0,60
I found it easy to select the target	4,25	0,83	3,79	0,95	4,00	0,58
I feel that I found quickly the target with this technique	3,88	0,82	4,04	0,59	4,00	0,70
I feel that I selected quickly the target with this technique	3,96	0,78	3,92	0,95	3,75	0,60
I found using this technique physically demanding	2,88	1,00	2,88	0,92	3,33	0,85
This technique is mentally demanding	2,00	0,82	1,67	0,85	1,67	0,75
I found this technique frustrating	1,75	0,72	1,92	1,11	1,92	1,95

Tableau 2.2 – Résultats de l'évaluation subjective des techniques (moyennes et écarts types sur une échelle de Likert à 5 points, 1="pas du tout d'accord" et 5="tout à fait d'accord").

nique TABLET est largement considérée comme la technique la plus fatigante de l'étude : 8 participants sur 12 ont choisi TABLET, contre 2 pour HMD_DT et 2 pour HMD_RC.

Préférences

Globalement, en considérant les résultats pour les 4 premières questions du Tableau 2.2, les participants trouvent les 3 techniques faciles à comprendre (> 4). Selon eux, toutes les techniques leur permettent de trouver et de sélectionner les cibles rapidement ($> 3,75$). Sur les 12 participants, 5 ont préféré HMD_RC, 5 ont préféré HMD_DT et 2 ont préféré TABLET. Les 2 participants préférant TABLET ont expliqué qu'ils n'appréciaient pas la vue immersive du casque et préféraient donc percevoir la scène virtuelle au travers de la tablette.

2.5 Étude expérimentale 1 : discussion

Cette section analyse les résultats obtenus au regard de nos hypothèses. De plus, nous rapportons les stratégies d'interaction observées pour l'interaction au toucher direct avec le casque.

Tâche de recherche : le casque plus efficace que la tablette.

Tout d'abord, nous avons observé que les participants avaient tendance à se pencher en avant (vers la table). Ceci réduisait leurs champs de vision, que ce soit avec le casque ou la tablette, et les obligeait à faire plus de mouvements pour trouver les cibles (s'ils restaient assis au fond de la chaise, les participants pouvaient voir 90% de la scène avec les deux dispositifs). Conformément à notre première hypothèse *H1*, les deux techniques d'interaction basées sur le casque sont plus rapides que la tablette pour la tâche de recherche (quelle que soit la zone). Les capacités du casque peuvent expliquer cette différence. En effet, le casque facilite la visualisation d'une scène 3D entière en un coup d'œil alors que la tablette nécessite d'amener le dispositif devant soi et de le déplacer : la tablette nécessite des mouvements des bras en plus des mouvements de tête.

Tâche de sélection et effet des zones : 1) impact de la zone sur les performances de HMD_DT ; 2) TABLET la moins performante et la plus fatigante même pour des cibles distantes.

Le partitionnement en zones révèle l'impact de la distance de la cible sur les performances des techniques. Une des hypothèses de l'étude (*H2*) était que les deux techniques (HMD_RC et TABLET) permettant d'interagir à distance étaient plus rapides et moins fatigantes que HMD_DT pour les cibles éloignées. Contrairement à HMD_RC et TABLET, la technique HMD_DT semble être impactée par les zones, en particulier pour la zone la plus éloignée de la position initiale (zone 1). En effet, les participants doivent se pencher en avant pour atteindre cette zone. Cela confirme donc l'avantage des techniques permettant une interaction à distance par rapport aux techniques reposant sur une interaction directe sur les cibles. Pourtant, malgré cet avantage, la technique avec la tablette reste la technique la plus lente, même pour les cibles distantes. Les participants rapportent également que tenir une tablette est plus fatigant et plus douloureux que de porter le casque ou de faire de grands mouvements avec HMD_DT par exemple. Ces résultats contredisent donc l'hypothèse *H2*.

	Casque + raycasting				Casque + toucher direct			
	[95]		Notre étude		[95]		Notre étude	
	<i>moy.</i>	σ	<i>moy.</i>	σ	<i>moy.</i>	σ	<i>moy.</i>	σ
I found it easy to select the target	2,5	1,04	4,25	0,83	2,88	0,76	3,79	0,95
I feel that I selected quickly the target with this technique	2,59	0,97	3,96	0,78	2,90	0,98	3,92	0,95
This technique is mentally demanding	3,59	0,92	2,0	0,82	3,33	0,92	1,67	0,85
I found this technique frustrating	3,44	0,92	1,75	0,72	3,04	0,99	1,92	1,11

Tableau 2.3 – Comparaison avec les résultats qualitatifs de Looser et al. [95], moyennes et écarts types sur une échelle de Likert à 5 points (de 1 à 5).

Comparaison avec l'étude de Looser et al.

L'hypothèse $H3$, basée sur les résultats de la technique *Lens* de Looser et al. [95], supposait que la technique avec tablette était en moyenne plus rapide et préférée par les participants pour la sélection de cibles virtuelles. Les résultats obtenus contredisent cette hypothèse. En effet, en plus d'être en moyenne plus rapides quelle que soit la position de la cible sur la table, les techniques avec le casque sont également préférées à celle avec la tablette. Le Tableau 2.3 oppose les résultats qualitatifs obtenus lors de cette étude et ceux de l'étude de Looser et al. [95]. Les techniques de cette étude basées sur le casque obtiennent de bien meilleurs résultats que celles de Looser et al. Notre étude permet donc de mettre à jour les résultats des travaux de Looser et al. et montre qu'avec les casques actuels, les techniques d'interaction au toucher direct et de raycasting sont beaucoup mieux perçues et utilisables qu'il y a 10 ans.

Analyse des mouvements du corps.

Les résultats obtenus en termes de performances et de retours qualitatifs sont bien différents selon le dispositif utilisé. Une analyse des mouvements du corps des participants a alors été réalisée afin de mieux comprendre ces différences.

Avec TABLET, nous avons observé que les participants (1) placent la tablette entre eux et la cible en tenant la tablette à *deux mains*, puis (2) lâche la tablette pour ne la tenir que d'une main et utiliser la seconde pour sélectionner la cible. Le fait d'utiliser les deux mains permet certainement de réduire la fatigue due au poids du dispositif et aussi d'améliorer la stabilité de la technique mais peut aussi faire perdre du temps aux participants (la main dominante devant être libérée pour la sélection de la cible).

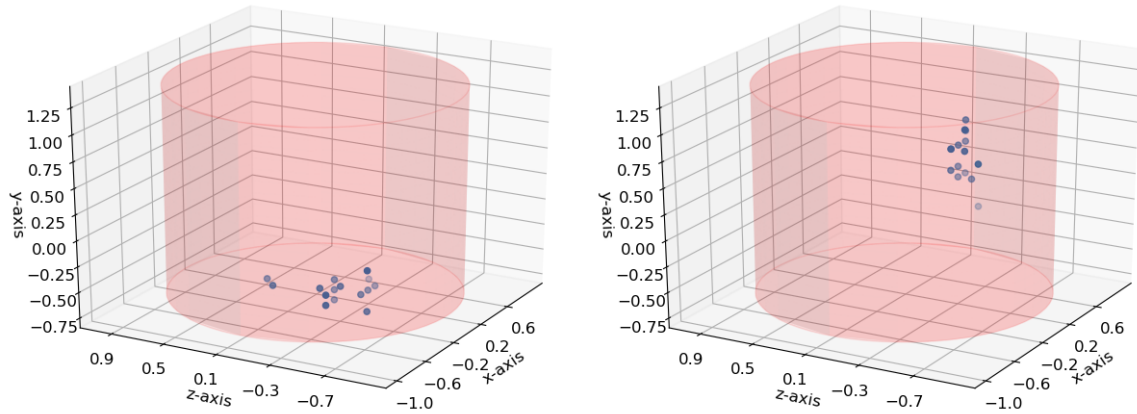


FIGURE 2.16 – Stratégies observées pour la sélection avec HMD_DT : interaction au toucher direct sur la table (à gauche) et interaction dans l’air (à droite). Le cylindre rouge représente la cible, les points bleus sont les positions du bout du doigt au moment de la validation de la sélection.

Alors que la sélection de cibles avec une tablette implique une interaction en deux étapes séquentielles (placement de la tablette puis sélection de la cible), le casque permet des mouvements découplés et synchronisés de la tête et des bras. Avec un casque de RA, les participants peuvent donc se tourner vers la cible et la sélectionner en même temps.

De plus, lors de l’analyse des mouvements des mains des participants, nous avons remarqué que la distance parcourue par les mains avec TABLET était similaire à celle avec HMD_DT, et également bien supérieure à celle avec HMD_RC. Contrairement à HMD_RC, les participants ne semblent pas utiliser la sélection à distance pourtant permise avec la tablette et préfèrent la déplacer proche des cibles pour les sélectionner.

L’interaction en deux étapes séquentielles induite par la tablette ainsi que les grands mouvements des bras des participants lorsqu’ils utilisent la tablette pourraient donc expliquer la différence de performance et de fatigue ressentie entre les deux dispositifs.

Stratégies d’interaction avec HMD_DT : interaction au toucher direct sur la table et interaction dans l’air.

Deux stratégies pour accomplir la tâche de sélection ont été observées. Nous nous attendions à ce que les participants mettent le doigt sur la table pour sélectionner une cible. En effet, sachant que les cylindres sont placés sur la table, avec cette stratégie les participants n’auraient qu’à gérer 2 des 3 dimensions pour le placement du doigt.

Cependant, seulement 4 participants sur 12 touchaient toujours la table pendant cette tâche tandis que 7 participants sélectionnaient toujours la cible dans l'air. La Figure 2.16 montre les positions du doigt de 2 participants par rapport à la cible au moment de la validation de la sélection. L'image de gauche montre les résultats d'un participant qui adopte la stratégie du toucher sur la table pour chaque sélection et l'image de droite montre les résultats d'un participant qui adopte la stratégie d'interaction dans l'air. Tous les participants sauf un ont exécuté une seule de ces deux stratégies au cours de l'étude. Nous émettons l'hypothèse que le premier groupe (stratégie au toucher sur la table) s'appuie à la fois sur le référentiel du monde réel et du monde virtuel pour accomplir la tâche, tandis que le second groupe (stratégie d'interaction dans l'air) effectue la tâche uniquement dans le référentiel du monde virtuel et oublie la table.

2.6 Étude expérimentale 2 : sélection d'objets 3D occultés en environnements denses

2.6.1 Objectif de l'étude et hypothèse

Nous avons observé précédemment qu'en environnements épars (étude 1, Section 2.3) les deux techniques avec le casque étaient plus performantes et préférées par les participants par rapport à la technique avec la tablette. Lorsque nous interagissons avec la carte augmentée d'une grande ville par exemple, plusieurs bâtiments peuvent être cachés derrière d'autres, il peut donc être difficile de les sélectionner. De plus, les cibles peuvent être plus ou moins grandes. Dans ces environnements denses (occultation, petites cibles), les techniques avec casque sont-elles toujours plus performantes que celle avec tablette ? Nous proposons donc une seconde étude expérimentale pour étudier l'impact de ce type d'environnements sur les performances des techniques et sur la manière dont nous utilisons ces techniques. Nous faisons l'hypothèse (H_4) que les techniques HMD_DT et HMD_RC sont toujours plus performantes et préférées à la technique TABLET. Cependant, la technique HMD_DT serait plus rapide pour sélectionner des cibles cachées que les deux autres techniques basées sur la métaphore du raycasting car le rayon s'arrête au premier obstacle rencontré.

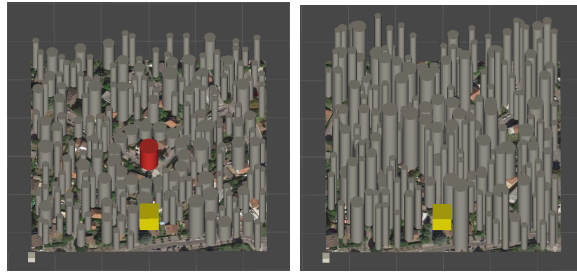


FIGURE 2.17 – Deux exemples de scènes denses pour l’étude 2. Le cube jaune indique la cible initiale. À gauche, la cible finale en rouge est visible de la position initiale. À droite, la cible est cachée par des distracteurs.

2.6.2 Système expérimental

Les dispositifs et les techniques sont les mêmes que pour la première étude. La configuration du système expérimental est également similaire à la précédente étude (Section 2.3.3). Cependant, sachant que les objets sont plus hauts et que les cibles peuvent être occultées, les participants devront par moment se pencher au-dessus de la table. Nous choisissons donc de laisser les participants debout pour cette étude. La chaise est retirée pour cette étude et une ligne est tracée au sol à 70 cm de la table pour indiquer la position de départ de la tâche. À cette position, les cibles occultées ne sont pas visibles et les participants doivent avancer pour les voir et les sélectionner.

2.6.3 Tâche

Les tâches et le protocole de cette étude sont inspirés de ceux de Vanacken et al. [163, 164]. Comme le présente la Figure 2.17, chaque scène de cette étude est composée de 200 distracteurs et de deux cibles : une cible initiale représentée par un cube jaune et une cible finale représentée par un cylindre rouge.

La cible initiale est toujours visible et est placée au même endroit dans toutes les scènes. L’effet de la position de la cible ayant déjà été étudié lors de la première étude, la position de la cible finale est fixe dans cette seconde étude : elle est toujours placée à une distance de 17 cm de la cible initiale. La cible finale peut être soit visible, soit cachée par des distracteurs placés autour d’elle comme illustré à la Figure 2.17. Les distracteurs, sauf ceux placés directement autour de la cible finale, sont placés aléatoirement et mesurent entre 2 et 8 cm de haut lorsque la cible finale est visible et entre 4 et 24 cm de haut lorsque la cible est occultée. Leur largeur est choisie aléatoirement entre 1 cm et 4 cm.

Chaque tâche de sélection comprend les étapes suivantes :

1. L'utilisatrice se place en position initiale derrière la ligne.
2. Pendant 3,5 secondes, seule la cible initiale (cube jaune) et la cible finale rouge sont affichées.
3. Pendant 4 secondes, tous les distracteurs grandissent jusqu'à atteindre leurs tailles finales.
4. La cible initiale passe du jaune au vert pour indiquer le début de la tâche. L'utilisatrice doit tout d'abord sélectionner la cible initiale. Lorsque la sélection de cette cible est terminée, cette cible disparaît. L'utilisatrice sélectionne ensuite la cible finale. Une barre de progression (identique à celle utilisée dans la première étude) est affichée dès qu'une collision avec un objet est détectée et indique l'avancement de la validation. L'utilisatrice doit maintenir la sélection pendant une seconde.
5. La tâche se termine par la validation de la sélection de la cible finale.

Les étapes 2 & 3 permettent d'éliminer la phase de recherche (déjà étudiée lors de la première étude). Les participants sont libres de se déplacer après l'étape 2.

2.6.4 Protocole

L'étude expérimentale est conçue comme une étude expérimentale intra-sujet, avec les quatre variables indépendantes (TECH, V, S, DS) suivantes :

- *La technique d'interaction utilisée (TECH)* : TABLET, HMD_DT, HMD_RC (Section 2.1).
- *La condition de visibilité de la cible (V)* : Lorsque les participants sont en position initiale (c'est-à-dire derrière la ligne de départ), la cible finale rouge est soit visible soit cachée (Figure 2.17).
- *La taille de la cible (S)* : Le diamètre de la cible cylindrique est de 1,5 cm (petite cible) ou 4 cm (grande cible). Les cibles mesurent toujours 6 cm de haut.
- *L'espacement de densité (DS)* : Pour chaque scène, plusieurs distracteurs sont positionnés de manière circulaire autour de la cible rouge (Figure 2.18), à une distance comprise entre 1 cm (DS petit) et 4 cm (DS grand). Pour une cible occultée, ces distracteurs mesurent 18 cm de haut. Pour une cible visible, ils mesurent 3 cm (hauteur de la cible rouge / 2).

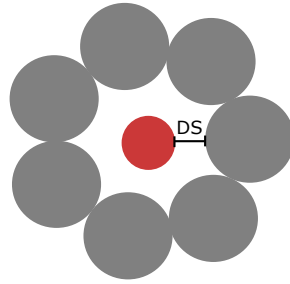


FIGURE 2.18 – Illustration de l’espacement de densité (DS) : la cible finale rouge est entourée par des distracteurs gris. DS représente la distance entre ces distracteurs et la cible.

Pour garantir des conditions expérimentales similaires entre les participants, un ensemble de 24 scènes 3D (c’est-à-dire la cible initiale, les cibles finales et les distracteurs) ont été calculées avant l’étude et sont partagées entre tous les participants et toutes les techniques. L’ordre des techniques est contrebalancé par un carré latin et les scènes sont présentées dans un ordre aléatoire pour chacune des techniques.

L’étude expérimentale commence par une brève introduction. Pour que les participants se sentent à l’aise avec toutes les techniques et comprennent bien les différentes étapes des tâches de sélection, ils débutent par une phase d’entraînement (autant de temps que nécessaire). Ils réalisent ensuite les 24 sélections pour chacune des trois techniques. Les participants remplissent également un questionnaire après chaque technique (Annexe C) et un questionnaire de synthèse à la fin de l’étude (Annexe D) pour comparer les trois techniques. L’étude a duré environ 50 minutes par participant.

Nous avons recruté 12 participants de notre étude précédente et de l’étude pilote sachant que ces deux études précédentes avaient une durée similaire de l’ordre de 40 minutes. Ainsi, tous les participants avaient déjà utilisé les trois techniques d’interaction pendant environ 40 minutes, permettant ainsi de réduire l’effet WoW et la phase d’apprentissage.

Au regard de notre hypothèse $H4$, nous considérons deux métriques objectives (le temps de réalisation de la tâche et la distance moyenne des mouvements des participants lors d’une tâche) et une métrique subjective (préférence des utilisateurs). Les participants peuvent se déplacer librement après l’étape 2 (voir étapes présentées dans la Section 2.6.3). Ils peuvent donc se déplacer pendant que les distracteurs apparaissent (étape 3) ou attendre que la cible initiale devienne verte (étape 4). Pour éviter tout biais et prendre en compte les déplacements de tous les participants, le temps de réalisation de la tâche est alors le temps entre l’étape 2 et la validation de la sélection de la cible finale. La métrique de distance représente la

distance parcourue par les participants. Elle est mesurée grâce à des marqueurs réfléchissants (suivis par les caméras OptiTrack) placés sur la tête des participants. Pour la mesure subjective, nous avons conçu un questionnaire pour chaque technique (voir Annexe C et Tableau 2.4 des résultats) pour évaluer la technique d'interaction avec les deux conditions de visibilité. Un questionnaire final (Annexe D) est aussi présenté aux participants pour choisir la meilleure technique pour chaque question du Tableau 2.4, pour classer les techniques de la moins précise à la plus précise ; et pour les classer par ordre de préférence.

2.7 Étude expérimentale 2 : résultats

Pour l'analyse, nous rapportons les résultats de 11 des 12 participants. Nous avons dû retirer un des participants car ses temps de réalisation étaient bien plus importants que les autres (la différence entre ses temps de réalisation et la moyenne du groupe était de plus de 6 fois l'écart type). La normalité des données est testée à l'aide du test Shapiro-Wilk. Pour les données normalement distribuées, nous utilisons un test statistique par analyse de variance (ANOVA) et des tests de Student (ou tests t) avec ajustement de Bonferroni pour les comparaisons par paires. Les données qui ne suivent pas une loi normale, sont pré-traitées par une transformation de rang aligné (ART) [175] avant d'utiliser une ANOVA avec une correction de Bonferroni pour les comparaisons par paires. Dans les graphiques suivants, nous représentons les valeurs moyennes et les intervalles de confiance (IC) à 95% calculés avec la méthode bootstrap Bca (la fonction du bootstrap étant la moyenne) [51].

2.7.1 Résultats quantitatifs

Temps moyens de réalisation de la tâche

Nous observons un effet statistiquement significatif de la technique *TECH* [$F_{2,20} = 19,96$, $p < 0,0001$], de la taille de la cible *S* [$F_{1,10} = 80,1$, $p < 0,0001$] et de la condition de visibilité de la cible *V* [$F_{1,10} = 39,65$, $p < 0,0001$] sur le temps de réalisation de la tâche. Cependant, nous n'avons pas observé d'effet statistiquement significatif de la densité *DS*. Le temps moyen de réalisation de cette tâche est de : 11349 ms ($IC = [10908, 11863]$) pour HMD_DT, 11549 ms ($IC = [11191, 11972]$) pour HMD_RC et 12580 ms ($IC = [11986, 13013]$) pour TABLET. Les comparaisons par paires montrent une différence statistiquement significative entre la technique TABLET et les deux techniques utilisant le casque (TABLET-HMD_DT : $p = 0,004$; TABLET-HMD_RC : $p = 0,016$).

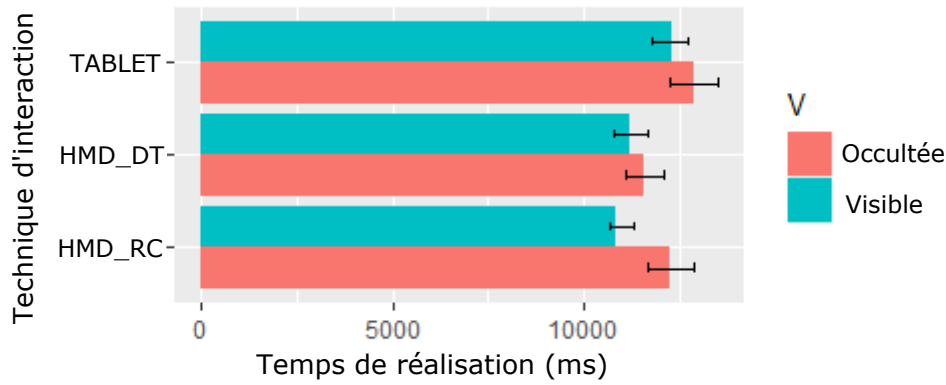


FIGURE 2.19 – Effet d’interaction entre la condition de visibilité de la cible et la technique sur le temps de réalisation de la tâche de sélection.

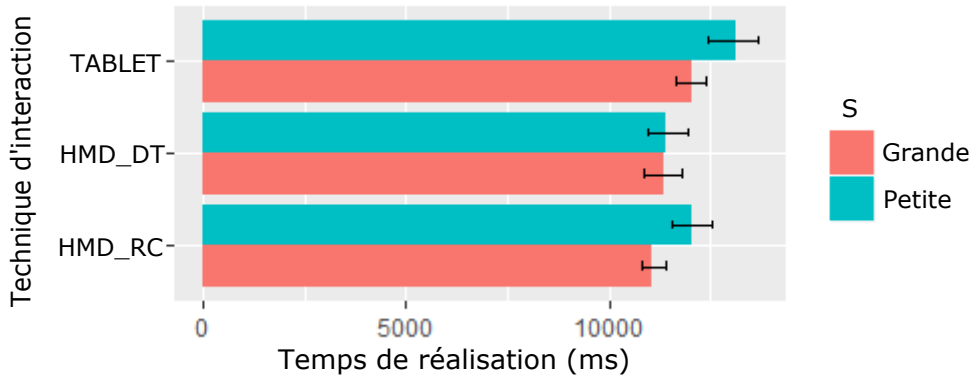


FIGURE 2.20 – Effet d’interaction entre la taille de la cible et la technique sur le temps de réalisation de la tâche de sélection.

Effet d’interaction $V \times TECH$. Nous observons un effet d’interaction statistiquement significatif entre V et $TECH$ [$F_{2,50} = 5,8$, $p = 0,005$]. Les résultats sont présentés à la Figure 2.19. Nous observons un effet statistiquement significatif de V uniquement sur HMD_RC ($p = 0,001$) pour le temps de réalisation de la tâche. Pour les cibles visibles, les comparaisons par paires montrent que les deux techniques avec le casque sont significativement plus rapides que TABLET (HMD_DT–TABLET : $p = 0,02$; HMD_RC–TABLET : $p = 0,006$). Pour les cibles cachées, les comparaisons par paires montrent que HMD_DT est significativement plus rapide que TABLET ($p = 0,01$). Nous n’observons pas de différence significative entre HMD_RC et TABLET, ni entre les deux techniques utilisant le casque.

Effet d’interaction $S \times TECH$. Nous observons également un effet d’interaction statistiquement significatif entre S et $TECH$ [$F_{2,50} = 6,13$, $p = 0,004$] sur le temps

de réalisation de la tâche (données pré-traitées avec une ART). Les résultats pour cet effet d'interaction sont présentés à la Figure 2.20. Les comparaisons par paires montrent que les temps des techniques HMD_RC ($p = 0,006$) et TABLET ($p = 0,011$) sont impactés par la taille de la cible. Nous n'observons pas de différence significative entre les tailles des cibles pour HMD_DT.

Dans le cas de la sélection de petites cibles, les deux techniques avec le casque sont significativement plus rapides que TABLET (HMD_DT-TABLET : $p = 0,0005$; HMD_RC-TABLET : $p = 0,033$). Il en est de même pour les grandes cibles : HMD_DT et HMD_RC sont significativement plus rapides que TABLET (HMD_DT-TABLET : $p = 0,045$; HMD_RC-TABLET : $p = 0,01$).

Distances parcourues

Nous observons un effet statistiquement significatif de la technique sur la distance parcourue par les participants [$F_{2,20} = 6,29$, $p = 0,008$]. Les comparaisons par paires montrent que la distance moyenne des participants est significativement plus courte ($p = 0,036$) avec TABLET ($m = 0,84$ m, $IC = [0,72; 0,96]$) qu'avec HMD_DT ($m = 1,09$ m, $IC = [1,03; 1,16]$). Nous n'observons pas de différence statistiquement significative entre HMD_RC ($m = 0,94$ m, $IC = [0,77; 1,09]$) et les deux autres techniques. Nous observons également un effet statistiquement significatif de V [$F_{1,10} = 29,3$, $p = 0,0003$] sur la distance moyenne parcourue par les participants. Par contre, nous n'avons pas observé d'effet statistiquement significatif de la taille de la cible S ou de la densité DS.

Effet d'interaction $V \times TECH$. Nous observons également un effet d'interaction statistiquement significatif entre V et TECH [$F_{2,50} = 10,25$, $p = 0,0002$] sur la distance moyenne parcourue par les participants. Les résultats pour cet effet d'interaction sont présentés à la Figure 2.21. Les comparaisons par paires montrent que la condition où les cibles sont occultées augmentent significativement la distance moyenne parcourue quelle que soit la technique utilisée (HMD_DT : $p = 0,01$; HMD_RC : $p = 0,0002$; TABLET : $p = 0,03$).

De plus pour les cibles occultées, la distance parcourue est significativement plus courte avec TABLET (TABLET-HMD_DT : $p = 0,006$; TABLET-HMD_RC : $p = 0,01$). Par contre pour les cibles visibles, la distance parcourue est significativement plus courte avec les techniques d'interaction à distance HMD_RC et TABLET qu'avec HMD_DT (HMD_DT-HMD_RC : $p = 0,009$; HMD_DT-TABLET : $p = 0,04$).

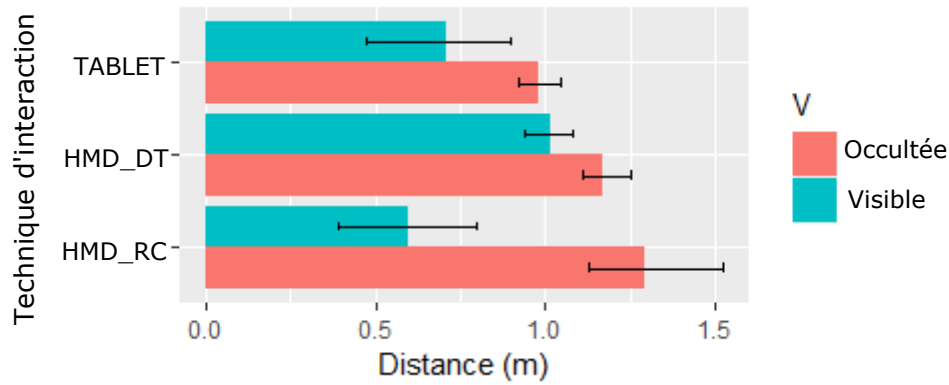


FIGURE 2.21 – Effet d’interaction entre la condition de visibilité de la cible et la technique sur la distance parcourue par les participants.

Questions	HMD_RC		HMD_DT		TABLET	
	Visible	Occultée	Visible	Occultée	Visible	Occultée
Q1. It is easy to select the target with this technique	4,9 ± 0,3	3,7 ± 1,0	4,7 ± 0,5	3,7 ± 1,1	4,0 ± 1,2	3,3 ± 1,4
Q2. I select the target quickly with this technique	4,8 ± 0,4	3,6 ± 0,9	4,8 ± 0,4	3,8 ± 0,9	4,2 ± 1,1	3,4 ± 1,2
Q3. This technique is physically demanding	2,3 ± 1,2	2,7 ± 1,2	3,2 ± 1,1	3,6 ± 1,1	2,6 ± 1,1	3,0 ± 1,1
Q4. This technique is mentally demanding	1,5 ± 0,7	1,6 ± 0,8	1,9 ± 1,4	2,1 ± 1,6	1,5 ± 0,7	1,6 ± 0,8
Q5. This technique is frustrating	1,3 ± 0,5	1,6 ± 0,7	1,4 ± 0,5	2,0 ± 1,9	2,3 ± 1,3	2,3 ± 1,6
Q6. I felt I was accurate with this technique	4,5 ± 0,5	3,6 ± 0,8	4,6 ± 0,5	4,0 ± 1,0	3,9 ± 0,9	3,6 ± 1,4

Tableau 2.4 – Résultats de l’évaluation subjective des trois techniques d’interaction selon la condition de visibilité de la cible. Nous reportons les moyennes et écarts types des notes des participants sur une échelle de Likert à 5 points (1=“pas du tout d’accord”, 5=“tout à fait d’accord”).

2.7.2 Résultats qualitatifs

Les participants ont trouvé que la sélection était simple et rapide quelle que soit la technique utilisée (Tableau 2.4). Les 3 techniques ont un classement plus bas pour la condition avec cibles occultées par rapport à la condition avec cibles visibles. De plus, les sélections avec les techniques utilisant le casque sont perçues comme plus simples et plus rapides que celle avec la tablette, quelle que soit la visibilité de la cible. En termes de précision, les participants se sont sentis précis avec toutes les techniques. Leurs réponses montrent cependant qu’ils s’estiment plus précis en condition visible et avec les deux techniques basées sur le casque.

Dans le questionnaire final, pour des cibles visibles, ils choisissent HMD_RC

comme la technique la plus simple (7 participants pour HMD_RC, 3 pour HMD_DT, 1 pour TABLET) et la plus rapide (6 participants pour HMD_RC, 2 pour HMD_DT, 3 pour TABLET). Pour les cibles cachées par contre, ils choisissent HMD_DT comme la technique la plus simple (6 participants pour HMD_DT, 3 pour HMD_RC, 2 pour TABLET) et la plus rapide (6 participants pour HMD_DT, 4 pour HMD_RC et 1 pour TABLET). Ils ont également reporté que les deux techniques d'interaction à distance semblaient être les moins précises (6 participants pour HMD_RC, 5 pour TABLET). La technique HMD_DT est la technique la plus précise. Enfin, ils préfèrent clairement les techniques avec casque à celle avec tablette (6 participants pour HMD_RC, 5 pour HMD_DT). La technique la moins préférée est celle avec la tablette pour 6 des 11 participants (2 participants pour HMD_RC, 3 pour HMD_DT).

2.8 Étude expérimentale 2 : discussion

Casque plus performant et plus apprécié par les participants que la tablette en environnements denses.

Cette deuxième étude montre qu'en environnements denses, la technique avec tablette reste moins appréciée et moins rapide que les deux techniques avec casque, quelles que soient la condition de visibilité et la taille de la cible. L'hypothèse $H4$ est confirmée.

Technique HMD_DT moins impactée par la visibilité et la taille de la cible.

HMD_DT et HMD_RC présentent des temps de réalisation similaires. Cependant, la technique HMD_DT ne semble pas être affectée par la taille de la cible, contrairement aux techniques d'interaction à distance HMD_RC et TABLET. Les cibles étant posées sur la table, nous supposons que le support de la table a pu aider les participants pour la sélection de petites cibles avec HMD_DT. Cela peut être un premier avantage de l'interaction au toucher direct avec casque par rapport à l'interaction par raycasting (HMD_RC). Contrairement à HMD_RC, HMD_DT est également plus robuste aux différentes conditions de visibilité de la cible, ce qui peut être un second avantage de cette technique.

Plusieurs stratégies observées avec HMD_DT.

Nous avons analysé les mouvements du bout du doigt des participants entre

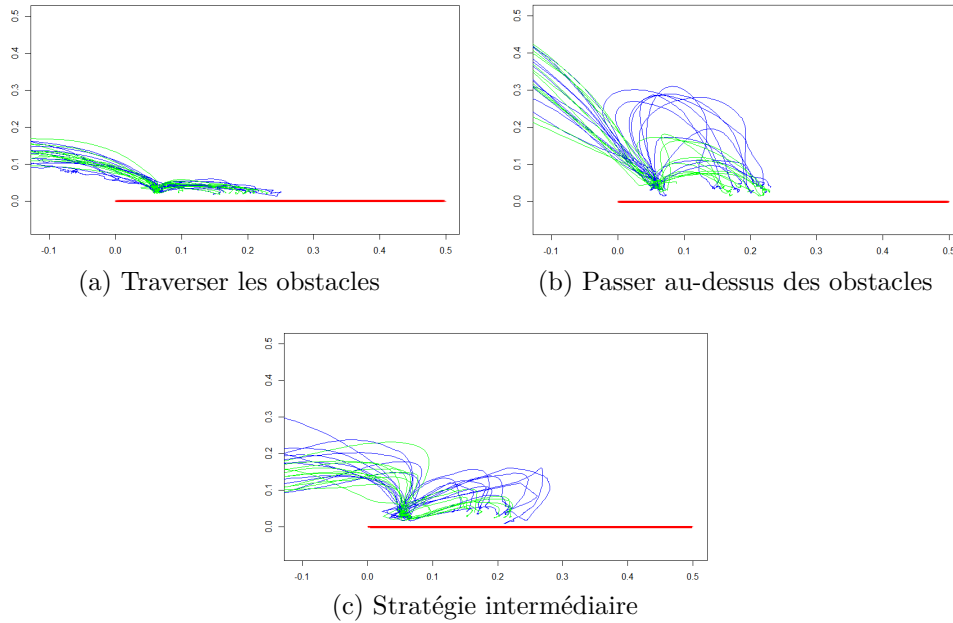


FIGURE 2.22 – Trajectoires du doigt de 3 participants : sélection de la cible initiale puis sélection de la cible finale au-dessus de la table (en rouge). En vert, les trajectoires pour les sélections de cibles visibles. En bleu, les trajectoires pour les sélections de cibles cachées.

la sélection de la cible initiale et la sélection de la cible rouge finale pour les deux conditions de visibilité. Pour illustrer ces mouvements, la Figure 2.22 présente les trajectoires de 3 participants (vue de côté). La ligne rouge représente la surface de la table, les lignes vertes sont les trajectoires pour les sélections de cibles visibles et les lignes bleues sont les sélections de cibles cachées. Alors que les trajectoires vertes sont assez droites et plates, trois stratégies sont observées pour sélectionner la cible rouge lorsqu'elle est cachée derrière des obstacles élevés :

- (a) *Traverser les obstacles.* La stratégie la plus évidente est d'atteindre directement la cible rouge en traversant les obstacles. Cette stratégie est adoptée par la majeure partie des participants (9 sur 11 participants). Un exemple est présenté à la Figure 2.22a : il n'y a aucune différence de trajectoires entre les deux conditions de visibilité.
- (b) *Passer au-dessus des obstacles.* Pour la sélection de cibles cachées, une possibilité est de passer au-dessus des obstacles pour atteindre la cible. Cette stratégie a été observée sur les trajectoires d'un seul participant. Les trajectoires sont présentées à la Figure 2.22b : il y a un saut du doigt entre les deux cibles pour la sélection de cibles cachées. Dans la plupart des cas, la hauteur du saut est supérieure à 20 cm ce qui correspond à la taille des obstacles (hauteur des

distracteurs étant 18 cm).

- (c) *Stratégie intermédiaire*. Un seul participant a utilisé une stratégie intermédiaire (Figure 2.22c). C'est-à-dire que nous observons un léger saut contrairement au cas a, pour la sélection de cibles cachées mais bien moindre que les sauts observés pour le cas b.

Lors du questionnaire final, nous avons demandé aux participants s'ils traversaient des objets virtuels pour sélectionner une cible cachée et à quelle fréquence (étaient-ils conscients de leurs stratégies?). Les réponses sont cohérentes avec l'analyse des mouvements : 1 participant a répondu "Jamais", 1 autre a répondu "Parfois", 2 ont répondu "Souvent" et les autres ont répondu "Toujours".

3. Résumé de l'étude sur les dispositifs d'affichage en réalité augmentée sur table

En étudiant les travaux sur la RA sur table (Section 1), nous avons souligné que les systèmes différaient en termes de dispositifs utilisés pour afficher la scène virtuelle : certains systèmes reposent sur un dispositif portable (téléphone, tablette) et d'autres sur un casque de RA. Pour les tâches de pointage d'objets 3D, ces deux types de dispositifs impliquent des techniques d'interaction complètement différentes. Le choix du dispositif est donc un choix de conception important car il impactera considérablement la manière dont l'utilisatrice va interagir avec les objets virtuels.

Cette étude est un premier pas vers une meilleure compréhension des avantages et inconvénients de ces dispositifs pour la sélection d'objets virtuels. Trois techniques de sélection ont été comparées, correspondant aux techniques les plus couramment utilisées avec les tablettes et les casques de RA :

- Une technique d'interaction tactile sur l'écran avec la tablette (*TABLET*),
- Une technique d'interaction au toucher direct sur l'objet avec un casque de RA (*HMD_DT*),
- Une technique d'interaction par raycasting avec un casque de RA (*HMD_RC*).

Deux études expérimentales ont permis de comparer ces techniques dans des environnements épars (Étude 1, Section 2.3) mais également dans des environnements denses (Étude 2, Section 2.6). Un résumé détaillé des résultats obtenus est présenté dans le Tableau 3.1. Nous rappelons les résultats clés à considérer lors de la conception d'un système de RA sur table :

Les deux techniques avec le casque sont plus rapides et préférées à la technique avec tablette quelle que soit la densité de l'environnement. Les tâches de recherche et de sélection sont réalisées plus rapidement avec un casque qu'avec la tablette, quelle que soit la position de la cible sur la table ou la densité de l'environnement. Contrairement à la tablette, seuls les mouvements de la tête permettent le parcours visuel de la scène ce qui permet à l'utilisatrice de pouvoir interagir avec les objets (via des mouvements des bras) en parallèle de la phase de

recherche. Ceci pourrait justifier les meilleures performances des techniques avec casque par rapport à la tablette, et ainsi être l'avantage principal du casque de RA face à la tablette. Les participants préfèrent également largement le casque et expliquent que le porter est moins fatigant que de tenir une tablette.

La technique HMD_DT perçue comme la plus simple et la plus rapide pour des cibles occultées. Dans la seconde étude, la densité n'a pas eu d'effet significatif sur le temps de sélection des techniques, ce qui confirme l'observation de [95]. Non étudié dans [95], nous avons également pris en compte la condition de visibilité (c'est-à-dire cible visible ou cachée) dans cette étude. Alors que les deux techniques HMD_RC et HMD_DT obtiennent de performances similaires, seule la technique HMD_RC est significativement impactée par l'occultation des cibles. En environnements denses et avec des cibles accessibles par le bras de l'utilisatrice, la technique HMD_DT semble bien appropriée car elle est directe, intuitive et peu impactée par l'occultation des cibles. Les résultats montrent qu'elle est également peu impactée par la taille des cibles. Cependant, ces études impliquaient des cibles placées directement sur la table, le support de la table a donc pu être une aide lors des sélections.

Stratégies des participants avec la technique HMD_DT. En analysant les mouvements du bout du doigt des participants, plusieurs stratégies d'interaction ont été observées avec HMD_DT. Pour rappel, les objets à sélectionner étaient toujours placés directement sur la table. Nous avons tout d'abord observé que certains participants sélectionnaient les cibles par des interactions dans l'air, alors que d'autres les sélectionnaient en plaçant leur doigt sur la table, réduisant la tâche 3D à une tâche 2D par interaction au toucher sur la base de l'objet. De plus, lors de la sélection d'objets cachés derrière des obstacles, nous avons aussi observé la réticence de certains participants à traverser des objets virtuels avec leur main.

	<i>Bénéfices</i>	<i>Limitations</i>
<i>Interaction avec la tablette</i>	<ul style="list-style-type: none"> • Plus courte distance parcourue par les participants pour sélectionner les cibles cachées • Technique ressentie comme précise, simple et rapide • Pas d'effet significatif observé de l'occultation des cibles sur le temps de réalisation de la tâche de sélection • Pas d'effet significatif observé de la zone (position de la cible) sur le temps de réalisation de la tâche de sélection 	<ul style="list-style-type: none"> • Plus lente que les techniques avec casque • Plus fatigante que les techniques avec casque • Moins préférée que les techniques avec casque • Perçue comme moins facile, moins rapide et moins précise que les techniques avec casque • Effet significatif de la taille de la cible sur le temps de réalisation de la tâche • Effet significatif de l'occultation de la cible sur la distance parcourue par les participants
<i>Interactions avec le casque</i>	<ul style="list-style-type: none"> • Techniques plus rapides que celle avec tablette • Techniques ressenties comme plus précises, plus simples et plus rapides que celle avec tablette • Techniques préférées à celle avec tablette • Techniques moins fatigantes que celle avec tablette 	<ul style="list-style-type: none"> • Effet significatif de l'occultation de la cible sur la distance parcourue par les participants • Effet significatif de la zone pour la tâche de recherche
<i>HMD_DT</i>	<ul style="list-style-type: none"> • Perçue comme la technique la plus simple et la plus rapide pour des cibles cachées • Perçue comme la technique la plus précise • Pas d'effet significatif observé de l'occultation et de la taille de la cible sur le temps de réalisation de la tâche 	<ul style="list-style-type: none"> • Plus lent que HMD_RC pour sélectionner des cibles distantes • Technique avec la plus longue distance parcourue par les participants pour des cibles visibles • Effet de la zone sur le temps de réalisation de la tâche de sélection
<i>HMD_RC</i>	<ul style="list-style-type: none"> • Plus rapide que HMD_DT pour les cibles distantes (zone 1) • Pas d'effet significatif observé de la zone sur le temps de réalisation de la tâche de sélection • Technique avec la plus courte distance parcourue par les participants pour les cibles visibles 	<ul style="list-style-type: none"> • Effet de l'occultation et de la taille de la cible sur le temps de réalisation de la tâche de sélection • Perçue comme la technique la moins précise de l'étude (pas de différence significative observé entre les techniques)

Tableau 3.1 – Bénéfices et limitations des techniques comparées.

Dispositifs de réalité augmentée	
Techniques d'interaction	TABLETTE
	CASQUE DE RA
	<div>Interaction au toucher direct sur l'écran</div> <div>Raycasting</div> <div>→ Chapitre II</div> <div>Toucher direct / Main virtuelle</div> <div>→ Chapitre III</div>

Tableau 3.2 – Espace de conception – Résultats obtenus pour l'étude sur les dispositifs de RA et organisation des autres travaux décrits dans ce manuscrit.

Au vu des résultats de ce Chapitre I, nous nous focalisons sur les interactions avec un casque de RA dans la suite de nos travaux (Tableau 3.2). Néanmoins, l'utilisation de techniques avec tablette n'est pas à exclure pour les systèmes de RA sur table, nos résultats remettent en question *l'implémentation standard* de la sélection avec tablette (technique la plus courante). Il serait intéressant de comparer avec d'autres techniques plus avancées utilisant la tablette, comme une technique permettant le zoom (par un geste de pincement sur l'écran par exemple). Cependant, une telle implémentation ne permettrait pas de coupler une vue augmentée avec des informations *physiques* sur la table : la correspondance entre informations physiques et informations virtuelles seraient en effet perdue après un zoom. Les résultats obtenus pour les deux techniques d'interaction basées sur le casque de RA soulignent des avantages et des limitations. Ces résultats nous ont conduits à considérer deux pistes de recherche complémentaires, chacune étant dédiée à une technique.

Tout d'abord les résultats montrent que l'interaction par raycasting (HMD_RC) permet une sélection efficace des cibles, et est particulièrement utile pour des cibles distantes. Elle devient cependant moins efficace en environnements denses et quand la cible est occultée. Dans le Chapitre II, nous étudions les améliorations qui peuvent être apportées au raycasting pour résoudre ces limitations.

De plus, l'avantage potentiel apporté par la table lors de la sélection de petites cibles au toucher direct (HMD_DT), ainsi que les stratégies observées avec l'interaction au toucher direct, nous ont également incité à étudier les techniques de sélection utilisant l'interaction tactile sur la table au lieu d'interactions 3D purement dans l'air. Ce travail est décrit dans le Chapitre III.

Chapitre II

Sélection d'objets distants : amélioration du raycasting pour la sélection en environnements denses

Introduction

Lors d'interaction sur une table interactive, certains objets ne sont pas directement accessibles par la main de l'utilisatrice en raison de la taille de la table. Pour résoudre ce problème, Banerjee et al. [17] préconisent d'utiliser l'espace au-dessus de la table pour interagir avec des objets distants. L'interaction classique au toucher direct sur table est alors remplacé par le pointage par raycasting pour la sélection d'objets distants. Dans le cas de réalité augmentée (RA) sur table, le chapitre précédent (Chapitre I) souligne les avantages du raycasting pour la sélection d'objets distants de l'utilisatrice. Cependant, nous avons également observé des limitations de cette technique d'interaction, en particulier en environnements denses et pour la sélection d'objets occultés. Ainsi, ce chapitre se concentre sur la tâche de sélection 3D par raycasting et considère différentes alternatives pour atténuer/éliminer ces limitations.

Le chapitre¹ est structuré en deux sections : la première section (Section 4) dresse un état de l'art des extensions du raycasting tandis que la section suivante (Section 5) présente une nouvelle technique d'extension du raycasting, appelée *RayLens*, qui combine une lentille grossissante et un rayon pour faciliter la sélection d'objets distants en environnements denses.

Sommaire

4	État de l'art sur les facilitations de pointage pour le raycasting	60
4.1	Facilitations de pointage pour le raycasting	60
4.2	Les lentilles en RA et RV	75
5	Contribution : extension du raycasting avec une lentille grossissante	83
5.1	Éléments de conception : <i>RayLens</i> et <i>RaySlider</i>	83
5.2	Étude expérimentale : évaluation de l'efficacité de <i>RayLens</i> face au raycasting et à <i>RaySlider</i>	89
5.3	Étude expérimentale : résultats	95
5.4	Étude expérimentale : discussion	103
6	Résumé de l'étude sur les améliorations du raycasting	108

1. Les travaux présentés dans ce chapitre ont été publiés à la conférence ACM-INTERACT 2021, the 2021 Conference on Human-Computer Interaction. Carole Plasson, Dominique Cunin, Yann Laurillau, & Laurence Nigay. A Lens-based Extension of Raycasting for Accurate Selection in Dense 3D Environments.

4. État de l’art sur les facilitations de pointage pour le raycasting

Dans cette section, nous passons en revue différentes extensions du raycasting et présentons les avantages et limitations pour une sélection précise d’objets 3D en environnements denses. Cette section explore également le principe de lentilles magiques virtuelles comme alternative pour améliorer la sélection par raycasting. Par exemple, un effet grossissant effectué par une lentille permet d’agrandir des objets situés dans une zone spécifique, facilitant ainsi leurs sélections. De plus, l’effet grossissant ne dépend pas de l’environnement immédiat de l’objet ciblé, il est alors “particulièrement pertinent sur des populations denses de cibles” [41].

Ainsi, cette section est composée de deux parties. La première partie, Section 4.1, explore les techniques de facilitation de pointage qui ont été appliquées au raycasting. La seconde partie, Section 4.2, explore l’utilisation des lentilles dans les environnements virtuels et augmentés.

4.1 Facilitations de pointage pour le raycasting

Les techniques basées sur la métaphore du raycasting sont particulièrement impactées par trois facteurs : la taille de la cible, sa condition de visibilité (cible visible, cible cachée) et la densité de l’environnement. L’impact des deux premiers facteurs a été observé sur la technique *HMD_RC* dans les études expérimentales du Chapitre I. En effet, en l’absence d’amélioration par une aide au pointage, le raycasting standard est peu performant voire inutilisable pour :

- *La sélection de petits objets.* La sélection de petits objets demande une plus grande précision. Aussi, les tremblements de la main sont accentués au bout du rayon ce qui augmente le risque de manquer la cible. Cette difficulté croît significativement avec la distance séparant l’utilisatrice de l’objet à sélectionner.
- *La sélection d’objets cachés.* Le rayon est généralement infini s’il ne croise aucun obstacle. Lorsqu’il rencontre un ou plusieurs objets, seul le premier objet rencontré est sélectionné, rendant donc difficile voire impossible la sélection d’un objet caché par d’autres.

- *La sélection d'objets dans une zone densément peuplée.* La densité de l'environnement autour d'une cible peut être estimée soit par le nombre de distracteurs autour de la cible (par exemple [15]), soit par la distance entre la cible et les plus proches distracteurs (par exemple [163]). Plus les distracteurs sont proches ou nombreux autour d'une cible, plus le risque de sélectionner un distracteur plutôt que la cible augmente.

Le raycasting standard n'est donc pas adapté pour des sélections *précises* en *environnements denses* où certains objets peuvent être *occultés*. Cette section présente les aides à la sélection qui ont été appliquées au raycasting¹ standard pour l'améliorer et surmonter ces difficultés. Nous les classons en deux catégories : les techniques indépendantes des cibles (*target-agnostic*) et les techniques sensibles aux cibles (*target-aware*). Alors que les techniques sensibles aux cibles ont besoin de connaissance *a priori* des cibles pour fonctionner, les techniques dites indépendantes des cibles n'en ont pas la nécessité. Ces dernières ont ainsi l'avantage d'être plus génériques. Nous présentons tout d'abord les techniques indépendantes des cibles (Section 4.1.1) puis celles dites sensibles aux cibles (Section 4.1.2).

4.1.1 Techniques indépendantes des cibles

Les techniques de pointage basées sur le raycasting exploitent et agissent principalement sur trois caractéristiques du rayon pour améliorer les performances :

- La longueur du rayon (Section 4.1.1.1),
- La forme du rayon (Section 4.1.1.2),
- La courbure du rayon (Section 4.1.1.3).

Cette section organise les techniques de pointage par raycasting indépendantes des cibles, c'est-à-dire ne nécessitant pas de connaître les cibles *a priori*, selon la caractéristique du rayon qu'elles exploitent.

4.1.1.1 Longueur du rayon

Rayon de longueur finie : fixe ou ajustable manuellement

Plusieurs techniques simulent un rayon de longueur finie en plaçant un curseur sur le rayon, ce dernier servant à la sélection des objets. En déplaçant le curseur

1. Pour plus de détails sur les aides à la sélection 3D en général, nous référençons le livre *3D User Interfaces* [89] ainsi que trois taxonomies : celle de Bowman et al. [29, 30], celle de Poupyrev & Ichikawa [126] et enfin celle de Argelaguet & Andujar [5].

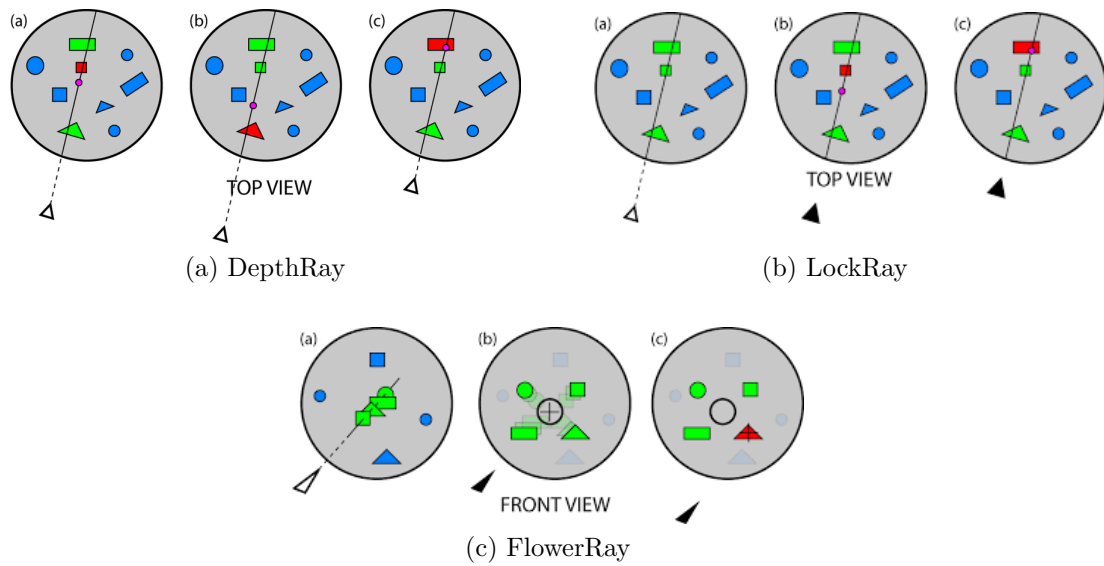


FIGURE 4.1 – Extensions du raycasting de Grossman et Balakrishnan [63] pour la sélection d'objets occultés. Images extraites de [63].

le long du rayon, l'utilisatrice contrôle la sélection en profondeur, ajustant ainsi la longueur "utile" du rayon.

Par exemple, afin de résoudre le problème de la sélection par raycasting d'objets cachés, Grossman et Balakrishnan [63] proposent la technique *DepthRay* implémentant un rayon de longueur finie et fixe, affiché dans un dispositif d'affichage volumétrique. Cette technique est illustrée à la Figure 4.1a. Le rayon virtuel est étendu au bout d'un *tracker* 3D physique (un *Ascension Flock-of-birds*) tenu par l'utilisatrice. Le curseur placé sur le rayon est représenté par une sphère rose. Tous les objets traversés par le rayon sont mis en évidence par un changement de couleur (ils passent en vert) mais seul l'objet le plus proche du curseur est sélectionné. Sa position étant fixe sur le rayon, le curseur est déplaçable par des mouvements de la main en avant ou en arrière. Ainsi, si une utilisatrice veut sélectionner le dernier objet intersecté par le rayon, elle déplace ce curseur en tendant le bras vers l'avant.

Par la suite, plusieurs travaux se sont appuyés sur le principe de la technique *DepthRay* et l'ont étudié en RA et RV [15, 136, 163]. Par exemple, Ro et al. [136] proposent un rayon dont la longueur est ajustable pour la sélection d'objets occultés en environnements 3D augmentés. Leur technique combine des lunettes de RA pour l'affichage de l'environnement virtuel et un téléphone pour le contrôle du rayon. Le rayon, de taille finie, prolonge le téléphone tenu par l'utilisatrice : un mouvement de la main change alors l'orientation du rayon. La longueur du rayon est ajustée par des interactions tactiles sur l'écran du téléphone (gestes de glissement sur l'écran).

Rayon de longueur infinie : présélection de tous les objets traversés par le rayon et désambiguïsation manuelle

D'autres techniques permettent de présélectionner plusieurs ou tous les objets intersectés par le rayon, y compris les objets cachés. Ensuite, grâce à une ou plusieurs étapes d'affinement, l'utilisatrice sélectionne manuellement l'objet de son choix parmi les objets présélectionnés [63, 113].

En suivant ce principe, Grossman et Balakrishnan [63] proposent deux extensions du raycasting pour la sélection d'objets cachés dans le cas de dispositifs d'affichage volumétrique 3D. Pour ces deux techniques, un rayon virtuel prolonge un *tracker* 3D physique tenu par l'utilisatrice.

- La technique *LockRay*, illustrée à la Figure 4.1b, repose sur une interaction en deux étapes bien *distinctes*. Tout d'abord, l'utilisatrice manie un rayon qui met en évidence (en vert) tous les objets traversés par ce rayon. Ensuite, elle bloque cet ensemble d'objets en appuyant sur un bouton placé sur le *tracker* 3D. Tant que ce bouton est enfoncé, et même si elle tremble, le rayon ne bouge plus : un curseur apparaît alors sur le rayon. Enfin, l'utilisatrice sélectionne l'objet de son choix de la même manière qu'avec *DepthRay* par de mouvements avant/ arrière de la main pour déplacer le curseur.
- La technique *FlowerRay*, illustrée à la Figure 4.1c, repose sur un réarrangement des objets traversés par le rayon dans un *marking menu* [87]. Tout d'abord, comme pour *DepthRay* et *LockRay*, l'utilisatrice manie un rayon qui met en évidence les objets traversés par le rayon. Ensuite, si l'utilisatrice appuie sur le bouton du *tracker* et le maintient enfoncé, les objets présélectionnés sont placés dans un menu. Enfin, l'utilisatrice sélectionne l'objet de son choix dans le menu en déplaçant un curseur 2D avec le *tracker*. L'objet sélectionné est l'objet le plus proche du curseur 2D lorsqu'elle relâche le bouton.

Une étude expérimentale a été menée pour comparer les deux techniques, *LockRay* et *FlowerRay*, avec la technique implémentant un curseur 3D standard et la technique *DepthRay* présentée précédemment. Pour cette étude, le curseur 3D standard est contrôlé par le *tracker* tenu par l'utilisatrice (mapping 1-1 entre les mouvements de l'appareil et les mouvements du curseur, et CD gain à 1). Les résultats de l'étude montrent que les techniques *DepthRay*, *LockRay* et *FlowerRay* réduisent le taux d'erreurs par rapport au curseur 3D. De plus, la technique *DepthRay* est la technique la plus rapide car elle permet, contrairement aux techniques *LockRay* et *FlowerRay*, de présélectionner un sous-ensemble d'éléments et de désambiguïser la cible simultanément.

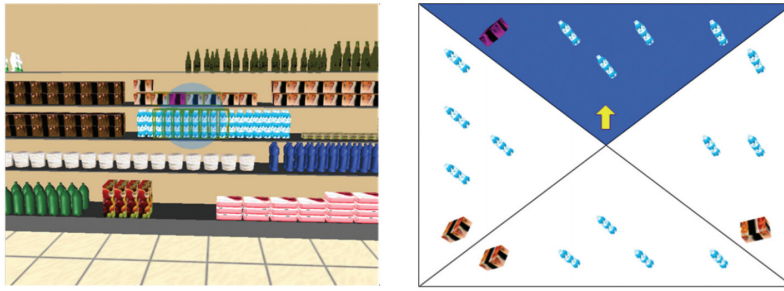


FIGURE 4.2 – *SQUAD* de Kopper et al. [84] et de Bacim et al. [7] : (1) présélection d'un ensemble d'objets par *sphere-casting* ; puis (2) affinement progressif en plaçant ces objets dans un menu en forme de quadrant. Images extraites de [84].

4.1.1.2 Forme du rayon : présélection volumétrique et désambiguïsation manuelle

Une autre approche pour faciliter la sélection par raycasting est d'utiliser un volume (un cône ou une sphère par exemple) plutôt qu'un rayon. Ainsi, au lieu de présélectionner les objets traversés par le rayon comme présenté à la Section 4.1.1.1, la présélection s'applique à tous les objets présents dans le volume [7, 39, 84, 107, 134]. L'utilisation d'un volume réduit ainsi le besoin de précision. Un mécanisme de désambiguïsation est ensuite nécessaire pour affiner cette présélection. Étant indépendantes des cibles, les techniques présentées dans cette section considèrent seulement des mécanismes de désambiguïsation *manuelle* (les mécanismes automatiques basés sur les positions des cibles étant présentés dans la Section 4.1.2.2).

Par exemple, la technique *SQUAD* de Kopper et al. [84] et de Bacim et al. [7] repose sur deux grandes phases. Tout d'abord, l'utilisatrice sélectionne un premier sous-ensemble d'objets à l'aide d'un volume de sélection sphérique (méthode appelée *sphere-casting*). Ensuite, ce sous-ensemble d'objets est réparti dans un menu en forme de quadrant (*quad-menu*), comme illustré à la Figure 4.2. La désambiguïsation consiste à sélectionner autant de fois que nécessaire la partie du menu contenant la cible jusqu'à ce qu'il ne reste plus qu'elle (affinement progressif).

Cashion et al. [39] comparent différentes extensions du raycasting utilisant une zone circulaire autour du pointeur pour présélectionner tous les objets présents dans cette zone :

1. La première est une technique implémentant un effet de grossissement visuel standard : (noté *zoom standard* par concision) : une utilisatrice choisit une zone à agrandir (c'est-à-dire la zone circulaire autour du pointeur) et cette zone zoomée est affichée sur un écran. Cette technique est efficace pour la sélection de petits objets. Cependant, le problème des objets cachés n'est pas

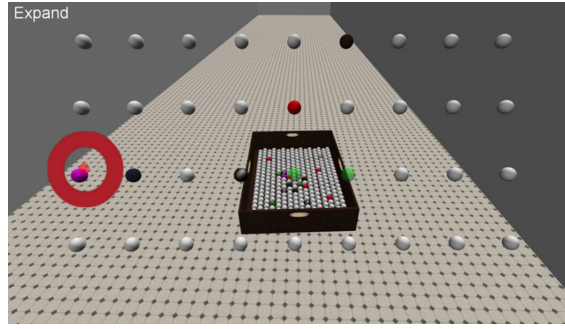


FIGURE 4.3 – *Expand* de Cashion et al. [39]. Image extraite de [39].

résolu.

2. La deuxième est la technique *SQUAD* [7, 84] (Figure 4.2).
3. La troisième technique, celle proposée par les auteurs, est appelée *Expand*. Cette technique est illustrée à la Figure 4.3. Contrairement à *SQUAD*, les objets sont répartis sur une grille semi-transparente, permettant à l'utilisatrice d'observer la scène initiale à travers la grille. Il n'y a donc pas de perte de contexte car il n'y a pas de changement de vue comme avec la technique de zoom standard ou la technique *SQUAD* (vue du placement initial puis vue dans le quadrant). De plus *Expand* ne comprend que deux étapes : l'affichage de la grille et la sélection dans la grille (contrairement à l'affinement progressif de *SQUAD* qui peut impliquer beaucoup d'étapes en environnements denses).

Ces extensions sont comparées au raycasting standard pour la sélection d'objets immobiles et en mouvement, et selon différents niveaux de densité. Leurs résultats montrent que le raycasting reste la meilleure alternative lorsque la scène est peu dense et que les objets sont immobiles. La technique de zoom requiert un temps similaire au raycasting mais réduit de moitié les erreurs de sélection. *SQUAD* est la plus rapide et précise en environnements dynamiques et en environnements peu denses mais lorsque la densité augmente, *Expand* surpasse *SQUAD* qui requiert beaucoup d'étapes allongeant considérablement le temps de sélection.

4.1.1.3 Courbure du rayon

Agir sur la courbure du rayon constitue une autre approche proposée par les techniques facilitant le pointage par raycasting. Par exemple, la technique *Sticky-Ray*² [151], illustrée à la Figure 4.4, permet de garder une cible active (c'est-à-dire sélectionnée) tant qu'une autre n'est pas rencontrée par le rayon. Un rayon flexible

2. À ne pas confondre avec la méthode d'expansion de cibles *sticky target* [176] (technique sensible aux cibles).

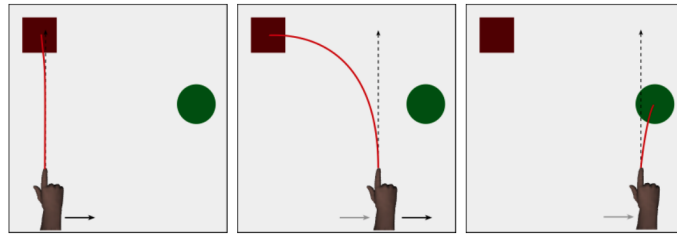


FIGURE 4.4 – *StickyRay* de Steinicke et al. [151]. Image extraite de [151].

est alors affiché pour que sa pointe reste “collée” à la cible tant qu’elle est active. Le *StickyRay* est ainsi une solution simple pour faciliter la sélection de petites cibles car il réduit le besoin de précision (même avec les tremblements de la main, le rayon reste collé à la cible). Dès qu’une nouvelle cible est sélectionnée, la pointe du rayon se déplace automatiquement sur celle-ci.

Un rayon flexible est également proposé par Olwal et Feiner [53] cette fois-ci pour gérer le problème de sélection des cibles occultées. Le rayon, dont la courbure et la taille sont calculées en fonction de l’orientation et de la position des mains de l’utilisatrice, peut donc contourner les obstacles pour sélectionner des objets cachés.

4.1.2 Techniques sensibles aux cibles

Cette section couvre les techniques de facilitation de pointage *sensibles aux cibles* utilisées pour le raycasting. Comme dans la section précédente (Section 4.1.1), nous classons les techniques selon la caractéristique du rayon qu’elles exploitent :

- La longueur du rayon (Section 4.1.2.1),
- La forme du rayon (Section 4.1.2.2),
- La courbure du rayon (Section 4.1.2.3).

4.1.2.1 Longueur du rayon : curseur étendu sur le rayon

Plusieurs techniques sensibles aux cibles ont été proposées pour faciliter la sélection 2D [41] et ont été appliquées à la sélection 3D. Une approche très étudiée consiste à agrandir la zone d’activation d’un curseur pour sélectionner l’objet le plus proche [62, 88, 166]. Par exemple, Vickers [166] utilise un *cube sensible* pour agrandir la zone d’activation d’un curseur 3D manipulé par une manette : lorsqu’un objet pénètre dans le cube, le curseur se déplace automatiquement sur cet objet. Une autre technique bien connue en 2D est le *Bubble Cursor* [62] dont la zone d’activation du curseur est représentée par un cercle autour du curseur. La taille de ce

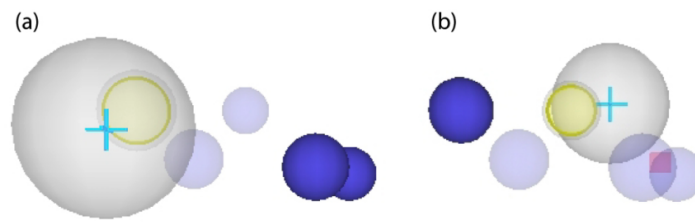


FIGURE 4.5 – *3DBubble* de Vanacken et al. [163]. La sphère grise représente la zone d'activation du curseur. (a) Un effet de transparence est appliqué à tous les objets proches du curseur. (b) La cible rouge occultée par des distracteurs devient ainsi visible à l'approche du curseur. Images extraites de [163].

cercle est ajustée dynamiquement afin de ne sélectionner qu'un seul objet à la fois. Une extension du *Bubble Cursor* pour les environnements virtuels 3D est proposée par Vanacken et al. [163, 164]. Leur technique, appelée *3D Bubble*, est illustrée à la Figure 4.5. *3D Bubble* utilise un curseur 3D déplacé selon la métaphore de la main virtuelle. À ce curseur est ajoutée une sphère représentant sa zone d'activation : cette sphère s'étend alors jusqu'à l'objet le plus proche du curseur et c'est cet objet qui est sélectionné. De plus, comme illustré à la Figure 4.5, toutes les cibles présentes à moins de 4 cm du curseur (par rapport au plan de coupe 2D vertical) deviennent semi-transparentes. Ceci permet de réduire l'occultation des objets proches du curseur. L'objet le plus proche du curseur est identifié par un bord opaque.

Comme nous l'avons vu dans la Section 4.1.1.1, d'autres techniques ajoutent un curseur sur le rayon, ce dernier permettant de contrôler la longueur effective du rayon. Ainsi, de nouvelles techniques étendent l'utilisation de curseurs étendus comme le *Bubble Cursor* au raycasting. Baloup et al. proposent *RayCursor* [13, 14, 15], une technique développée pour le casque de RV HTC Vive qui combine un rayon et un curseur étendu déplaçable de long de ce rayon. Cette technique est présentée à la Figure 4.6. Le rayon prolonge la manette de contrôle du casque tenue par l'utilisatrice : un curseur est déplacé le long du rayon par des mouvements avant-arrière du pouce sur le pavé tactile de la manette (interaction au toucher). L'objet le plus proche du curseur est automatiquement sélectionné. Le déplacement du curseur le long du rayon permet ainsi d'ajuster la longueur du rayon. Cette technique repose sur trois principaux éléments de conception : 1) un retour visuel pour identifier la sélection courante, 2) une fonction de gain pour le déplacement du curseur le long du rayon et, 3) un filtrage du rayon (fonction *1€ filter* [40]) pour réduire les tremblements de la main et ainsi stabiliser le rayon. Par la suite, cette fonction de filtrage est appliquée à toutes les techniques développées par les auteurs.

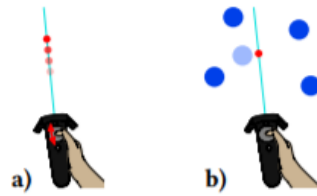


FIGURE 4.6 – *RayCursor* de Baloup et al. [15]. Un rayon virtuel étend la manette du HTC Vive. L'utilisatrice déplace un curseur le long du rayon grâce à des interactions au toucher sur le pavé tactile de la manette. Images extraites de [15].

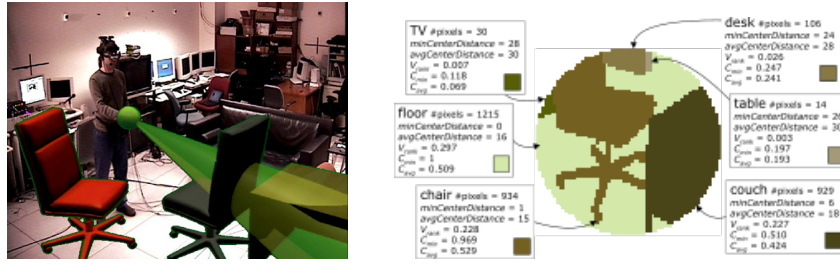
Une première étude menée par les auteurs comparent plusieurs retours visuels et montre qu'une simple mise en surbrillance de la cible la plus proche est le meilleur retour visuel car il encombre peu l'espace visuel, confirmant ainsi les résultats de Guillon et al. [65].

Baloup et al. comparent ensuite une technique de raycasting *RayCasting* (technique de référence) utilisant un simple rayon filtré (*1€ filter*), à trois extensions du raycasting :

- Une version du *RayCursor* avec déplacement manuel du curseur par des interactions au toucher sur le pavé tactile de la manette.
- Une version du *RayCursor* avec déplacement semi-automatique du curseur. Le *RayCursor* semi-automatique permet deux modes de déplacement du curseur : un mode automatique et un mode manuel. Le mode automatique, activé lorsque l'utilisatrice ne touche pas le pavé tactile, fonctionne comme le raycasting standard car le curseur saute directement à la première cible rencontrée par le rayon. L'utilisatrice passe en mode manuel en touchant le pavé tactile et déplace le curseur avec le pavé tactile (*RayCursor* manuel).
- Une technique inspirée de celle de Ro et al. [136] et implémentant un rayon de longueur ajustable (la technique de Ro et al. [136] est détaillée à la Section 4.1.1.1). La longueur du rayon est ajustée par des mouvements avant-arrière sur le pavé tactile de la manette.

Les deux versions de *RayCursor* reposent sur la proximité des cibles pour faciliter le pointage. Les deux autres techniques, le raycasting avec rayon ajustable et *RayCasting*, sélectionnent par contre le premier l'objet rencontré par le rayon.

Les auteurs comparent ces quatre techniques pour des tâches de sélection en faisant varier la densité de l'environnement et la taille des cibles. Dans leur étude, la densité de l'environnement est implémentée de la manière suivante : 30 distracteurs pour une faible densité et 60 distracteurs pour une forte densité sont répartis aléatoirement dans deux sphères de 60 cm de diamètre. La première sphère est placée à 1 m

FIGURE 4.7 – *SenseShapes* de Olwal et al. [117]. Images extraites de [117].

devant l'utilisatrice, la seconde est à 4 m devant elle. Les objets à sélectionner sont placés alternativement dans ces deux sphères.

Les résultats montrent tout d'abord que *RayCursor* avec déplacement manuel du curseur est plus rapide et permet de faire moins d'erreurs que la technique de rayon ajustable (Ro et al. [136]). La seule différence entre ces deux techniques est l'utilisation de la sélection de proximité avec *RayCursor*. Le fait que les performances de *RayCursor* surpassent celles du rayon ajustable montre donc l'avantage du principe de curseur étendu. En revanche ces deux techniques sont en moyenne plus lentes que la version semi-automatique de *RayCursor* et que *RayCasting*. En effet, les utilisateurs perdent beaucoup de temps à déplacer manuellement le curseur le long du rayon (*RayCursor* manuel) ou à ajuster la longueur du rayon (Ro et al. [136]). La version semi-automatique de *RayCursor* présente des temps de réalisation similaires à *RayCasting* mais permet de diminuer considérablement le nombre d'erreurs de sélection. Enfin, alors que la taille de la zone d'activation du curseur est impactée par la proximité des distracteurs, les auteurs trouvent peu voire aucun impact de la densité sur les performances des techniques. Seul *RayCasting*, plus lent en forte densité, est impacté et aucun impact n'est observé sur le taux d'erreurs. Le peu d'impact de cette variable est peut-être dû au choix d'implémentation de la densité qui ne prend pas en compte la distance entre la cible et ses plus proches distracteurs (contrairement au *density spacing* [163, 164] par exemple, voir Section 5.2.4).

4.1.2.2 Forme du rayon : présélection volumétrique

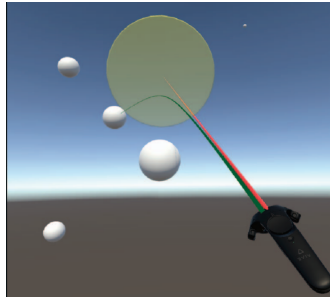
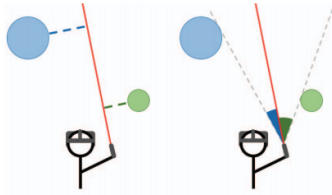
L'utilisation d'un volume de sélection (un cône, une sphère) plutôt qu'un rayon pour faciliter la sélection s'applique également aux techniques sensibles à la cible. Comme décrit dans la Section 4.1.1.2, ce type de sélection implique ensuite un mécanisme de désambiguïsation pour affiner cette sélection.

La deuxième étape de désambiguïsation étant souvent très impactée par le nombre d'éléments présélectionnés lors de la première étape (affinement progressif avec

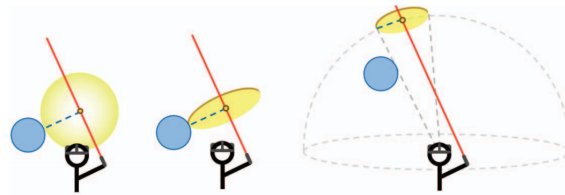
SQUAD [84] par exemple), Bacim et al. proposent la technique *Double Bubble* [8] qui utilise un *Bubble Cursor* 3D [62, 163] comme volume de sélection pour la première étape : la sphère (*bubble*) se redimensionne automatiquement jusqu'à l'objet le plus proche du curseur. Cependant, pour éviter le problème du *Bubble Cursor* en environnements denses (sphère très réduite), la taille de la sphère ne peut pas diminuer en dessous d'une taille prédéfinie. Si un seul objet est présent dans la sphère, celui-ci est automatiquement sélectionné. Dans le cas contraire, les objets sont placés dans un menu pour une phase de désambiguïsation. Le mécanisme de *Bubble Cursor* est alors à nouveau utilisé pour aider la sélection d'un objet dans ce menu.

Contrairement aux techniques avec désambiguïsation manuelle présentées à la Section 4.1.1.2, certaines techniques utilisent des approches heuristiques ou comportementales [158] pour prédire quelle cible, parmi celles présélectionnées par le volume, l'utilisatrice souhaite sélectionner. Pour les techniques basées sur un cône [56, 91, 92, 141, 150], la solution la plus utilisée est de considérer la distance (distance euclidienne ou angulaire) entre chaque objet contenu dans le cône et l'axe central du cône. L'objet le plus proche de cet axe est ainsi sélectionné. Cependant, bien d'autres mesures peuvent être considérées. Olwal et al. [117] présentent ainsi *SenseShapes* un ensemble de données statistiques permettant d'identifier automatiquement la cible souhaitée par l'utilisatrice (Figure 4.7). Pour cela, le système mémorise par un mécanisme d'historique tous les objets qui sont entrés dans le volume au moins une fois (le volume pouvant être un cube, une sphère, un cône ou un cylindre). Avec un cône par exemple, le système calcule et met à jour les données statistiques suivantes :

- *Le temps de présence des objets dans le cône.* Le temps passé dans le volume est calculé pour chacun des objets présents dans le cône depuis le début de l'expérimentation.
- *La distance entre le centre des objets et le centre du cône.*
- *La stabilité des objets dans le cône.* La stabilité d'un objet est le nombre de fois où l'objet entre et ressort du cône. Un objet avec beaucoup d'entrées et sorties sera considéré comme peu stable.
- *La visibilité des objets dans le cône.* Une image 2D en faible résolution, illustrée à la Figure 4.7-droite, est calculée et représente la projection 2D des objets présents dans le cône (comme la projection 2D du champ de vision d'une caméra qui serait attachée au sommet du cône). L'objet présentant le plus de pixels dans cette image est considéré comme étant l'objet le plus visible.
- *La distance entre le centre de la portion visible des objets dans l'image 2D et le centre de cette image.*

(a) *Bubble Ray*.

(b) Distance euclidienne (gauche) ou angulaire (droite) entre les objets et le rayon.



(c) Trois retours visuels étudiés.

FIGURE 4.8 – *Bubble Ray* de Lu et al. [97]. Images extraites de [97].

Un rang pour chacun des objets est calculé en considérant soit une seule de ces données statistiques, soit une combinaison de plusieurs ou de toutes ces données. Enfin, l'objet avec le rang le plus élevé parmi ceux présents dans le volume, est sélectionné. Une des principales limitations de ce type de calculs est qu'ils peuvent être très coûteux en environnements denses.

4.1.2.3 Courbure du rayon

La technique *Bubble Ray* [97] considère un rayon flexible dont la pointe est collée sur l'objet le plus proche du rayon (contrairement au *StickyRay* présenté à la Section 4.1.1.3 qui reste “collé” à la dernière cible intersectée). Cette technique est illustrée à la Figure 4.8a. La distance entre un objet et un rayon peut être calculée de deux manières différentes, comme illustré à la Figure 4.8b : soit en considérant la distance euclidienne, soit en considérant la distance angulaire. Deux variantes de *Bubble Ray* sont développées pour chacune de ces deux mesures de distance.

Inspirés du *Bubble Cursor* [62], les auteurs étudient également trois types de retours visuels (en plus du rayon flexible), illustrés à la Figure 4.8c, pour indiquer la sélection en cours d'un objet. Le premier retour visuel considère une sphère (Figure 4.8c-gauche) comme avec *3D Bubble* de Vanacken et al. [163]. Cependant, plus l'objet sélectionné est distant du rayon, plus la sphère est grande. Ceci peut fortement encombrer l'espace visuel. Les auteurs ont alors considéré une forme 2D telle

qu'un disque à la place de la sphère (Figure 4.8c-centre). Le disque est affiché entre le rayon et l'objet le plus proche. Ce second retour visuel est moins encombrant que la sphère. Cependant, lorsque l'objet sélectionné change, le disque saute d'un coup au niveau de la nouvelle cible : ces nombreux changements de profondeur du disque sont perçus comme gênants par plusieurs participants. Avec la troisième et dernière version, le disque n'est plus affiché au niveau de l'objet sélectionné mais est affiché sur une surface sphérique distante, comme illustré à la Figure 4.8c-droite. La profondeur du disque (distance entre l'utilisatrice et le disque) ne change donc plus.

Une première étude expérimentale pour une tâche de sélection simple en environnements épars montre de bonnes performances pour les deux variantes de *Bubble Ray*. Les performances de *Bubble Ray* avec la distance angulaire sont également légèrement meilleures que celles du *Bubble Ray* avec la distance euclidienne. Une seconde étude expérimentale évalue ces techniques en environnements denses. Leurs résultats montrent que les performances de *Bubble Ray* sont très impactées par la densité de l'environnement.

4.1.3 Bilan sur les facilitations de pointage existantes pour le raycasting

Une synthèse des techniques décrites dans la Section 4.1 est présentée dans le Tableau 4.1. Nous observons que les travaux existants proposent différentes solutions pour faciliter la sélection de petites cibles et de cibles cachées pour la sélection par raycasting. Cependant, il apparaît que ces facilitations de pointage restent généralement peu performantes en environnements denses. Par exemple, les techniques basées sur la sélection de l'objet le plus proche sont directement impactées par la proximité des distracteurs (la zone d'activation du curseur étant réduite en forte densité [97]). Les techniques d'affinement progressif (*SQUAD* [84] par exemple) sont également très impactées par la densité de l'environnement : le nombre d'étapes d'affinement étant dépendant du nombre d'éléments présents dans la scène.

Des techniques de zoom ont également été considérées pour faciliter la sélection de petites cibles. Ces techniques ont l'avantage de ne pas être impactées par la densité de l'environnement. Par exemple, une technique utilisant un zoom standard comme celle présentée par Cashion et al. [39] permet tout d'abord de sélectionner une zone d'intérêt dans la scène puis la caméra zoome sur cette zone. Les objets de cette zone sont grandis et ainsi plus facilement sélectionnables. Cette technique de zoom standard implique cependant la perte du contexte autour de la zone d'intérêt et ne résout pas non plus le problème de sélection des cibles cachées. Une autre

solution également proposée par Cashion et al. [39] est de réarranger les éléments présents dans la zone d'intérêt sous la forme d'une grille semi-transparente proche de l'utilisatrice : les objets sont agrandis pour faciliter leur sélection et il est toujours possible de voir la scène globale à travers la grille semi-transparente. Ce réarrangement des objets en grille permet aussi de rendre visibles des cibles qui étaient occultées dans la scène initiale. Cependant, contrairement au zoom standard, le réarrangement des objets dans une grille (ou dans un menu comme *FlowerRay* [63] par exemple) engendre une perte de la position relative des éléments par rapport aux autres. Dans de nombreux domaines de RA comme les cartes augmentées ou la chirurgie augmentée, cette perte de liens spatiaux entre les objets n'est pas acceptable car les liens spatiaux entre objets sont centraux à la tâche.

	Caractéristique exploitée du rayon	Techniques étendant le raycasting	Sélection des petites cibles facilitée	Sélection des cibles cachées facilitée	Robuste à la proximité des distracteurs	Robuste à un grand nombre de distracteurs
Techniques indépendantes des cibles	Longueur	<i>DepthRay</i> Grossman et Balakrishnan [63] <i>LockRay</i> Grossman et Balakrishnan [63] <i>FlowerRay</i> Grossman et Balakrishnan [63] Ro et al. [136]		✓ ✓ ✓ ✓	✓	✓
	Forme	<i>SQUAD</i> Kopper et al. [84] <i>Expand</i> Cashion et al. [39] Zoom standard comme présenté dans [39]	✓ ✓ ✓	✓ ✓ ✗	✓ ✓ ✓	✗ ✓ ✓
	Courbure	<i>StickyRay</i> Steinicke et al. [151] <i>Flexible Ray</i> Feiner et Olwal [53]	✓	✓		
Techniques sensibles aux cibles	Longueur	<i>RayCursor</i> Baloup et al. [15]	✓	✓		✓
	Forme	<i>Double Bubble</i> Bacim [8] <i>SenseShape</i> Olwal et al. [117]	✓ ✓	✓ ✓	✓	✓
	Courbure	<i>Bubble Ray</i> Lu et al. [97]	✓		✗	

Tableau 4.1 – Résumé des techniques d'extension du raycasting : améliorations permises (petites cibles, cibles cachées) et effet de la densité (proximité et nombre de distracteurs).

Une alternative aux techniques précédentes est l'utilisation d'un mécanisme de lentille grossissante avec sélection par raycasting directement sur la lentille. Une lentille grossissante a tout d'abord tous les avantages liés à l'utilisation d'un zoom qui sont mises en évidence dans le Tableau 4.1 (sélection de petites cibles facilitée, robuste à la densité). De plus, en utilisant une lentille déplaçable dans l'espace 3D de la scène, il serait possible d'atteindre et de visualiser des objets initialement cachés. Ce type de lentilles appliquées au raycasting pourraient ainsi être une bonne alternative pour la sélection précise d'objets 3D en environnements denses.

Tout d'abord et comme souligné dans le Tableau 4.1 :

1. L'effet de zoom est indépendant des cibles.
2. L'effet de zoom permet de faciliter la sélection de petites cibles.
3. L'effet de zoom n'est pas impacté par la densité de l'environnement.

De plus :

4. L'effet de zoom garde la position relative des objets ciblés.
5. Le mécanisme de désambiguïsation (sélection sur la lentille) est composé d'une seule étape (pas d'affinement progressif).
6. Le contexte global de l'environnement est gardé. Les techniques avec lentilles offrent simultanément une double représentation de la scène : l'utilisatrice a une vue zoomée de la zone d'intérêt sur la lentille, tout en gardant la scène globale visible (*focus+context* [46]).
7. Le déplacement 3D de la lentille permet de faciliter la sélection de cibles cachées.

Ainsi, la section suivante (Section 4.2) présente et analyse les techniques utilisant le mécanisme de lentilles dans des environnements augmentés et virtuels avant de présenter la technique conçue à la Section 5.

4.2 Les lentilles en RA et RV

Dès les années 90, Bier et al. [22] proposent d'utiliser une interface transparente, incorporant un filtre *lentille magique*, en environnements 2D. Bier et al présentent cet outil comme une surcouche permettant par exemple de “modifier la présentation des objets pour révéler des informations cachées, pour améliorer des données d'intérêt ou pour supprimer des informations superflues”[22]. Par la suite, ces propriétés font de la lentille un outil interactif de plus en plus employé en particulier dans le domaine de la visualisation d'information. En effet, ces dernières années, la visualisation d'information doit faire face à une augmentation importante de la taille des jeux de données à visualiser (*Big Data*). Les lentilles montrent alors tout leur intérêt car elles constituent un moyen simple et intuitif pour réduire l'encombrement des visualisations (c'est-à-dire réduire la quantité d'informations affichées), notamment en s'appuyant sur le principe de visualisation *focus + contexte* [46]. Suivant ce principe, les lentilles peuvent par exemple offrir une vue détaillée (*focus*) d'une zone spécifique à l'intérieur d'une vue globale (*contexte*) des informations entourant cette zone (“*displaying the focus within the context in a single continuous view*”[46]). Un examen détaillé des systèmes de visualisation basés sur des lentilles interactives est présenté dans les articles de Tominski et al. [158, 159].

Ainsi, les lentilles sont principalement utilisées en environnements 2D [57, 58, 82, 124, 137, 143] et pour des tâches de visualisation de données. Cependant, que ce soit en environnements de bureau ou sur des appareils mobiles, des lentilles *grossissantes* ont également été considérées pour la facilitation de pointage 2D (par exemple : *Fisheye* [66], *Shift* [168], *Pointing Lenses* [128], *Sigma Lenses* [123], *WidgetsLens* [2]). Mott et Wobbrock [115] présentent *Bubble Lens*, une amélioration du *Bubble Cursor* pour les environnements denses. Cette technique fonctionne comme le *Bubble Cursor* lorsque les cibles sont grandes et espacées. Mais quand les cibles sont petites et rapprochées, la technique zoome automatiquement autour du curseur afin de rendre les cibles plus visibles et plus facilement sélectionnables. Payne et al. [120] proposent *Expansion Cursor*, une technique 2D pour faciliter la sélection de petites cibles et basée sur une lentille à la demande. Avec *Expansion Cursor*, une utilisatrice a le choix entre cliquer directement sur la cible ou, grâce à un clic long, zoomer autour du curseur de la souris et ainsi sélectionner la cible agrandie.

Récemment, la création du domaine appelé *immersive analytics*³ [37, 52] a en-

3. Définition de *Immersive Analytics* extraite de <https://aviz.fr/~bbach/immersiveanalytics/> : “a topic that will explore the applicability and development of emerging user-interface technologies for creating more engaging experiences and seamless work-flows for data analysis applications”

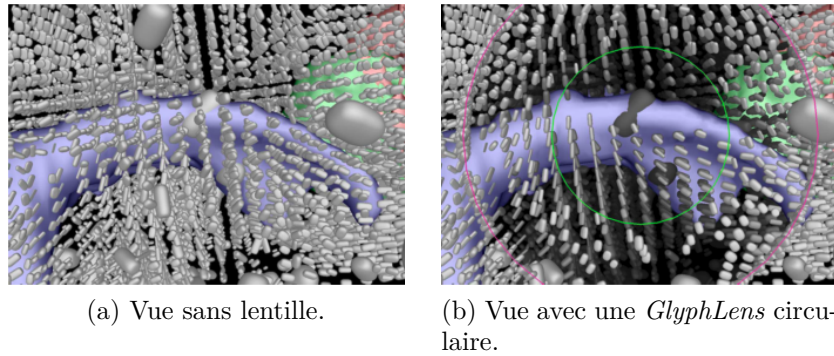


FIGURE 4.9 – *GlyphLens* de Tong et al. [160]. Images extraites de [160].

gendré de nombreuses recherches sur l'utilisation des technologies de RA et RV pour la visualisation d'information 3D. Le mécanisme de lentilles est alors beaucoup étudié dans des environnements 3D immersifs, et plus seulement pour la visualisation d'information en environnements 2D.

Les sections suivantes explorent la métaphore de la lentille appliquée à des environnements 3D immersifs. Nous passons en revue plusieurs études exploitant des lentilles, classées selon les tâches que les lentilles prennent en charge (Section 4.2.1), leur forme (Section 4.2.2) et la façon dont l'utilisatrice les déplace (Section 4.2.3).

4.2.1 Tâches d'interaction permises par les lentilles

Dans le domaine de RA [94, 108, 160] et de RV [129, 114], les lentilles permettent plusieurs tâches classées par Tominski et al. [159] en 7 catégories⁴ :

- *Explore*. La lentille permet l'exploration de l'espace d'information tout en montrant d'autres informations sur la zone couverte par la lentille. Par exemple, Bane et Höllerer [16] proposent une lentille pour avoir une vision à rayons X. Elle permet de voir la pièce à travers le mur.
- *Reconfigure*. La lentille permet de réarranger des éléments présents dans la zone couverte par la lentille. Par exemple, la *GlyphLens* [160], illustrée à la Figure 4.9, réduit l'occultation de glyphs en déplaçant les éléments occultants sur les bords de la lentille.
- *Encode*. La lentille permet de représenter différemment les données présentes sous la lentille. Un exemple est la lentille de Looser et al. [94] qui affiche la représentation fil-de-fer des objets présents sous la lentille.

4. Les sept catégories sont des catégories clés d'intentions des utilisateurs pour l'interaction, extraites de [178].



FIGURE 4.10 – Lentilles pour la visualisation et l’exploration de données [94]. Images extraites de [94].

- *Abstract/Elaborate*. La lentille permet d’abstraire ou de détailler les informations incluses dans la lentille. Par exemple une lentille grossissante peut permettre d’afficher plus d’informations : sur une carte n’affichant que les grands axes routiers, une lentille grossissante pourrait permettre de zoomer sur une zone et afficher également les rues (zoom sémantique⁵).
- *Filter*. La lentille permet de filtrer les données selon des critères : tout objet répondant à ces critères et présent dans la zone couverte par la lentille est affiché, les autres sont effacés. Par exemple, sur une carte de France, une lentille filtrant les grandes villes pourrait n’afficher que les villes de plus de 100 000 habitants.
- *Connect*. La lentille permet d’afficher les éléments liés à ceux présents sous la lentille. Par exemple, dans le cas de la visualisation du trafic aérien mondial, la sélection de certaines villes (celles présentes sous la lentille) entraîne l’affichage de toutes les villes reliées par avion à celles-ci. Ces dernières sont alors rapprochées autour de la lentille pour faciliter la visualisation des liaisons aériennes (technique Bring & Go de Moscovich [112]).
- *Select*. La lentille permet d’identifier un élément dans les données : tâche de pointage, de sélection, d’annotation.

Pour les six premières catégories, les lentilles sont principalement utilisées pour la visualisation et l’exploration de données (Figure 4.10) : elles permettent d’ajouter des informations supplémentaires, de supprimer des informations pour rendre la vue obtenue moins complexe ou d’afficher des vues alternatives [83, 94, 108], mais aussi de gérer l’occultation [22, 129, 160, 161] ou de zoomer dans une partie des données [94, 146, 148]. Si plusieurs lentilles sont utilisées en parallèle, une propriété intéressante est de pouvoir les combiner. Viega et al. [167] et plus récemment Kluge et al. en 2020 [83] proposent cette fonctionnalité : une utilisatrice peut superposer plusieurs lentilles et ainsi combiner leurs effets.

5. voir wiki InfoVis https://infovis-wiki.net/wiki/Semantic_Zoom

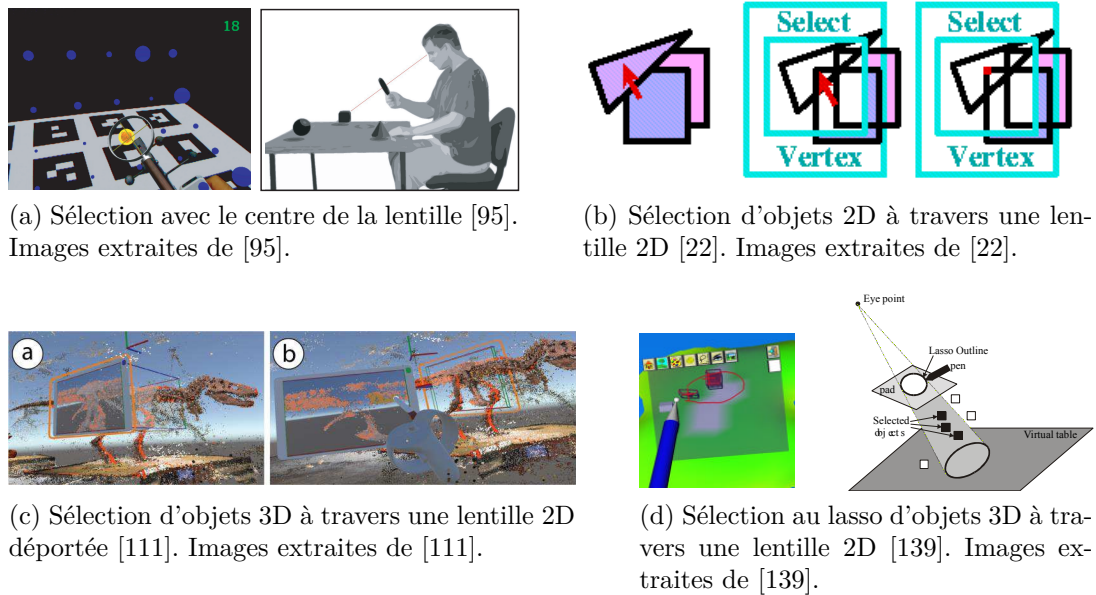


FIGURE 4.11 – Lentilles pour des tâches de sélection.

D'après la classification de Tominski et al. [159], outre l'exploration des données, les lentilles peuvent également être utilisées pour identifier des éléments dans des données (tâche *Select*, Figure 4.11). Les techniques de sélection avec lentille peuvent être réparties en deux groupes : celles basées sur une lentille passive, et celles basées sur une lentille active.

Sélection avec une lentille passive. La lentille passive ne sert qu'à viser l'objet ciblé : par exemple le centre de la lentille est utilisé comme pointeur. Looser et al. [94] proposent ce type de sélection en RA sur table. Une croix est affichée au centre dans la lentille et la métaphore du raycasting est implémentée pour sélectionner des objets : le rayon part de l'œil de l'utilisatrice et passe par le centre de la lentille [95] (technique présentée au Chapitre I, Section 1.3 et illustrée à la Figure 4.11a).

Sélection avec une lentille active – interaction sur la lentille. Une lentille active est un composant interactif sur lequel l'utilisatrice peut interagir. Par exemple, la sélection d'un objet peut être réalisée en deux étapes : (1) placer la lentille sur la zone d'intérêt et (2) sélectionner l'objet sur la lentille. Les techniques permettant de sélectionner des objets cachés [22, 111, 90] fonctionnent généralement de cette manière : la lentille est placée de manière à voir l'objet caché, puis la sélection est réalisée sur la lentille. En environnements 2D, Bier et al. [22] proposent une lentille qui affiche une vue en fil de fer des formes géométriques. À la Figure 4.11b, la lentille met en transparence l'intérieur des triangles pour pouvoir voir à travers.

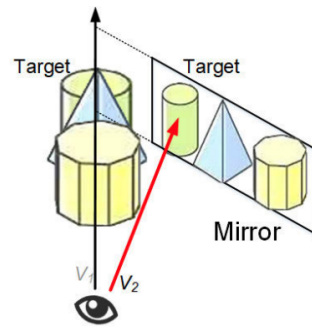


FIGURE 4.12 – Lentille avec effet miroir de Lee et Park [90]. Image extraite de [90].

L'utilisatrice peut ensuite cliquer sur le coin caché derrière les triangles.

La récente technique hybride 2D / 3D appelée *Slicing Volume* [111] est une technique de sélection multi-cibles en RV. La technique est illustrée à la Figure 4.11c. Une lentille 2D est initialement attachée à un cube pour présélectionner les éléments présents dans ce volume. Le cube est déplaçable en 3D afin d'atteindre des objets cachés. La lentille 2D est ensuite détachée du cube et déplacée à côté de la main de l'utilisatrice pour la phase de sélection. La sélection se fait avec un stylet sur la lentille déportée : un rayon est lancé dans le cube à partir du point d'intersection entre le stylet et la lentille. Tous les objets contenus dans le cube qui rencontrent ce rayon sont sélectionnés.

Contrairement aux deux lentilles précédentes qui éliminent les objets occultants, Lee et Park implémentent une lentille avec un effet miroir [90], comme illustré à la Figure 4.12. Placée dans l'espace 3D, cette lentille permet à l'utilisatrice de voir la cible cachée sous un autre point de vue, un point de vue à partir duquel elle n'est pas occultée. La sélection est alors faite sur le miroir.

Une autre possibilité de sélection proposée par Schmalstieg et al. [139] est la sélection par lasso. Une lentille tangible et transparente est tenue par l'utilisatrice. Si elle regarde la scène 3D à travers la lentille et dessine un cercle sur le plan 2D de cette lentille, tous les éléments visibles dans le cercle sont sélectionnés (Figure 4.11d).

Une dernière possibilité est la sélection “filet de pêche” également présentée par Schmalsteil et al. [139]. Avec cette technique, l'utilisatrice déplace la lentille tangible dans l'espace 3D de la scène. Tous les objets touchés par cette lentille sont alors “attrapés” par la lentille.

Des techniques pour des tâches de manipulation d'objets ont aussi été développées. Après avoir sélectionné un objet avec la lentille, Looser et al. [94] proposent différentes manipulations de l'objet ciblé comme la possibilité de le déplacer, de changer sa taille, de le dupliquer etc.

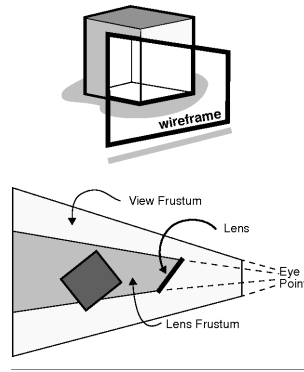


FIGURE 4.13 – *Lens frustum*. Illustration de la zone couverte par une lentille plate. Images extraites de [167].

4.2.2 Dimension des lentilles : plates ou volumétriques ?

Dans le contexte de la visualisation de données 3D en RA et RV, deux formes de lentilles [167] sont considérées : les *lentilles plates* et les *lentilles volumétriques*. Tel que défini par Schmalstieg et Höllerer [140], les lentilles volumétriques [27, 45, 114] appliquent leurs effets à chaque objet à l'intérieur de la région de la lentille et les lentilles plates [83, 95, 139, 152] appliquent leurs effets à chaque objet présent dans le champ de vision de la lentille magique. Cette zone, appelée également *lens frustum* par Viega et al. [167] est illustrée à la Figure 4.13.

Un des principaux avantages des lentilles volumétriques est de pouvoir extraire ou filtrer une partie de la scène 3D. Par exemple, *Slicing-Volume* [111] utilise un cube déplaçable afin d'extraire un sous-ensemble d'éléments et de n'interagir qu'avec cette partie des données. L'utilisatrice peut ajuster les dimensions du cube à l'aide d'un contrôleur tenu dans sa main non dominante. Cependant pour la tâche de sélection l'utilisatrice interagit sur l'une des faces de ce cube, elle passe donc d'une lentille volumétrique à une lentille plate. Comme le montre cette étude ainsi que d'autres travaux [94, 95, 139], l'utilisation d'une lentille plate est souvent privilégiée pour le pointage. Ceci permet de réduire la difficulté de la tâche de pointage, en passant d'une tâche de sélection dans l'espace 3D à une tâche de sélection dans l'espace 2D de la lentille.

Le fonctionnement de la lentille étant différent selon que la lentille est plate (2D) ou volumétrique (3D), ce choix est majeur dans la conception d'une technique.

Enfin, la forme des lentilles peut être prédéfinie mais également adaptative. De nombreuses études se sont concentrées sur des lentilles capables d'adapter leurs formes en fonction des données visualisées. Ce type de lentilles dédiées à la visualisation sort du cadre de notre travail sur la sélection d'objets 3D.

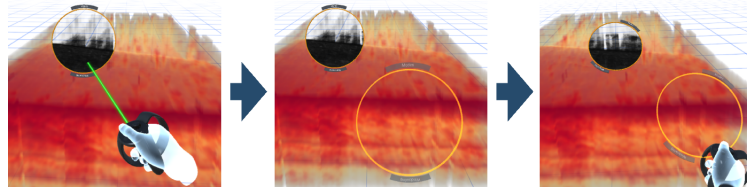


FIGURE 4.14 – Déplacement d'une lentille distante à l'aide d'un proxy à portée de main de l'utilisatrice [83]. Images extraites de [83].

4.2.3 Déplacement des lentilles

Dans la majorité des techniques basées sur des lentilles, celles-ci suivent la main ou l'accessoire tenu par l'utilisatrice (métaphore de la main virtuelle standard). Par exemple, Brown et al. [33] et Brown et Hua [32] utilisent un accessoire physique en forme de miroir 2D comme lentille. De la même manière, des marqueurs physiques (images imprimées souvent semblables à des QRcodes) sont suivis dans [94, 96, 108] pour afficher la lentille au niveau des marqueurs. Mota et al. [114], Kluge et al. [83] et Montano et al. [111] utilisent les manettes de RV pour déplacer la lentille dans la scène 3D. Des gestes à main levée peuvent également être utilisés pour activer et déplacer la lentille, la lentille suivant la position de la main de l'utilisatrice [129, 160].

Reipschlager et al. [132] implémentent une lentille virtuelle affichée au-dessus d'une surface 2D interactive. La lentille est seulement affichée dans le casque : la vue de la lentille est donc personnelle à l'utilisatrice. Cependant, le déplacement de la lentille est géré par des interactions au toucher directement sur la surface et la lentille suit le doigt de l'utilisatrice. Ces approches permettent de déplacer la lentille de manière intuitive. Néanmoins, le placement de la lentille dans une zone hors de portée devient difficile et fastidieux car il implique généralement un déplacement de l'utilisatrice.

Face à ce problème, Kluge et al. [83] développent un système de visualisation de données 3D en RV qui utilise des lentilles pouvant être déplacées à distance. La lentille suit la manette, comme vu précédemment, quand elle est à sa portée. Cependant, comme l'utilisatrice se déplace régulièrement pour explorer les données, certaines lentilles peuvent devenir hors de sa portée. Pour lui éviter de se déplacer (perte de temps et fatigue physique de déplacement) pour atteindre ces lentilles, Kluge et al. ajoutent la possibilité de déplacer la lentille à distance. Comme illustré à la Figure 4.14, les lentilles distantes sont sélectionnées par raycasting. Un *proxy* semi-transparent est ensuite affiché près de la main de l'utilisatrice. C'est ce proxy qui est manipulé par la métaphore de la main virtuelle pour déplacer la lentille distante. Tout déplacement ou rotation de ce *proxy* est appliqué à la lentille distante.

4.2.4 Bilan sur l'utilisation des lentilles en RA et RV

Dans cette section (Section 4.2), nous avons présenté les lentilles en environnements augmentés et virtuels classées selon leur fonction, leur forme et la manière dont l'utilisatrice les déplace. Après étude des lentilles existantes, nous faisons trois constats. Tout d'abord, ces outils sont principalement employés pour l'exploration de données et assez peu pour des interactions telles que le pointage et la sélection. De plus, lorsqu'elles permettent la sélection d'objets 3D, les lentilles sont généralement plates. Même dans l'étude de Montano et al. [111], bien qu'un cube 3D (*Slicing volume*) soit manipulé pour extraire et filtrer un sous-ensemble de données, c'est une lentille 2D qui est utilisée pour les tâches de sélection multi-cibles. En effet, l'usage de lentilles plates au lieu de lentilles volumétriques permet de simplifier la tâche : une sélection dans l'espace 3D est alors réduite à une sélection sur le plan 2D de la lentille. Enfin, nous constatons que les lentilles sont souvent tenues par l'utilisatrice pour interagir avec des objets à portée de main, très peu d'études les déplacent à distance.

En outre, nous avons vu dans la Section 4.1.3 qu'une technique de sélection avec une lentille grossissante pourrait être une bonne alternative pour faciliter la sélection d'objets distants en environnements denses : (1) il serait possible avec une telle technique de sélectionner des petites cibles et des cibles cachées ; (2) en utilisant le principe de zoom, le comportement de cette technique serait indépendant des cibles et de la densité de l'environnement.

Bien que prometteur, ce type d'extension du raycasting n'a jamais été étudié (Section 4.1) dans ce contexte, et donc plusieurs questions restent sans réponse : une extension du raycasting basée sur une lentille est-elle performante ? Le déplacement de la lentille dans l'espace 3D engendre-t-il une charge cognitive trop importante ? La section suivante (Section 5) est consacrée à ces questions.

5. Contribution : extension du raycasting avec une lentille grossissante

La section précédente a dressé un état de l'art des approches pour étendre le raycasting afin de faciliter la sélection à distance de petites cibles et de cibles cachées.

Comme expliqué dans la Section 4.1, nous constatons qu'aucune des extensions proposées n'est performante en environnements denses tout en préservant l'arrangement spatial des objets dans l'environnement.

Ce constat nous a conduit à considérer un mécanisme de lentille grossissante avec sélection par raycasting directement sur la lentille. Les lentilles grossissantes ont des propriétés intéressantes pour faciliter la sélection 3D d'objets distants, qu'ils soient petits ou cachés (pas d'impact de la densité, préservation du contexte autour de la cible et également des liens entre les objets). Pourtant, comme nous le soulignons dans la Section 4.2, aucune étude n'a considéré ce mécanisme comme extension du raycasting. La question est alors de savoir si une telle extension est performante et simple à utiliser.

Ainsi dans cette section, nous présentons une nouvelle technique d'extension du raycasting, appelée *RayLens*, qui combine une lentille grossissante et un rayon pour faciliter la sélection d'objets distants en environnements denses. Nous présentons également une étude expérimentale qui compare cette technique au raycasting standard, et à *RaySlider*, une autre extension du raycasting utilisant le mécanisme d'expansion de cibles. Cette étude expérimentale évalue à la fois l'effet de la densité de l'environnement et de la taille des cibles sur ces trois techniques. Les deux extensions du raycasting (*RayLens* et *RaySlider*) sont présentées en Section 5.1. Le protocole de l'étude ainsi que les résultats obtenus sont détaillés dans les Sections 5.2 à 5.4.

5.1 Éléments de conception : *RayLens* et *RaySlider*

Dans cette section, nous présentons les choix de conception des deux techniques d'extension du raycasting *RayLens* et *RaySlider*.

5.1.1 RayLens : une lentille grossissante couplée à un rayon virtuel

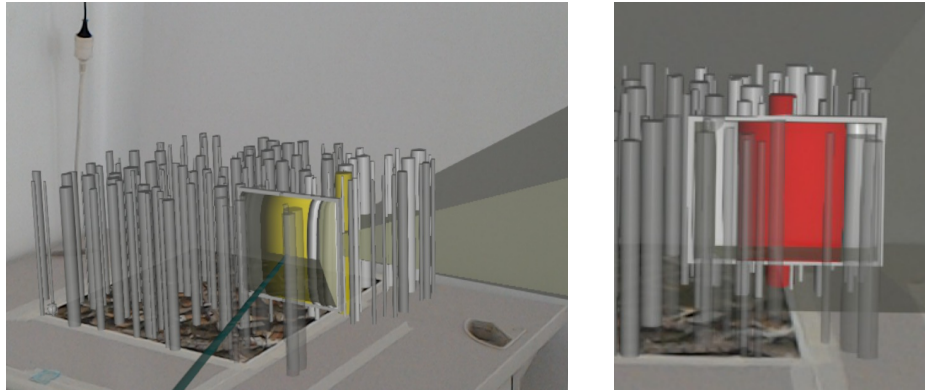
RayLens combine deux composants indépendants : un rayon et une lentille plate grossissante. Nous avons conçu *RayLens* comme une technique bi-manuelle : l'utilisatrice déplace le rayon en utilisant la main dominante et déplace la lentille en utilisant l'autre main. Ce choix de conception est conforme aux principes du modèle du contrôle bi-manuel humain de Guiard [64]. La technique bi-manuelle conçue est "asymétrique et dépendante" [78] : "les deux mains ont des rôles très différents à jouer qui dépendent l'une de l'autre" [78].

Le rayon virtuel prolonge un bâton physique tenu par la main dominante de l'utilisatrice. Pour sélectionner un objet 3D, l'utilisatrice vise avec le rayon la projection 2D de l'objet agrandi affichée sur la lentille distante, comme illustré à la Figure 5.1. *RayLens* est une extension de la technique de raycasting standard dont la lentille ne doit être utilisée qu'en cas de besoin. Néanmoins, afin de comparer *RayLens* avec la technique de raycasting standard dans notre étude (Section 5.2), un objet ne peut être sélectionné que sur la lentille : une sélection par raycasting standard n'est donc pas permise avec *RayLens* pendant l'étude expérimentale. Dans la lentille virtuelle, les projections 2D des objets intersectés par le rayon sont mises en surbrillance (coloration en jaune). Une sélection est validée en appuyant sur un bouton fixé sur le bâton.

La conception de la technique *RayLens* repose sur la combinaison de trois éléments de conception fondamentaux. Le premier est l'effet grossissant de la lentille pour faciliter la sélection des petits objets. Le second élément est le déplacement à distance de la lentille dans l'espace 3D pour faciliter la sélection des objets occultés. Le troisième élément de conception est l'effet de transparence appliqué aux objets devant la lentille pour éliminer l'occultation de la lentille.

Effet grossissant de la lentille virtuelle pour faciliter la sélection des petits objets

Le composant principal de *RayLens* est la lentille virtuelle. Elle est représentée dans la scène par un écran virtuel interactif, semblable à une tablette physique. Comme une tablette, elle fournit un second point de vue grâce à une caméra virtuelle attachée derrière. Ce point de vue supplémentaire, implémentant la métaphore de "l'œil dans la main" [153, 154], est toujours fixe par rapport à la lentille. Ainsi, la vue affichée n'est pas un prolongement de l'œil de l'utilisatrice. Le choix de conception sur la vue de la lentille a trois principaux avantages :



(a) Représentation du champ de vision de la lentille (*“lens frustum”* [167]) par un cône 3D et éclaircissement des objets présents dans ce cône.

(b) Vue zoomée des objets placés dans le champ de vision de la lentille.

FIGURE 5.1 – Illustration de la lentille grossissante et de son champ de vision.

- Il permet à l'utilisatrice d'avoir un autre point de vue que le sien. Elle peut donc visualiser sur la lentille des objets cachés de son point de vue.
- Il offre un point de vue indépendant de la position de l'utilisatrice. Ainsi, si la lentille ne bouge pas, les éléments affichés sur son écran sont stables quels que soient les mouvements de l'utilisatrice.
- Dans un contexte collaboratif, ce point de vue fixe permet à toutes les utilisatrices de partager la même vue sur la lentille.

La Figure 5.1a présente le champ de vision de la lentille (cône 3D affiché derrière la lentille) : tous les objets présents dans cette zone sont éclaircis pour que l'utilisatrice puisse facilement les identifier parmi tous les objets de la scène.

Afin de faciliter la sélection de petites cibles, la lentille agrandit visuellement la taille des objets présents dans son champ de vision. Comme illustré à la Figure 5.1b, un plan d'image virtuel (*virtual image-plane* [89]) contenant les projections 2D de ces objets agrandis est affiché sur la lentille.

Déplacement 3D de la lentille virtuelle pour faciliter la sélection des objets distants et cachés

L'utilisatrice déplace la lentille dans les trois dimensions à l'aide d'une télécommande tenue de la main non-dominante. Lorsqu'elle appuie sur le bouton de la télécommande, et tant que ce bouton n'est pas relâché, les mouvements de la télécommande sont suivis et traduits par un déplacement de la lentille. Si un seul déplacement ne suffit pas, pour atteindre les objets les plus éloignés par exemple, l'utilisatrice peut également faire des déplacements successifs (*clutching*).

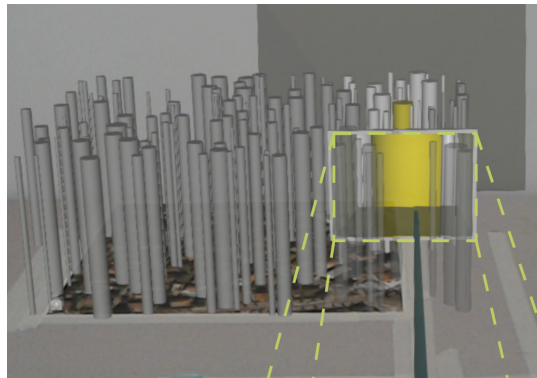


FIGURE 5.2 – Gestion de l’occultation de la lentille : mise en transparence des objets placés devant la lentille.

Le déplacement de la lentille dans l’espace est un choix de conception essentiel pour pouvoir visualiser et sélectionner des objets cachés. En effet, l’utilisatrice peut déplacer la lentille à travers les obstacles et visualiser des objets placés derrière. De plus, cela permet de placer la lentille à proximité de cibles distantes sans que l’utilisatrice ait besoin de se déplacer (contrairement à une tablette physique).

Effet de transparence pour réduire l’occultation de la lentille

Lorsque l’utilisatrice déplace la lentille à travers la scène 3D, des distracteurs peuvent apparaître devant elle. Cela entraîne une occultation partielle ou totale de la lentille, rendant donc difficiles voire impossibles les sélections sur la lentille. Pour réduire l’occultation visuelle de la lentille, tous les objets placés entre elle et l’utilisatrice sont rendus semi-transparents, comme illustré à la Figure 5.2. Sachant que l’utilisatrice est debout devant la table, nous approximons la zone de filtrage comme étant la zone devant la lentille, voir l’encadré jaune à la Figure 5.2. Les objets semi-transparents ne peuvent plus être sélectionnés par le rayon (le rayon les traverse). Avec cet effet de transparence appliqué devant la lentille, l’utilisatrice peut plus facilement voir et atteindre l’écran de la lentille avec le rayon.

En résumé, les choix de conception de *RayLens* sont basés sur les propriétés suivantes pour faciliter la sélection d’objets virtuels 3D :

- *Simplification de la tâche de sélection 3D.* La sélection d’un objet se fait en pointant sa projection 2D affichée sur la lentille et non en pointant l’objet 3D lui-même : la tâche de sélection 3D est donc réduite à une tâche de sélection 2D sur la lentille.

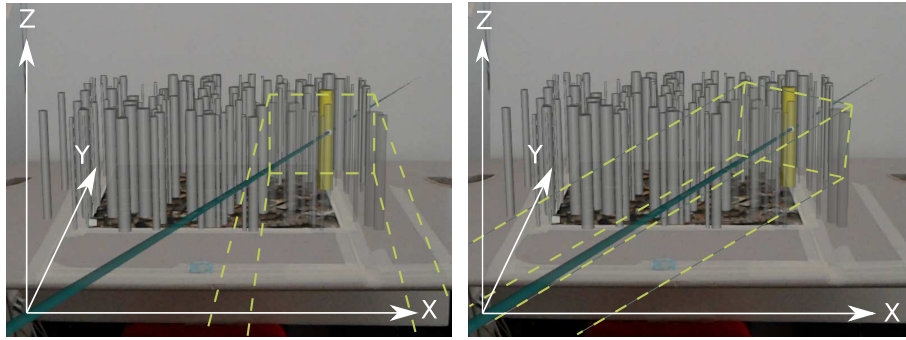


FIGURE 5.3 – Deux gestions différentes de l’occultation du curseur.

- *Précision.*
 1. L’effet de zoom facilite la sélection de petits objets.
 2. Sans mouvement volontaire de la lentille par l’utilisatrice, la position de la lentille reste fixe.
 3. Comme la sélection d’un objet se fait sur la lentille et non directement sur l’objet 3D, la distance entre l’utilisatrice et la cible est réduite. Le problème de la stabilité du rayon est par conséquent également réduit.
- *Occultation et densité.* Le mouvement de la lentille dans la scène 3D permet de sélectionner des cibles cachées en passant à travers des obstacles. La lentille agit comme un “plan de coupe” de la même manière que [57, 111]. En positionnant la lentille, l’utilisatrice peut ainsi effacer les distracteurs devant une cible pour augmenter sa visibilité.

5.1.2 RaySlider : un curseur déplaçable sur un rayon virtuel

RaySlider est une autre extension du raycasting qui partage les éléments de conception de *RayLens* tout en implémentant un mécanisme d’expansion de cibles. *RaySlider* utilise un curseur qui se déplace le long du rayon pour passer à travers les distracteurs. Comme *RayLens*, cette technique est bi-manuelle : l’utilisatrice déplace le rayon avec la main dominante et déplace le curseur le long du rayon avec l’autre main. En considérant un curseur attaché au rayon, les deux mains impactent la position du curseur :

- Un mouvement 3D du curseur est contrôlé par la main dominante via le mouvement du rayon,
- Un mouvement 1D du curseur le long du rayon est contrôlé par la main non dominante.

Comme la main dominante contrôle déjà la position 3D du curseur, l'utilisatrice n'a pas besoin de déplacer en plus le curseur sur le rayon lorsque la précision est requise. De plus, un contrôle simultané avec les deux mains pourrait "augmenter la charge sur le système moteur de l'utilisatrice" et la tâche pourrait "devenir plus difficile" comme observé dans [12]. Ainsi, avec une telle conception, la phase de précision devrait être effectuée par la main dominante qui tient le rayon, et le mouvement 1DOF du curseur le long du rayon devrait être principalement utilisé pendant la phase balistique. En conséquence, nous concevons *RaySlider* pour placer le curseur rapidement et non précisément dans la scène avec la main non-dominante. La main dominante contrôlant le rayon finit la tâche de pointage par un positionnement précis du curseur. Ce choix de conception pour l'interaction bi-manuelle est conforme au modèle du contrôle bi-manuel humain de Guiard [64] : en effet la main non-dominante est macro-métrique tandis que la main dominante est micro-métrique (résolution plus fine de la main dominante) et la main non-dominante intervient avant la main dominante.

Comme *RayLens*, le rayon virtuel prolonge le bâton physique tenu par la main dominante de l'utilisatrice. Pour obtenir un mouvement facile et rapide du curseur le long du rayon à l'aide de la main non-dominante, l'utilisatrice tient dans sa main non-dominante un curseur physique ou *slider* (pour plus de concision et de clarté, nous utiliserons le terme anglais *slider* dans la suite du manuscrit). Grâce à la mémoire musculaire, le slider physique tenu par la main non-dominante permet à l'utilisatrice de placer rapidement le curseur près de la cible.

Suivant le principe d'expansion de cibles pour faciliter la sélection d'objets, *RaySlider* met en surbrillance l'objet le plus proche du curseur. La sélection est ensuite validée en appuyant sur le bouton fixé sur le bâton. Enfin, le déplacement du curseur en profondeur (le long du rayon) permet de passer à travers des obstacles et d'atteindre des objets cachés. Un filtre de transparence, comme celui utilisé avec *RayLens*, est alors appliqué pour minimiser l'occultation du curseur.

Dans une étude pilote, nous avons comparé deux positions possibles pour ce filtre. Les deux positions sont illustrées à la Figure 5.3.

1. *Filtre appliqué entre le curseur et l'utilisatrice, Figure 5.3-gauche.* La zone de filtrage est parallèle à la surface de la table et à l'axe Y. La face arrière du parallélépipède est centrée autour du curseur.
2. *Filtre appliqué le long du rayon, entre la main de l'utilisatrice et le curseur, Figure 5.3-droite.* La zone de filtrage est parallèle à la surface de la table. Le parallélépipède pivote autour de l'axe Z pour qu'il suive l'orientation du rayon.

La face arrière du parallélépipède est centrée autour du curseur.

Dans un contexte de scènes de taille moyenne (scènes de 70 cm de large et 60 cm de profondeur), les participants n'ont observé aucune différence entre les deux positions. Une étude plus approfondie avec plusieurs tailles de scènes virtuelles serait nécessaire pour déterminer la meilleure option de conception selon la taille de la scène. Par la suite, et pour la taille de scène virtuelle considérée, nous avons choisi la première option (Figure 5.3-gauche) qui est similaire au filtre de *RayLens*.

5.2 Étude expérimentale : évaluation de l'efficacité de *RayLens* face au raycasting et à *RaySlider*

5.2.1 Objectif de l'étude et hypothèses

Cette étude expérimentale a pour objectif de comparer *RayLens* avec *RayCasting*, une implémentation de la technique de raycasting standard, et *RaySlider* pour étudier son efficacité (performances, retours utilisateurs) pour une tâche de sélection d'objets 3D virtuels. Nous évaluons également l'impact de la densité de l'environnement et de la taille des cibles sur les trois techniques étudiées.

Comme plusieurs études telles que [15, 39] proposant des améliorations de la métaphore du pointeur virtuel, nous choisissons une technique de raycasting standard comme base de référence. La technique *RayCasting* est donc implémentée comme un rayon qui sélectionne le premier objet intersecté. Sans collision, ce rayon est infini. Comme pour les deux autres techniques, le rayon prolonge le bâton physique tenu par l'utilisatrice, et la sélection est validée en appuyant sur le bouton placé sur ce bâton.

Pour cette étude comparative, nous formulons quatre hypothèses :

- *H1.* Les deux techniques utilisant un mécanisme de facilitation de pointage, *RayLens* & *RaySlider*, sont plus rapides et plus précises que la technique *RayCasting* pour sélectionner des petites cibles.
- *H2.* Si l'environnement autour de la cible est peu dense, *RaySlider* est la technique la plus rapide. En effet, lorsque la distance entre la cible et les distracteurs les plus proches est grande, la sélection automatique de l'objet le plus proche est très efficace.
- *H3.* Si l'environnement autour de la cible est dense, la technique *RayLens* est plus rapide et plus précise que la technique *RaySlider*. En effet, les techniques d'expansion de cibles comme *RaySlider* sont impactées par la densité

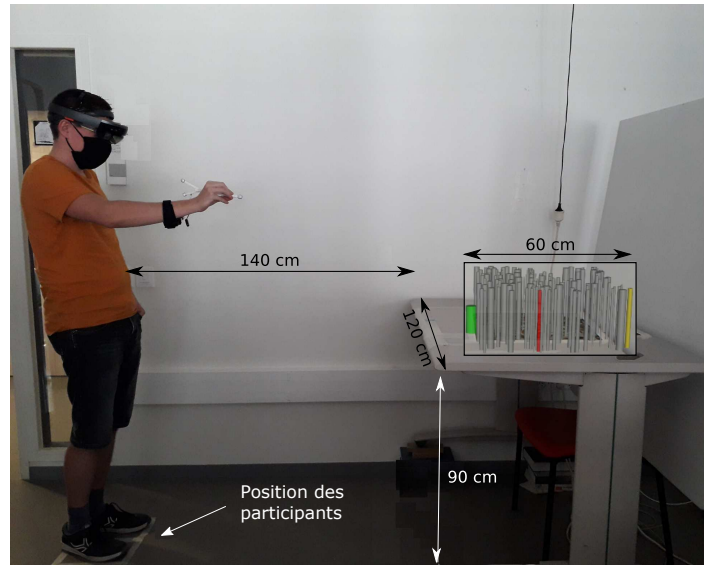


FIGURE 5.4 – Système expérimental.

de distracteurs [23]. Contrairement au mécanisme d’expansion de cibles, l’effet grossissant de la technique *RayLens* est indépendant de la densité de l’environnement. *RayLens* est donc peu impactée par la densité.

- *H4*. Les techniques *RayCasting* et *RayLens* nécessitent respectivement la charge de travail la plus faible et la plus élevée. Avec *RayLens*, basculer entre la vue 3D de la scène et la vue 2D affichée sur la lentille peut nécessiter un effort cognitif supplémentaire important de la part de l’utilisatrice.

5.2.2 Système expérimental

Le système expérimental de cette étude, présenté à la Figure 5.4, reprend le système de RA sur table décrit dans le Chapitre I, Section 2.2. Les participants sont debout à 1,40 m de la table physique (qui est non-interactive), centrés par rapport à la scène virtuelle. La position des participants est fixe et est indiquée par un marquage au sol (un rectangle de 30 cm × 40 cm au sol). Comme expliqué dans le Chapitre I, Section 2.2, l’image suivie par Vuforia est collée sur la table et permet de placer tous les objets au-dessus de la table.

Pour les trois techniques étudiées, le rayon virtuel prolonge un bâton physique suivi par les caméras OptiTrack. Comme illustré à la Figure 5.5, un bouton Arduino est fixé sur le bâton physique et permet à l’utilisatrice de valider une sélection.

Implémentation de la technique *RayLens*

La lentille virtuelle de *RayLens* mesure 22,5 cm × 17,5 cm, soit la taille standard

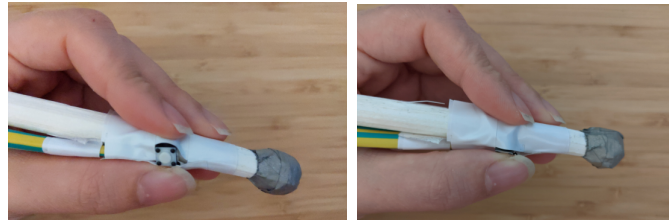


FIGURE 5.5 – Bouton Arduino placé sur le bâton physique et permettant de valider une sélection facilement avec le pouce.

d'une tablette physique en orientation paysage. Elle est déplacée dans l'espace 3D grâce au dispositif de clic sans fil, appelé *clicker*, de l'HoloLens. Le système HoloLens ne permettant pas de suivre la position du *clicker* en 3D, le suivi est réalisé par le système OptiTrack. Pour cela, le *clicker* est inséré dans un socle composé de quatre marqueurs réfléchissants (Figure 5.6a). Ce socle a été spécialement conçu pour cette étude et imprimé avec une imprimante 3D.



(a) RayLens : Socle + marqueurs pour le suivi du clicker de l'HoloLens par le système OptiTrack

(b) RaySlider

FIGURE 5.6 – RaySlider : slider utilisé pour déplacer le curseur le long du rayon

Lorsqu'un clic est détecté sur le *clicker*, celui-ci se comporte comme une télécommande 3D : son mouvement est traduit par un mouvement de la lentille tant que le bouton est maintenu appuyé. Le gain pour le déplacement de la lentille est de 1 :3, pour un déplacement du *clicker* de 1 cm, la lentille se déplace de 3 cm dans la même direction. Avec un tel ratio, un mouvement de 30 cm du *clicker* permet à l'utilisatrice de déplacer la lentille de sa position initiale (30 cm devant la scène) à la position des objets les plus éloignés (90 cm de la position initiale de la lentille).

L'implémentation actuelle de *RayLens* ne permet pas de tourner la lentille. L'étude expérimentale est donc conçue pour que sa rotation ne soit pas nécessaire. Nous rappelons que les participants sont debout devant la table, centrés par rapport à la scène virtuelle. Pour visualiser les objets les plus au bord de la scène, la lentille



FIGURE 5.7 – Le boîtier, contenant le nécessaire pour le fonctionnement du bouton du bâton, est attaché au poignet d’une utilisatrice.

sera déplacée au maximum de 20 degrés vers la gauche ou la droite par rapport à la position fixe des participants.

Implémentation de la technique *RaySlider*

Le curseur virtuel est déplacé le long du rayon grâce à un slider tangible tenu par l’utilisatrice. Il s’agit d’un potentiomètre de 7,3 cm de long, comme illustré à la Figure 5.6b. La valeur minimale du slider (sa position basse) correspond à un curseur virtuel placé à l’extrémité du bâton (proche de l’utilisatrice). La valeur maximale du slider (sa position haute) correspond à un curseur virtuel à 3 mètres de la pointe du bâton sur le rayon. Les cibles les plus éloignées de cette étude étant à 2 mètres de l’utilisatrice, la valeur minimale et maximale du slider implique donc de n’utiliser que les 2/3 bas du slider, permettant ainsi un usage confortable de cet outil.

Les événements produits par le bouton du bâton et le slider sont envoyés via une communication Bluetooth à un récepteur connecté à l’HoloLens. Pour cela, le bouton et le slider sont chacun connectés à une radio NRF24L01 enfin d’envoyer leur état (appuyé/relâché, valeur du slider) à une troisième radio qui réceptionne les événements pour l’HoloLens. Pour éviter des fils qui pourraient gêner l’utilisatrice durant la réalisation de la tâche, les deux cartes Arduino émettrices reliées au bouton et au slider sont alimentées grâce à une batterie fixée au poignet. La Figure 5.7 présente le boîtier du bouton Arduino qui a été créé pour contenir la carte Arduino, la radio, la batterie et également masquer les fils. Ce boîtier est attaché par une bande scratch au poignet de l’utilisatrice.

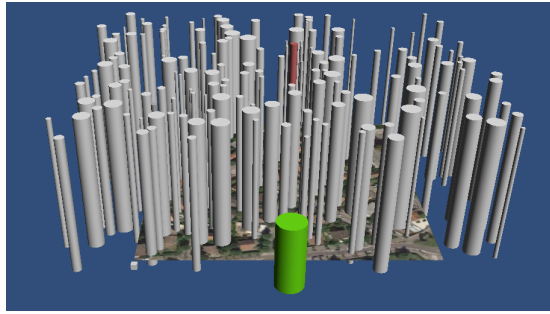


FIGURE 5.8 – Exemple d’une scène 3D créée pour l’étude expérimentale.

5.2.3 Tâche

Dans cette étude expérimentale, les participants doivent sélectionner des cibles cylindriques dans des scènes virtuelles 3D placées sur une table. Chaque scène 3D est composée de deux cibles à sélectionner (une cible initiale et une cible finale) et de 120 distracteurs. Tous les distracteurs sont gris, la cible initiale est verte et la cible finale est rouge. Un exemple de scène est présenté à la Figure 5.8. Lorsqu’un cylindre est pointé par le rayon, il passe du gris au jaune. Quelle que soit la scène, la cible initiale est toujours placée au même endroit, face à l’utilisatrice et devant tous les autres cylindres. La cible finale est positionnée au hasard mais toujours à une distance fixe de 40 cm de la cible initiale. Avant chaque essai avec *RayLens*, la lentille est positionnée devant la cible initiale, à 30 cm des premiers distracteurs. Tous les distracteurs sont positionnés aléatoirement sous les contraintes suivantes : ils ne doivent ni se chevaucher ni occulter complètement la cible. De plus, leur diamètre est choisi aléatoirement entre 0,5 cm et 3,5 cm. Leur hauteur est de 24 cm.

Chaque tâche de sélection comprend les étapes suivantes :

1. L’utilisatrice se place à la position indiquée par le marquage au sol (Figure 5.4), à 1,40 m de la table. Elle restera à cette place tout au long de l’étude expérimentale.
2. L’utilisatrice doit tout d’abord sélectionner la cible initiale verte. La validation de la sélection est faite en appuyant sur le bouton du bâton. Une fois sélectionnée, la cible initiale disparaît. Le temps de réalisation de la tâche est décompté à partir de la validation de la cible initiale.
3. L’utilisatrice doit ensuite sélectionner la cible finale rouge le plus rapidement possible tout en minimisant le nombre d’erreurs de sélection. Le temps de réalisation de la tâche est décompté jusqu’à la validation de la cible finale.

5.2.4 Protocole

L'étude expérimentale est conçue comme une étude intra-sujet, avec les trois variables indépendantes (TECH, SIZE, DS) suivantes :

- *La technique d'interaction utilisée (TECH) : RayLens, RaySlider, RayCasting.*
- *La taille de la cible (SIZE) :* la taille correspond au diamètre du cylindre. Le diamètre de la cible est 0,5 cm, 1,5 cm ou 3,5 cm. Les cibles mesurent toujours 24 cm de haut (hauteur identique à celle des distracteurs).
- *L'espacement de densité (DS) :* la densité est représentée par la distance entre la cible finale et les distracteurs les plus proches (*density spacing* [163]). Cette distance est de 1 cm (forte densité) ou de 5 cm (faible densité). Dans toutes les scènes, quatre distracteurs cylindriques sont positionnés autour de la cible finale. Deux sont positionnés sur une ligne définie par la cible finale et la position fixe des participants : un devant et un derrière la cible ; les deux autres sont positionnés sur une ligne perpendiculaire, un de chaque côté de la cible finale. Afin de réduire l'occultation visuelle de la cible, ces quatre distracteurs ont un petit diamètre (0,5 cm). De plus, nous faisons pivoter ces quatre distracteurs autour de la cible pour que les participants puissent la voir facilement (rotation de 45 degrés).

Pour garantir des conditions expérimentales similaires entre les participants, un ensemble de scènes 3D (c'est-à-dire la cible initiale, les cibles finales et les distracteurs) ont été calculées avant l'étude et sont partagées entre tous les participants et toutes les techniques. L'ordre de présentation des techniques a été contrebalancé par un carré latin et les scènes sont présentées dans un ordre aléatoire pour chacune des techniques.

L'étude expérimentale commence par une brève introduction. Avant chaque technique, les participants effectuent une phase d'entraînement pour expérimenter chacune des 6 combinaisons de $DS \times SIZE$.

Pour observer un potentiel effet d'apprentissage, les essais sont groupés en trois blocs par technique. Dans chaque bloc, les 6 combinaisons de $DS \times SIZE$ ont été répétées 3 fois dans un ordre aléatoire pour un total de $3 SIZE \times 2 DS \times 3$ répétitions = 18 essais par bloc. Les participants réalisent donc 54 essais par technique. Ils peuvent prendre une pause à la fin de chaque technique. Les participants remplissent également un questionnaire après chaque technique (Annexe F) et un questionnaire de synthèse à la fin de l'étude (Annexe G) pour comparer les trois techniques. L'étude se termine par un entretien. En particulier, les participants sont interrogés sur leur

utilisation de *RayLens* et de *RaySlider* (positionnement de la lentille et du curseur). L'étude a duré environ 75 minutes par participant.

Nous avons recruté 12 participants, âgés de 22 à 38 ans (moyenne = $28,7 \pm 4,23$). Tous les participants étaient droitiers. Aucun d'entre eux n'était un expert en RA. L'étude s'étant déroulée pendant l'épidémie de COVID-19, nous avons appliqué des mesures sanitaires préventives (masque, gel hydroalcoolique).

Au regard de nos hypothèses, nous considérons deux métriques objectives (temps de réalisation, nombre d'erreurs) et deux métriques subjectives (charge de travail, préférence des utilisateurs). Le temps de réalisation de la tâche est le temps écoulé entre la validation de la sélection de la cible initiale et la validation de la sélection (avec succès) de la cible finale. Le nombre d'erreurs est le nombre de validations erronées (c'est-à-dire lorsqu'une utilisatrice appuie sur le bouton alors qu'aucun objet n'est sélectionné ou qu'un distracteur est sélectionné) avant la sélection de la cible finale.

Pour les métriques subjectives, nous avons conçu un premier questionnaire (Annexe F) pour chaque technique. Ce questionnaire combine des questions pour évaluer :

- La charge de travail perçue par l'utilisatrice à partir d'une version simplifiée du questionnaire NASA-TLX, appelée Raw TLX ou RTLX. Selon l'enquête de Hart [69], les deux versions (NASA-TLX et RTLX) sont applicables ;
- L'utilité du filtre de transparence (*RaySlider* et *RayLens*) mesurée sur une échelle de 1 à 7 points ;
- L'utilité du retour visuel (cône + éclaircissement des objets présents dans le champ de vision) représentant le champ de vision de la lentille (*RayLens*) mesurée sur une échelle de 1 à 7 points.

Le questionnaire final (Annexe G) permet de déterminer la technique la moins exigeante mentalement et physiquement, la technique la plus performante, la plus précise et la plus rapide. Enfin, les participants classent les techniques par ordre de préférence.

5.3 Étude expérimentale : résultats

Nous présentons dans cette section les résultats quantitatifs et qualitatifs obtenus. La normalité des données est testée à l'aide du test Shapiro-Wilk. Pour les données normalement distribuées (temps de réalisation) nous utilisons un test sta-

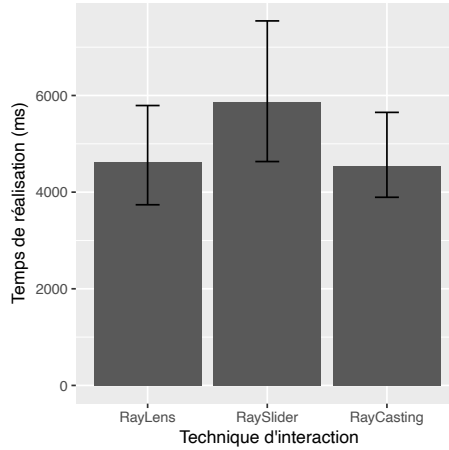


FIGURE 5.9 – Temps moyen de réalisation des tâches de sélection pour chacune des techniques.

tistique par analyse de variance (ANOVA) et des tests de Student (ou tests t) avec ajustement de Bonferroni pour les comparaisons par paires. Les données qui ne suivent pas une loi normale (nombre d'erreurs) sont pré-traitées par une transformation de rang aligné (ART) [175] avant d'utiliser une ANOVA avec une correction de Bonferroni pour les comparaisons par paires. Dans les graphiques suivants, nous représentons les valeurs moyennes et les intervalles de confiance (IC) à 95% calculés avec la méthode bootstrap Bca (la fonction du bootstrap étant la moyenne) [51].

5.3.1 Résultats quantitatifs

Temps de réalisation moyens

Nous n'avons pas observé d'effet statistiquement significatif de la technique TECH [$F_{2,22} = 2,51$, $p = 0,1$] sur le temps de réalisation de la tâche (Figure 5.9). Le temps moyen de réalisation de la tâche de sélection est de 5,9 s pour *RaySlider*, 4,6 s pour *RayLens* et 4,5 s pour *RayCasting*. Aussi, en comparant les performances des participants entre les 3 blocs, nous n'observons pas d'effet d'apprentissage ou de fatigue.

Nous observons un effet statistiquement significatif de la taille de la cible SIZE [$F_{2,22} = 103$, $p < 0,0001$, $\eta_G^2 = 0,31$] et de la densité DS [$F_{1,11} = 94,8$, $p < 0,0001$, $\eta_G^2 = 0,12$] sur le temps de réalisation.

Effet d'interaction SIZE \times TECH. Nous observons un effet d'interaction statistiquement significatif entre SIZE et TECH [$F_{4,44} = 26,3$, $p < 0,0001$, $\eta_G^2 = 0,15$]. Les résultats sont présentés à la Figure 5.10. Pour *RayCasting*, les résultats montrent

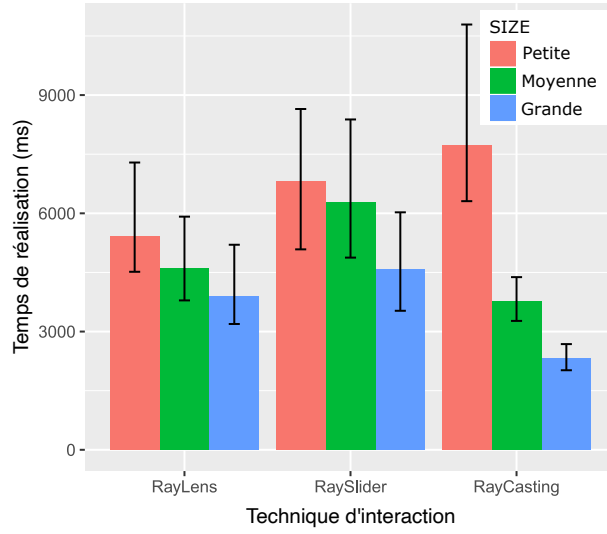


FIGURE 5.10 – Effet d’interaction entre la taille de la cible et la technique sur le temps moyen de réalisation des tâches de sélection.

de grandes variations de temps de réalisation selon les tailles ($RC_{petite} = 7,6$ s, $RC_{moyenne} = 3,7$ s, $RC_{grande} = 2,3$ s) : la technique est très rapide pour sélectionner de grandes cibles mais devient très lente avec des petites cibles. Les comparaisons par paires montrent des différences statistiquement significatives entre les petites et grandes cibles ($p < 0,0001$) et entre les moyennes et grandes cibles ($p = 0,004$). Les techniques *RaySlider* ($RS_{petite} = 6,8$ s, $RS_{moyenne} = 6,4$ s, $RS_{grande} = 4,5$ s) et *RayLens* ($RL_{petite} = 5,5$ s, $RL_{moyenne} = 4,6$ s, $RL_{grande} = 3,9$ s) sont légèrement impactées par la taille des cibles. Une différence statistiquement significative est observée seulement entre les petites et les grandes cibles ($p = 0,02$ pour *RaySlider*, $p = 0,035$ pour *RayLens*).

Pour les grandes cibles, les comparaisons par paires montrent des différences statistiquement significatives entre les techniques. *RayCasting* est la technique de sélection la plus rapide ($p = 0,005$ pour *RayCasting-RayLens* ; $p = 0,0003$ pour *RayCasting-RaySlider*). Aucune différence statistiquement significative n’est observée entre *RayLens* et *RaySlider*.

Pour les cibles moyennes, *RayCasting* est significativement plus rapide que *RaySlider* ($p = 0,008$). Aucune différence statistiquement significative n’est observée entre *RayCasting* et *RayLens*, ni entre *RayLens* et *RaySlider*.

Pour les petites cibles, *RayLens* semble être la technique de sélection la plus rapide ($1,38\times$ plus rapide que *RayCasting* et $1,24\times$ plus rapide que *RaySlider*). Cependant, une différence statistiquement significative est observée seulement entre *RayLens* et *RayCasting* ($p = 0,038$).

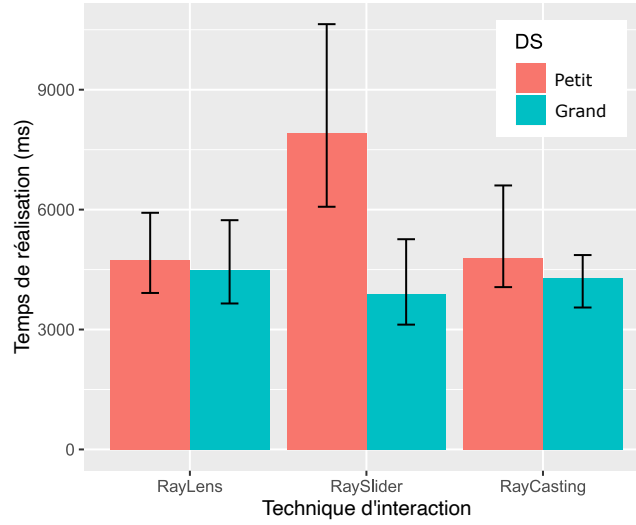


FIGURE 5.11 – Effet d’interaction entre l’espace de densité et la technique sur le temps moyen de réalisation des tâches de sélection.

Effet d’interaction DS \times TECH. Nous observons un effet d’interaction statistiquement significatif entre DS et TECH [$F_{2,22} = 35,4$, $p < 0,0001$, $\eta_G^2 = 0,13$]. Les résultats sont présentés à la Figure 5.11. Les comparaisons par paires montrent un effet statistiquement significatif de DS sur les temps de *RaySlider* ($p = 0,0008$, $RS_{DS_{\text{petit}}} = 7,9$ s, $RS_{DS_{\text{grand}}} = 3,9$ s). Aucun effet statistiquement significatif de DS n’est observé pour les techniques *RayCasting* ($RC_{DS_{\text{petit}}} = 4,8$ s, $RC_{DS_{\text{grand}}} = 4,3$ s) et *RayLens* ($RL_{DS_{\text{petit}}} = 4,7$ s, $RL_{DS_{\text{grand}}} = 4,5$ s).

Lorsque la densité est faible (DS grand), les comparaisons par paires ne montrent aucune différence statistiquement significative entre les techniques. En forte densité (DS petit), *RaySlider* semble être la technique la plus lente (*RayLens* et *RayCasting* sont $1,6\times$ plus rapides que *RaySlider* en moyenne). Cependant, aucune différence statistiquement significative n’est observée entre *RaySlider* et *RayLens* ($p = 0,06$), ni entre *RaySlider* et *RayCasting* ($p = 0,08$).

Pour résumer, l’étude montre que le temps de réalisation de la tâche de sélection avec la technique *RayCasting* est fortement impacté par la taille de la cible. Le temps de réalisation avec la technique *RaySlider* est également fortement impacté par la densité. Enfin, pour la technique *RayLens*, l’étude montre peu d’impact de ces deux variables sur le temps de réalisation de la sélection.

Précision de la tâche de sélection

Le nombre d’erreurs est le nombre de fois où les participants appuient sur le bouton de validation avant la validation avec succès de la cible. Nous observons un

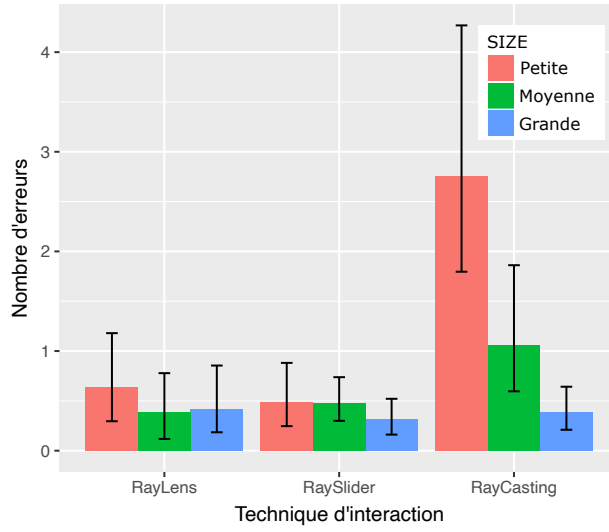


FIGURE 5.12 – Effet d’interaction entre la taille de la cible et la technique sur le nombre moyen d’erreurs.

effet statistiquement significatif de la technique TECH [$F_{2,187} = 46,12$, $p < 0,0001$], de la taille de la cible SIZE [$F_{2,187} = 33,51$, $p < 0,0001$] et de la densité DS [$F_{1,187} = 14,61$, $p = 0,0002$] sur le nombre d’erreurs.

Le nombre d’erreurs moyen est de 0,43 erreur pour *RaySlider*, 0,48 erreur pour *RayLens*, et 1,35 erreur pour *RayCasting*. Les comparaisons par paires montrent que la technique *RayCasting* est statistiquement moins précise que *RaySlider* ($3,1\times$ plus d’erreurs, $p = 0,004$) et que *RayLens* ($2,8\times$ plus d’erreurs, $p = 0,001$). Aucune différence statistiquement significative n’est observée entre *RaySlider* et *RayLens*.

Effet d’interaction SIZE \times TECH. Nous observons un effet d’interaction statistiquement significatif entre SIZE et TECH [$F_{4,187} = 23,71$, $p < 0,0001$]. Les résultats obtenus sont présentés à la Figure 5.12. Pour *RayCasting*, les résultats montrent de grandes variations du nombre d’erreurs selon la taille de la cible ($RC_{petite} = 2,7$ erreurs, $RC_{moyenne} = 1,05$ erreur, $RC_{grande} = 0,38$ erreur) : le nombre d’erreurs lors de la sélection de petites cibles est $7\times$ plus élevé en moyenne qu’avec de grandes cibles. Les comparaisons par paires montrent des différences statistiquement significatives entre les petites et grandes cibles tailles ($p = 0,0002$) et entre les petites et moyennes cibles ($p = 0,016$), et entre les moyennes et grandes cibles ($p = 0,05$). Aucune différence statistiquement significative n’est observée entre les tailles pour *RayLens* ($RL_{petite} = 0,63$ erreur, $RL_{moyenne} = 0,39$ erreur, $RL_{grande} = 0,40$ erreur) et *RaySlider* ($RS_{petite} = 0,49$ erreur, $RS_{moyenne} = 0,47$ erreur, $RS_{grande} = 0,31$ erreur).

Pour les petites cibles, les comparaisons par paires montrent que le nombre d’er-

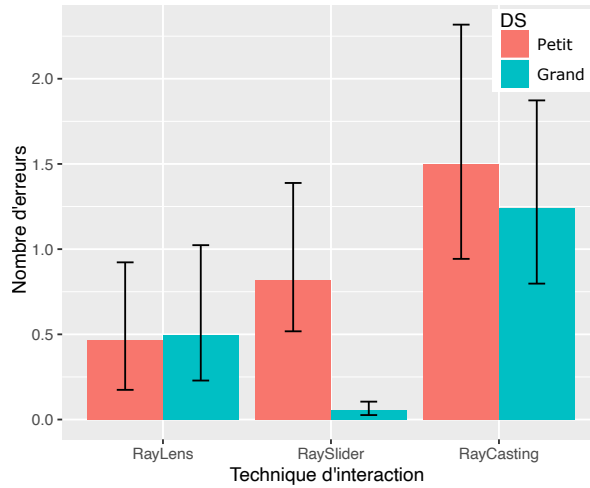


FIGURE 5.13 – Effet d’interaction entre l’espace de densité et la technique sur le nombre d’erreurs.

reurs avec *RayCasting* est significativement plus élevé qu’avec *RayLens* ($p = 0,001$) et avec *RaySlider* ($p = 0,0006$). Aucune différence statistiquement significative n’est observée entre *RayLens* et *RaySlider* pour les petites cibles. Pour les cibles moyennes, *RayLens* est significativement plus précis que *RayCasting* ($p = 0,04$). Aucune différence statistiquement significative n’est observée entre *RaySlider* et les deux autres techniques. Enfin, pour les grandes cibles, aucune différence statistiquement significative n’est observée entre les techniques.

Effet d’interaction DS \times TECH. Un effet d’interaction statistiquement significatif est observé entre DS et TECH [$F_{2,187} = 11,39, p < 0,0001$]. Les résultats obtenus sont présentés à la Figure 5.13. Les comparaisons par paires montrent une différence statistiquement significative entre les densités pour *RaySlider* ($RS_{DS_{\text{petit}}} = 0,82$ erreur, $RS_{DS_{\text{grand}}} = 0,05$ erreur, $p < 0,0001$). Aucune différence statistiquement significative n’est observée pour *RayLens* ($RL_{DS_{\text{petit}}} = 0,47$ erreur, $RL_{DS_{\text{grand}}} = 0,49$ erreur) et *RayCasting* ($RC_{DS_{\text{petit}}} = 1,48$ erreur, $RC_{DS_{\text{grand}}} = 1,23$ erreur).

En forte densité (DS petit), *RayLens* est significativement la technique la plus précise. Les comparaisons par paires montrent des différences statistiquement significatives entre *RayLens* et *RayCasting* ($p = 0,009$) et entre *RayLens* et *RaySlider* ($p = 0,046$). Aucune différence statistiquement significative n’est observée en forte densité entre *RayCasting* et *RaySlider*.

En faible densité (DS grand), *RaySlider* est de loin la technique la plus précise. Les comparaisons par paires montrent une différence statistiquement significative entre *RaySlider* et *RayLens* ($p = 0,006$) et entre *RaySlider* et *RayCasting* ($p < 0,0001$).

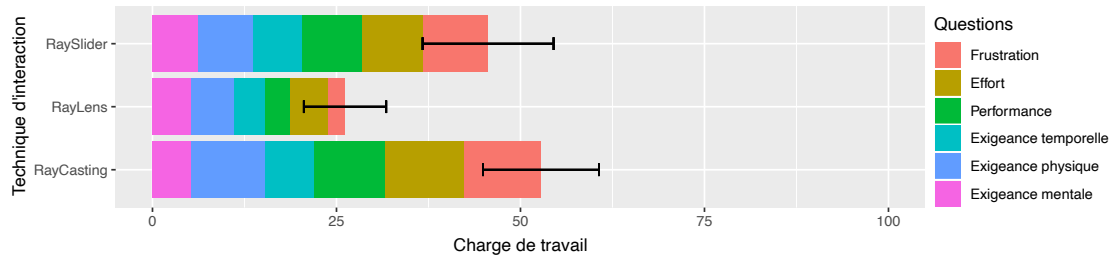


FIGURE 5.14 – [RTLX] Charge de travail totale entre 0 et 100.

La technique *RayLens* est également significativement plus précise que *RayCasting* ($p = 0,009$).

En résumé, les résultats de cette étude montrent que *RayCasting* est la technique la moins précise (nombre moyen d'erreurs de sélection le plus élevé), tandis que les techniques *RayLens* et *RaySlider* présentent un nombre équivalent d'erreurs en moyenne. Pour la technique *RayLens*, les résultats ne montrent pas d'impact de la taille de la cible et de la densité sur le nombre d'erreurs. En revanche, le nombre d'erreurs avec la technique *RaySlider* est très impacté par la densité. De plus, le nombre d'erreurs avec la technique *RayCasting* est très impacté par la taille de la cible. Enfin, en environnements peu denses, *RaySlider* est la technique la plus précise de l'étude.

5.3.2 Résultats qualitatifs

Charge de travail

Les résultats obtenus à partir du questionnaire RTLX (Annexe F) sont reportés à la Figure 5.14. Les participants estiment que la technique *RayLens* nécessite de loin la charge de travail la plus faible ($26,16/100 \pm 5,6$). Les deux autres techniques présentent une charge de travail similaire ($45,6/100 \pm 8,9$ pour *RaySlider* et $52,78/100 \pm 7,9$ pour *RayCasting*).

En particulier, les participants estiment que la technique *RayLens* est beaucoup moins frustrante (≈ 4 fois moins) et qu'elle demande moins d'effort (≈ 2 fois moins) que les autres techniques. Ils estiment également réussir mieux avec cette technique qu'avec les deux autres (≈ 2 fois plus).

Classements des techniques

Le questionnaire final (Annexe G) confirme les résultats du questionnaire RTLX : tous les participants choisissent *RayLens* comme étant la technique la moins frus-

trante, la plus précise et celle avec laquelle ils ont le mieux réussi les tâches. Nous pensions que déplacer la lentille pour trouver une cible dans un environnement dense serait difficile pour les participants. Cependant, les participants ne rapportent aucune difficulté liée au déplacement de la lentille.

Le classement des techniques par les participants montre que *RayLens* est également perçue comme étant la technique la moins exigeante physiquement (*RayLens* : 9/12 participants, *RaySlider* : 3/12 participants) et demandant le moins d'effort (*RayLens* : 9/12 participants, *RaySlider* : 2/12 participants, *RayCasting* : 1/12 participant).

Ce classement montre également que *RayCasting* est perçue comme étant la technique la moins exigeante mentalement (*RayCasting* : 8/12 participants, *RaySlider* : 3/12 participants, *RayLens* : 1/12 participant).

Enfin, tous les participants préfèrent la technique *RayLens*. Ils déclarent aimer cette technique car les sélections sont faciles et stables. À l'opposé, le classement des techniques montre que la technique *RayCasting* est la moins appréciée pour 7/12 participants. Les 5 participants restants choisissent *RaySlider* comme la technique la moins appréciée.

Utilité des retours visuels utilisés avec les techniques *RaySlider* et *RayLens*

Nous avons questionné les participants sur l'utilité (mesure sur une échelle de 1 à 7) des retours visuels des techniques *RayLens* et *RaySlider* : l'effet de transparence et la représentation du champ de vision de la lentille. Pour la technique *RaySlider*, le filtre de transparence appliqué sur les objets placés devant le curseur est jugé utile avec une note moyenne de 5/7. Pour la technique *RayLens*, le filtre de transparence devant la lentille est jugé utile avec une note moyenne de 5,6/7. Le retour visuel représentant le champ de vision de la lentille (cône + éclaircissement des éléments dans le champ) est également jugé utile par les participants avec une note moyenne de 4,6/7.

5.3.3 Analyse des temps de déplacement de la lentille et du curseur

Pour les techniques *RayLens* et *RaySlider*, nous avons analysé a posteriori les temps moyens de déplacement de la lentille et du curseur par rapport au temps total de réalisation de la tâche de sélection. Ces différents temps sont reportés à la Figure 5.15, par technique et par taille de cibles ou espacement de densité de l'environnement. Nous observons qu'en moyenne, les participants ont passé $54,9\% \pm$

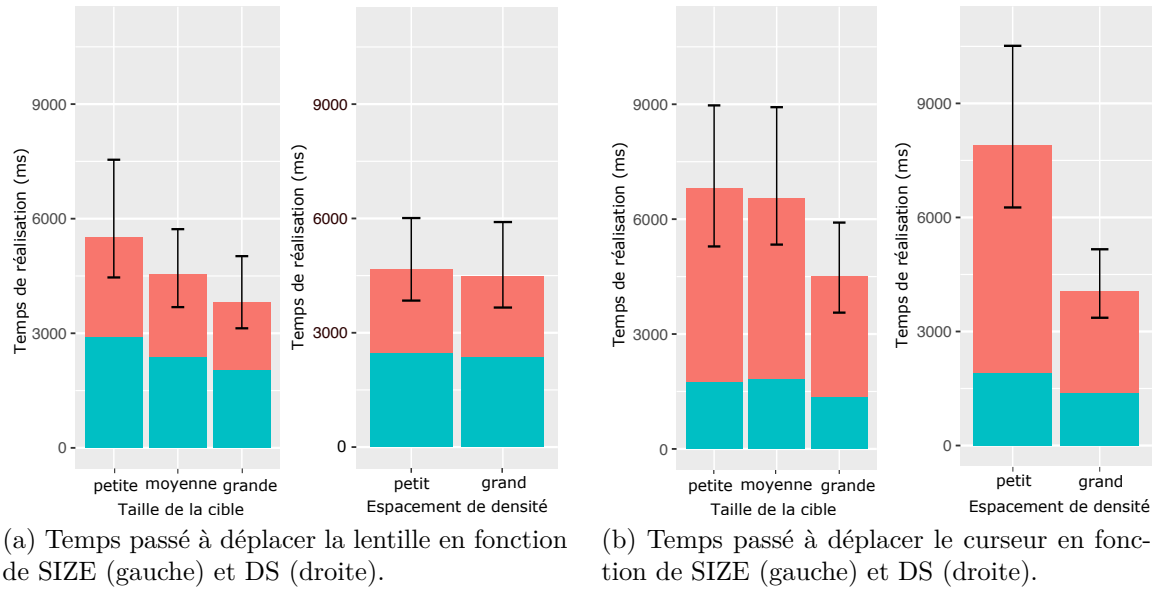


FIGURE 5.15 – Temps moyen passé à déplacer la lentille et le curseur (en bleu) sur le temps total passé à sélectionner les cibles. En rouge, le temps moyen pendant lequel la lentille et le curseur ne sont pas déplacés.

15% des temps de sélection à déplacer la lentille, contre seulement $34\% \pm 9\%$ à déplacer le curseur.

Pour la technique *RayLens*, le temps passé à déplacer la lentille est légèrement impacté par la taille de la cible : le temps moyen de déplacement de la lentille est de 2,04 s pour les grandes cibles, 2,37 s pour les cibles de taille moyenne, et 2,91 s pour les petites cibles (Figure 5.15a – gauche). Cependant, nous n’observons pas d’effet statistiquement significatif de SIZE ni de DS. De même, nous n’observons pas d’effet statistiquement significatif de SIZE ni de DS sur le temps passé à déplacer le curseur avec *RaySlider* (Figure 5.15b).

5.4 Étude expérimentale : discussion

5.4.1 Hypothèses

Les techniques *RaySlider* et *RayLens* plus rapides et plus précises que la technique *RayCasting* pour sélectionner de petites cibles.

Pour les sélections de grandes cibles, la technique *RayCasting* reste la meilleure option : elle permet à une utilisatrice de sélectionner de grandes cibles plus rapidement et avec la même précision que les deux extensions du raycasting, *RaySlider*

et *RayLens*. La première hypothèse *H1* suggérait que les techniques *RaySlider* et *RayLens* seraient plus performantes que la technique *RayCasting* pour la sélection de petites cibles. Les résultats de cette étude mettent tout d’abord en évidence les limites de la technique *RayCasting* et montrent l’avantage des techniques *RayLens* et *RaySlider* pour sélectionner des petites cibles. Les résultats montrent également que la technique *RayLens* est plus rapide que la technique *RayCasting*. Cependant, comme la technique *RaySlider* n’est que légèrement plus rapide que la technique *RayCasting*, l’hypothèse *H1* est partiellement validée.

Enfin, pour les cibles de taille moyenne, les techniques *RayCasting* et *RayLens* présentent des temps moyens de sélection similaires : cela pourrait révéler l’existence d’une taille seuil en dessous de laquelle la technique *RayLens* devient plus précise que la technique *RayCasting*.

La technique *RayLens* peu impactée par la densité de l’environnement.

Cette étude montre que la technique *RayLens* présente des temps moyens de réalisation et un nombre moyen d’erreurs similaires quelle que soit la densité. En revanche, la technique *RaySlider* est fortement impactée par la densité (2 fois plus d’erreurs en forte densité). La seconde hypothèse *H2* supposait qu’en faible densité (grand espacement de densité) les performances de la technique *RaySlider* surpasseraient celle de la technique *RayLens*. Les résultats obtenus montrent que les trois techniques présentent des temps moyens de réalisation similaires en faible densité, cela invalide donc l’hypothèse *H2*.

La troisième hypothèse *H3* supposait qu’en forte densité (petit espacement de densité) les performances de la technique *RayLens* surpasseraient celles de la technique *RaySlider*. Les résultats obtenus montrent que *RayLens* est la technique la plus précise en environnements denses. Bien que nous n’observions pas de différence statistiquement significative en termes de temps moyen de réalisation, les résultats suggèrent tout de même qu’en moyenne la technique *RayLens* est 1,6 fois plus rapide que la technique *RaySlider* en forte densité. Ainsi, les résultats obtenus ne permettent que de valider partiellement l’hypothèse *H3*.

Charge de travail plus faible pour la technique *RayLens*.

La quatrième hypothèse (*H4*) de l’étude supposait que la technique *RayCasting* et la technique *RayLens* nécessitaient respectivement la charge de travail la plus faible et la plus forte. Pourtant, les résultats du questionnaire RTLX montrent que la technique *RayLens* nécessite la charge de travail la plus faible. Elle est également la technique préférée de tous les participants. À l’inverse, les techniques *RaySlider*

et *RayCasting* présentent une charge de travail équivalente et également bien plus élevée que celle de la technique *RayLens*. Ainsi, basculer entre la vue 3D de la scène et la vue projetée 2D affichée sur la lentille ne semble pas entraîner un effort cognitif supplémentaire. Les résultats invalident donc l’hypothèse *H4*.

5.4.2 Contrôle du curseur, de la lentille, et du rayon

***RaySlider* : contrôle couplé du curseur et du rayon, le slider principalement utilisé en phase balistique.**

Le placement du curseur dans la scène à l’aide du slider est rapporté comme rapide par certains participants (le participant 8 rapporte un “mouvement automatique avec le slider pour amener le curseur dans la scène”). Comme expliqué dans la Section 5.1.2, cette technique a été conçue selon le principe que seule la main dominante est utilisée pour la phase de précision (le curseur étant utilisé uniquement pour la phase balistique). Ce comportement est confirmé par les données enregistrées. En calculant le ratio : *temps passé entre la dernière utilisation du curseur et la fin de la tâche* sur le *temps total de réalisation de la tâche*, nous obtenons 0,48. Ce ratio montre que le curseur n’est pas utilisé pendant la seconde moitié du temps total de sélection.

Largement inspirée de la variable d’espacement de densité de [163], nous avons contrôlé la densité de l’environnement en plaçant 4 distracteurs autour de la cible : un devant, un derrière, mais aussi un à gauche et un à droite. Pour la phase de précision, et particulièrement en forte densité, les participants doivent alors être précis le long de l’axe de profondeur mais également sur l’axe gauche-droite. La fin de la tâche s’apparente donc à une tâche de positionnement précis d’un curseur en 3D, ce qui explique la difficulté de sélectionner des objets dans des environnements denses avec cette technique.

***RayLens* : contrôle indépendant de la lentille et du rayon.**

La conception de la technique *RayLens* suit les principes d’une technique bi-manuelle “asymétrique et dépendante” [78]. Comme expliqué pour la première fois dans [78], “une telle cohérence est un bon point de départ pour identifier l’usage à deux mains qui semble naturel”. Ceci est confirmé par les résultats expérimentaux (Section 5.3, aucun effet d’apprentissage) : les participants atteignent rapidement de bonnes performances avec la technique *RayLens*. De plus, avec cette conception bi-manuelle, une utilisatrice peut affiner la position de la lentille (par exemple, pour modifier le grossissement de la cible sur la lentille) et pointer vers la cible en même

temps, sans basculer entre le mouvement de la lentille et le mouvement du rayon. Cette possibilité a été observée au cours de l'étude. Néanmoins, sur la base des mouvements enregistrés de la lentille et du rayon, nous n'avons pas pu quantifier cette utilisation synergique des deux mains. Cependant, 5/12 participants ont demandé pendant la phase d'entraînement s'il était possible de pointer et de sélectionner la cible tout en déplaçant/tenant la lentille. Ils ont également indiqué que cette fonctionnalité était utile. L'utilisation synergique (par exemple, pour la sélection relativement facile d'une cible affichée sur la lentille) est aussi utile que l'utilisation séquentielle consistant à placer la lentille à une position fixe puis à sélectionner la cible dessus (par exemple, pour la sélection difficile d'une petite cible même après grossissement).

5.4.3 Avantages de la technique *RayLens*

Une bonne précision.

Alors que les performances des techniques *RayCasting* et *RaySlider* sont respectivement fortement impactées par la taille de la cible et la densité de l'environnement, les résultats montrent que la technique *RayLens* présente des performances équivalentes, quelle que soit la taille de la cible ou de la densité de l'environnement. En moyenne, dans un environnement dense, *RayLens* reste globalement une technique précise et est 1,7 fois plus précise que *RaySlider* et presque 3,2 fois plus précise que la technique *RayCasting*.

Deux éléments de conception de la technique *RayLens* peuvent expliquer ces résultats. Premièrement, le grossissement facilite la tâche de pointage, en particulier pour les petites cibles. Par exemple, si l'on considère une lentille positionnée dans l'espace moteur à 35 cm d'une petite cible de 0,5 cm de diamètre, la largeur de la cible projetée sur la lentille est de 2 cm : plus de 4 fois son diamètre initial. De plus, le grossissement augmente la distance entre les distracteurs proches et la cible, à la fois dans l'espace moteur et dans l'espace visuel. Enfin, le filtre de transparence appliqué sur les objets placés devant la lentille supprime toute occultation visuelle et motrice de la cible. Par conséquent, avec la technique *RayLens*, les cibles deviennent plus grandes, la densité et l'occultation sont diminuées, contribuant à une bonne précision.

Une sélection 2D facile sur la lentille.

D'après les résultats obtenus, les tâches de sélection avec la technique *RayLens* sont faciles et rapides en environnements denses et avec des petites cibles. La concep-

tion de cette technique peut expliquer ces résultats. Tout d’abord, la sélection d’un objet se fait sur la lentille et non directement sur l’objet 3D. La distance de pointage est ainsi plus courte ce qui simplifie la tâche et contribue également à réduire le problème de stabilité du rayon. Deuxièmement, lorsqu’une utilisatrice pointe sur le plan défini par la lentille, la tâche de pointage est plus simple car, d’une part, le nombre de degré de liberté est réduit puisque la tâche de pointage 3D devient une tâche de pointage 2D¹; et d’autre part, la cible est élargie en raison du grossissement.

Une technique facile à apprendre et à utiliser.

La technique *RayLens* nécessite la charge de travail la plus faible et est la technique préférée des participants de cette étude. Comme indiqué dans la Section 5.4.2, sa conception bi-manuelle “asymétrique et dépendante” [78] implique un usage à deux mains qui semble “naturel” [78] et évite une phase d’apprentissage (Section 5.3).

1. La réduction du nombre de degrés de liberté est un des conseils proposés dans le livre 3DUI [89], Section 7.11 Design Guidelines.

6. Résumé de l'étude sur les améliorations du raycasting

Dans ce chapitre, nous avons tout d'abord présenté l'état de l'art des techniques permettant de faciliter la sélection par raycasting des petites cibles ou des cibles cachées (Section 4.1). Cependant, la majeure partie des techniques existantes sont peu performantes en environnements denses. Par exemple, le mécanisme d'expansion de cibles est efficace pour la sélection de petites cibles en environnements épars, mais il est également très impacté par la proximité de distracteurs autour de la cible. De la même manière, les techniques d'affinement progressif sont très impactées par la densité de l'environnement : le nombre d'étapes d'affinement étant dépendant du nombre d'objets présents autour de l'objet ciblé.

D'autres techniques utilisant le concept de zoom ont l'avantage de ne pas être impactées par la densité. Mais une technique de zoom standard n'est pas efficace pour la sélection d'objets cachés (un objet initialement caché restant caché dans une vue agrandie). Pour résoudre ce problème, des techniques réarrangent spatialement les objets agrandis (sous la forme d'une grille par exemple). Néanmoins, ces techniques perdent le contexte global de la scène ou le contexte autour de la cible, ce qui peut être inacceptable dans certains cas d'utilisation comme les cartes augmentées.

Partant de ce constat, nous avons étudié une technique basée sur un mécanisme de lentille grossissante déplaçable dans l'espace 3D. L'utilisation d'une lentille grossissante pour faciliter la sélection d'objets distants n'avait jamais été considérée dans le contexte de RA sur table. Notre rationnel de conception était que la technique serait indépendante des cibles et peu impactée par la densité de l'environnement. De plus, le déplacement 3D de la lentille permet d'atteindre des cibles cachées. Enfin, un déplacement à distance de la lentille combiné à une sélection par raycasting sur la lentille permet la sélection d'objets distants sans que l'utilisatrice n'ait besoin de se déplacer. Pour revenir sur le tableau de synthèse, Tableau 4.1, une technique basée sur une lentille grossissante répondait aux quatre critères.

Ainsi, nous avons conçu et développé *RayLens*, la première technique de facilitation de pointage étendant le raycasting avec une lentille pour une sélection précise en environnements 3D denses. *RayLens* est une technique bi-manuelle qui combine un rayon tenu par la main dominante, avec une lentille virtuelle 2D grossissante

contrôlée par l'autre main. La lentille est déplaçable à distance et un filtre de transparence est appliqué sur tous les objets présents entre la lentille et l'utilisatrice afin de réduire l'occultation de la lentille.

Une étude expérimentale a ensuite permis de comparer *RayLens* avec une technique de raycasting standard ainsi qu'avec *RaySlider*, une technique bi-manuelle étendant le raycasting avec le mécanisme d'expansion de cibles largement utilisé pour faciliter le pointage des petites cibles. Les résultats de cette étude confirment tout d'abord que le raycasting standard est rapide pour la sélection de grands objets et met également en évidence les limites de cette technique avec des petits objets. Les résultats montrent également que les deux extensions *RayLens* et *RaySlider* ont une précision bien meilleure que le raycasting standard pour la sélection de petits objets. De plus, nous avons observé que les performances des techniques *RayLens* et *RaySlider* différaient en environnements denses. Alors que la technique *RaySlider* est fortement impactée par la densité (proximité des distracteurs), *RayLens* offre des performances homogènes quelles que soient la densité et la taille de la cible. Enfin, les résultats qualitatifs montrent que *RayLens* est la technique préférée des participants, qu'elle est facile à comprendre et qu'elle nécessite la plus faible charge de travail.

Au vu de nos résultats, nous suggérons d'utiliser une lentille 2D grossissante comme une aide pour une sélection distante précise. Alors que l'utilisation d'un simple rayon reste à privilégier pour des sélections faciles (sélections de grandes cibles par exemple), l'extension que nous proposons avec *RayLens* ne doit être utilisée qu'en cas de besoin (c'est-à-dire pour la sélection de petites cibles ou la sélection en environnements denses). De plus, comme la technique *RayLens* est facile à apprendre et intuitive, nous pensons qu'elle peut être employée par des utilisateurs novices, par exemple dans des applications publiques où les utilisateurs doivent maîtriser rapidement la technique (pour explorer une carte 3D d'un site à visiter par exemple).

Chapitre III

Sélection d'objets proches : trois
espaces d'interaction disponibles
en réalité augmentée sur table

Introduction

Le Chapitre II est consacré à la sélection précise d'objets 3D distants par la métaphore du pointeur virtuel. Nous nous sommes alors concentrés sur des techniques améliorant le raycasting, toutes reposant sur l'espace d'interaction 3D dans l'air en entrée. Ce chapitre complète le précédent en étudiant la sélection d'objets à portée de main de l'utilisatrice et les améliorations possibles de la métaphore de la main virtuelle. De plus, le fait que les objets 3D à sélectionner soient à portée de main permet aussi d'explorer l'utilisation de la table comme espace d'interaction en entrée et comme aide à la sélection précise d'objets 3D.

Le chapitre¹ est structuré en deux sections. La première section (Section 7) dresse un état de l'art des techniques d'interaction en RA sur table pour la sélection d'objets à portée de main qui est organisé selon l'espace d'interaction en entrée : 3D au-dessus de la table, 2D sur la table ou hybride 2D/3D combinant l'espace d'interaction sur et au-dessus de la table. La deuxième section (Section 8) est dédiée à l'étude expérimentale comparative de techniques 3D dans l'air, 2D sur la table et hybride qui combine l'interaction dans l'air avec l'interaction au toucher sur la table.

Sommaire

7	État de l'art sur les techniques d'interaction 2D/3D en RA sur table : focus sur les espaces d'interaction en entrée	112
7.1	Espace 3D dans l'air au-dessus de la table	114
7.2	Espace 2D sur la table	114
7.3	Espace hybride : combiner l'espace 2D sur table et l'espace 3D dans l'air .	117
7.4	Études comparatives existantes	121
7.5	Bilan	126
8	Contribution : Comparaison des trois espaces d'interaction disponibles	128
8.1	Éléments de conception : trois techniques étudiées	129
8.2	Étude expérimentale 1 : sélection d'objets 3D visibles ancrés à la table dans un environnement éparé	133
8.3	Étude expérimentale 1 : résultats	137
8.4	Étude expérimentale 2 : sélection d'objets 3D occultés ancrés à la table dans un environnement dense	146
8.5	Étude expérimentale 2 : résultats	148
8.6	Études expérimentales 1 et 2 : discussion	152
9	Résumé de l'étude sur les trois espaces d'interaction	156
10	Mise en contexte : vers une installation artistique exploitant les trois espaces d'interaction disponibles en RA sur table	159

1. Les travaux présentés dans ce chapitre ont été en partie publiés à la conférence ACM-AVI 2020, the 2020 International Conference on Advanced Visual Interfaces. Carole Plasson, Dominique Cunin, Yann Laurillau, & Laurence Nigay. 3D Tabletop AR : A Comparison of Mid-Air, Touch and Touch+Mid-Air Interaction. <https://doi.org/10.1145/3399715.3399836>.

7. État de l’art sur les techniques d’interaction 2D/3D en RA sur table : focus sur les espaces d’interaction en entrée

Dans cette section, nous étudions les techniques d’interaction existantes pour sélectionner un objet 3D à portée de main en réalité augmentée (RA) sur table. Complémentaire à l’état de l’art du chapitre précédent qui est dédié aux extensions de la métaphore du pointeur virtuel (ou raycasting) pour la sélection d’objets distants, nous nous concentrons dans cette section sur les techniques étendant la métaphore de la main virtuelle pour la sélection d’objets à portée de main.

De plus, le fait que les objets 3D à sélectionner soient à portée de main permet d’explorer l’utilisation de la table comme espace d’interaction en entrée. Nous l’avons déjà observé dans l’expérimentation décrite dans le Chapitre I : certains participants ont sélectionné un objet 3D placé sur la table par toucher direct sur la table (stratégies avec la technique *DirectTouch*, Section 2.5). De plus, l’utilisation de “références spatiales” telles que la table et l’utilisation de “contraintes physiques” comme la suppression de degrés de liberté superflus sont des pistes proposées par Hinckley et al. [72] pour limiter les problèmes rencontrés en interaction 3D (problèmes pour percevoir et interagir avec les environnements virtuels 3D).

Ainsi la sélection d’objets à portée de main permet d’étudier plusieurs espaces d’interaction, 2D (toucher sur la table), 3D (interaction dans l’air au-dessus de la table) ou hybride combinant 2D et 3D. Cette variété d’espaces d’interaction en entrée, 2D, 3D ou hybride renvoie à la question de la représentation mentale ou structure perceptive de la tâche de sélection d’objets 3D.

Au début des années 90, Jacob et al. [76, 77] ont montré “l’importance de faire correspondre la structure perceptive de la tâche et la structure de contrôle de la tâche”. Dans une étude, ils ont comparé deux tâches à trois attributs :

- *Tâche 1* (x , y , *taille*). Les participants devaient faire correspondre la position (x , y) et la taille d’un carré à celles d’un carré de référence.
- *Tâche 2* (x , y , *couleur*). Les participants devaient faire correspondre la position (x , y) et le niveau de gris d’un carré à ceux d’un carré de référence.

Les deux tâches étaient réalisées soit avec un Polhemus 3SPACE permettant une

interaction intégrant les trois dimensions, soit par une souris d'ordinateur et un slider pour une interaction 2D et une interaction 1D distinctes. Leurs résultats montrent de meilleures performances pour les conditions *Polhemus/tâche1* et *souris+slider/tâche2*. Selon les auteurs, cette différence serait due à la perception de la tâche. Pour la première tâche, les trois dimensions sont intrinsèquement liées dans la représentation mentale de la tâche faite par l'utilisatrice (tâche intégrale). Pour la deuxième tâche, les dimensions sont au contraire perçues comme étant séparées (tâche séparable) : une tâche 2D (gestion de la position en X et en Y) d'une part et une tâche 1D (gestion de la couleur) d'autre part. Les performances seraient donc meilleures lorsque la structure perceptive de la tâche (intégrale ou séparable) et la technique d'interaction correspondent.

Cependant, il n'est pas toujours évident de savoir comment une utilisatrice va percevoir et se représenter mentalement une tâche. Cette représentation mentale est de plus variable d'une utilisatrice à l'autre. Alors qu'une majeure partie des systèmes de RA interagissent dans l'air (3D), la RA sur table a l'avantage de mettre à notre disposition d'autres espaces d'interaction comme l'espace de la table (2D) et l'espace combinant le plan de la table et l'espace au-dessus de celle-ci (2D+3D). Ces différents espaces d'interaction permettent de concevoir des techniques d'interaction pour la sélection d'un objet 3D à portée de main qui peuvent correspondre à différentes perceptions de la tâche de sélection :

- Comme une tâche 3D intégrale : les trois dimensions XYZ sont intrinsèquement liées dans la représentation mentale de la tâche par l'utilisatrice ;
- Comme une tâche séparable : une tâche 2D (gestion de la position XY sur le plan 2D de la table) d'une part et une tâche 1D (gestion de la hauteur Z de la sélection par rapport à la table) d'autre part.

Dans cette section, nous présentons différentes utilisations de la table dans le contexte de la sélection d'objets 3D placés au-dessus de la table. Pour cela, nous passons en revue des techniques de sélection et de positionnement d'objets 3D de la littérature utilisées en RA sur table. Notre revue des techniques est organisée selon leur espace d'interaction en entrée :

- Espace 3D dans l'air au-dessus de la table (Section 7.1) ;
- Espace 2D sur la table (Section 7.2) ;
- Espace hybride combinant l'espace 2D et l'espace 3D (Section 7.3).

Pour terminer cette section, nous décrivons en Section 7.4 des études qui ont été menées pour comparer des techniques reposant sur ces espaces d'interaction.

7.1 Espace 3D dans l'air au-dessus de la table

Pour sélectionner des objets proches de soi, la métaphore de la main virtuelle est communément appliquée : l'utilisatrice pointe directement sur l'objet d'intérêt avec sa main (la technique *DirectTouch* étudiée au Chapitre I-Section 2.1 implémente la version standard de cette métaphore). Dans le cas de la RA sur table, cette interaction se fait dans l'espace 3D situé au-dessus de la table. Cette technique est la plus évidente pour interagir avec des objets 3D car elle intègre les trois dimensions simultanément et permet donc une correspondance exacte (*mapping 1-1*) entre la main réelle de l'utilisatrice et sa main virtuelle. Ceci définit donc une manière intuitive de sélectionner un objet car elle simule la façon dont nous manipulons les objets dans le monde réel [31]. Mais interagir dans l'air n'est pas sans difficulté [104] en raison de la limitation de la précision du pointage humain, des tremblements de la main et de la fatigue [26, 68, 71]. De plus, certains problèmes de perception d'une scène augmentée, tels que la difficulté d'estimer la distance des objets virtuels [86, 135, 173] peuvent fortement impacter les performances d'une tâche de sélection.

7.2 Espace 2D sur la table

Du fait des limitations de l'interaction 3D dans l'air, les recherches en RA sur table se sont alors tournées vers l'interaction tactile [19, 145, 155] pour interagir avec des objets virtuels 3D. Ces techniques d'interaction tactile impliquent une décomposition de la tâche¹ de sélection 3D en une tâche 2D et une tâche 1D [101, 103].

Benko et Feiner [19] proposent une technique multi-touch, appelée *Balloon Selection*, pour la sélection d'objets 3D au-dessus d'une table. Cette technique utilise la métaphore d'un ballon d'hélium attaché à une ficelle, comme illustré à la Figure 7.1a. Une personne tient à deux mains une ficelle à laquelle est attaché un ballon gonflé à l'hélium. Le ballon flotte au-dessus de sa main droite, celle la plus proche du ballon. Elle peut ensuite faire monter ou descendre le ballon en rapprochant ou en éloignant ses mains l'une de l'autre tout en faisant glisser la corde entre les doigts de sa main droite.

La technique *Balloon Selection*, présentée à la Figure 7.1b, implémente cette métaphore. Le ballon est représenté par une sphère virtuelle attachée à l'extrémité d'une ficelle virtuelle tenue par les deux index de l'utilisatrice en contact avec la table.

1. Selon la terminologie de Jacob et al. [76, 77], ces techniques impliquent des *tâches séparables*. Dans la suite de ce manuscrit nous utiliserons plutôt la notion de *décomposition de la tâche* comme utilisée dans de nombreux travaux.

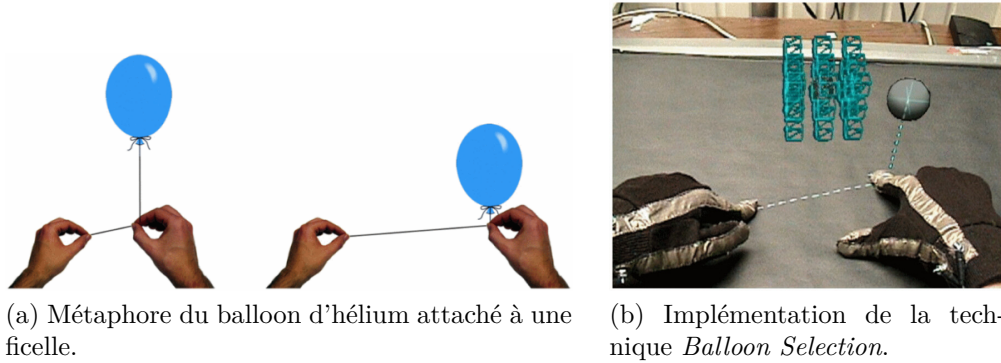


FIGURE 7.1 – *Balloon Selection* de Benko et Feiner [19]. Images extraites de [19].

Le ballon est utilisé comme un collisionneur pour sélectionner des objets virtuels : une sélection se produit lorsque ce ballon entre en collision avec un objet (et *un* objet seulement). Le positionnement 3D du ballon est décomposé en une tâche de positionnement 2D et une tâche de positionnement 1D, les deux tâches pouvant être réalisées en parallèle. D'une part, le mouvement de la main proche du ballon (Main 1) place celui-ci sur le plan 2D de la table (tâche 2D). D'autre part, la distance entre les index (Main 1 et Main 2) de l'utilisatrice ajuste la hauteur du ballon par rapport à la table (tâche 1D). La validation d'une sélection est faite en rapprochant le pouce et l'index de la main (Main 2) tenant le bout de la ficelle (action semblable à un "clic"). De l'autre main (Main 1), celle la plus proche du ballon, l'utilisatrice peut diminuer ou agrandir la taille de la sphère en rapprochant ou en éloignant le pouce de son index.

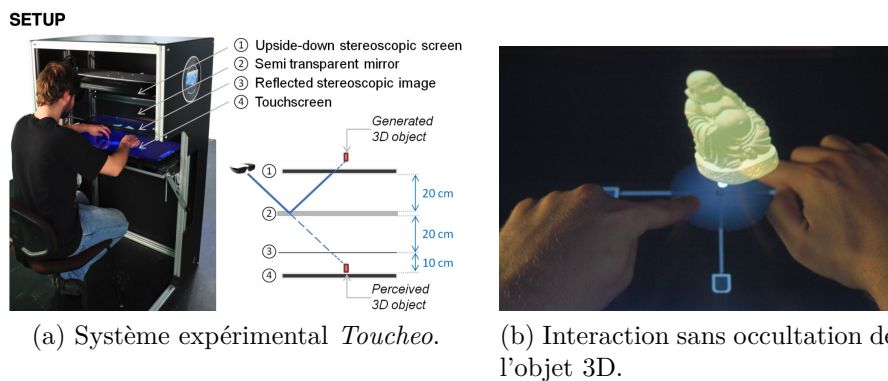


FIGURE 7.2 – *Toucheo* de Hachet et al. [67]. Image extraite de [67].

Hachet et al. proposent *Toucheo* [67], un système combinant une visualisation stéréoscopique d'objets 3D et des interactions multi-touch sur une table. Ce système est présenté à la Figure 7.2a. La particularité de ce système est qu'il permet des

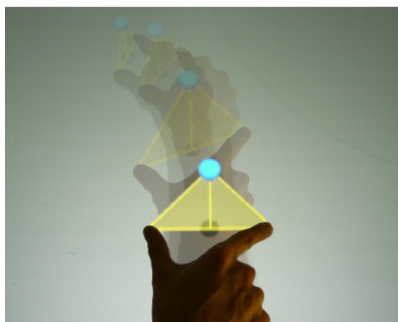


FIGURE 7.3 – *Triangle Cursor* de Strothoff et al. [155]. Image extraite de [155].

interactions tactiles avec des objets 3D virtuels sans aucune occultation causée par les mains (Figure 7.2b). Pour interagir avec un objet virtuel 3D, un disque est affiché sur la table juste en dessous de l'objet et un lien virtuel permet de lier visuellement l'objet et ce disque (ancrage de l'objet à la table). L'utilisatrice touche simplement ce disque pour sélectionner l'objet. Des *widgets* sont affichés autour de ce disque (Figure 7.2b) et permettent de faire des rotations et de changer la taille de l'objet. L'utilisatrice peut également déplacer l'objet dans l'espace 3D au-dessus de la table de la même manière que *Balloon Selection* [19] : l'utilisatrice touche le centre du disque avec une main, et un widget apparaît pour l'ajustement de la hauteur. Le déplacement 2D de l'objet sur la table ainsi que l'ajustement de sa hauteur peuvent ensuite être réalisés simultanément.

Par la suite, Strothoff et al. proposent *Triangle Cursor* [155] qui permet, comme Benko et Feiner [19], de positionner une sphère virtuelle (c'est-à-dire un collisionneur) dans l'espace 3D en interagissant seulement sur le plan 2D d'une table interactive. Cependant, contrairement à *Balloon Selection*, cette technique n'utilise qu'une seule main : quand une utilisatrice place deux doigts sur la table, deux triangles isocèles apparaissent entre ses doigts, comme illustré à la Figure 7.3. La sphère flotte alors dans l'air au sommet de ces deux triangles, sa position 2D sur le plan de la table est le centre des doigts de l'utilisatrice. Enfin, la hauteur des triangles, et par extension la hauteur de la sphère, dépend de l'espacement des doigts de l'utilisatrice : si l'utilisatrice rapproche ses doigts la sphère monte, si elle les éloigne alors la sphère descend. Le déplacement 2D de la sphère sur la table ainsi que l'ajustement de sa hauteur peuvent être réalisés simultanément.

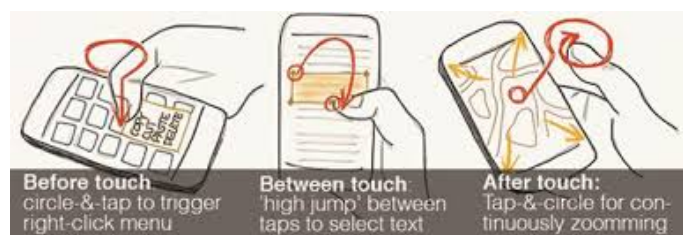


FIGURE 7.4 – Exemples d’application de la modalité *Air+Touch* de Chen et al. [42]. Image extraite de [42].

7.3 Espace hybride : combiner l’espace 2D sur table et l’espace 3D dans l’air

Combiner l’interaction tactile et l’interaction dans l’air pour créer de nouvelles possibilités d’interaction a été étudié autant dans le contexte des appareils mobiles que dans le contexte de la RA sur table.

7.3.1 Interaction avec des appareils mobiles

Dans le contexte d’appareils mobiles, Chen et al. [42] implémentent un prototype permettant de suivre précisément les mouvements du doigt dans l’air au-dessus de l’écran d’un téléphone à l’aide d’une caméra de profondeur attachée au téléphone. Les auteurs créent alors une nouvelle modalité appelée *air+touch* permettant de combiner des interactions dans l’air et des interactions tactiles sur un téléphone, et identifient plusieurs tâches et applications de cette modalité (Figure 7.4). Par exemple, l’utilisatrice d’une application de cartes interactives comme google maps sélectionne un commerce en appuyant directement sur celui-ci puis zoome ou dézoome autour de ce commerce en faisant continuellement des cercles dans l’air, dans un sens ou dans l’autre, au-dessus de l’écran [99].

7.3.2 Réalité augmentée : cinq classes d’utilisation de l’espace hybride

De l’étude des systèmes existants en RA sur table combinant un espace 2D et un espace 3D, nous identifions cinq principales classes d’utilisation de l’espace hybride. Les trois premières classes d’utilisation sont tirées des travaux de Marquardt et al. [100] présentant différentes techniques d’interaction pour unifier ces deux espaces en un seul et même espace, appelé *the continuous interaction space*.

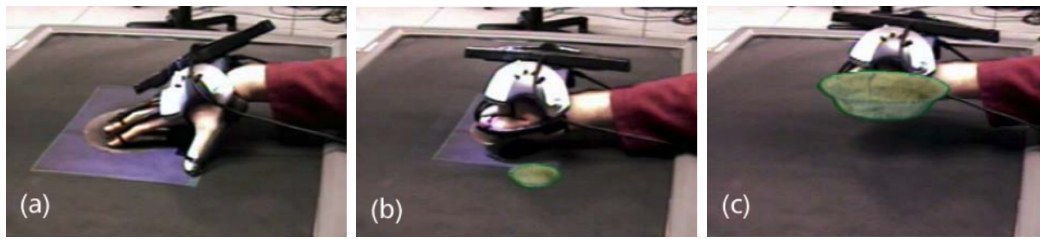


FIGURE 7.5 – Environnement immersif hybride de Benko et al. [20]. Images extraites de [20].

Sélection d'un objet par interaction dans l'air, manipulation par interaction tactile

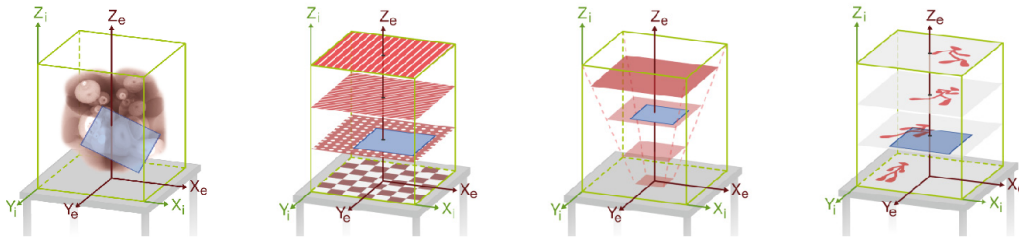
Banerjee et al. [17] proposent *Pointable*, une technique bi-manuelle permettant :

- De sélectionner et de déplacer un objet par des interactions dans l'air avec la main dominante.
- De changer la taille ou l'orientation de cet objet par des interactions tactiles sur la table avec la main non dominante.

En utilisant la métaphore du pointeur virtuel dans la main dominante, l'utilisation de l'espace 3D permet alors d'étendre la zone d'interaction d'une utilisatrice (un des exemples d'utilisation proposés par Marquardt et al. [100]) : l'utilisatrice peut alors sélectionner et déplacer des objets hors de sa portée.

Interaction avec des objets virtuels 2D affichés sur la table par des gestes naturels dans l'air

L'intégration de l'espace 2D de la table et de l'espace 3D au-dessus permet également d'interagir dans l'air au-dessus des objets 2D virtuels affichés sur la table. En imitant des gestes de la vie quotidienne, une utilisatrice peut par exemple faire le geste d'attraper un objet (fermeture du poing sur la table) au-dessus d'un objet 2D pour l'attraper. Elle peut ensuite le déplacer et le relâcher en ouvrant le poing. De la même manière, l'environnement immersif hybride de Benko et al. [20] combinant un casque de RA et une table interactive simule de tels comportements physiques. Dans cet environnement, l'utilisatrice passe d'une représentation 2D à une représentation 3D d'un objet en effectuant des gestes 2D sur la table, puis des gestes dans l'air. Par exemple, si l'utilisatrice pose sa main sur la table, au-dessus d'un objet 2D et ferme le poing comme pour prendre cet objet, la représentation 2D de l'objet disparaît et sa représentation 3D apparaît au niveau de sa main. Cet exemple est présenté à la Figure 7.5.

FIGURE 7.6 – *PaperLens* de Spindler et Dachsel [146]. Images extraites de [146].

Interaction avec des couches discrètes et lentilles magiques

Un autre exemple liant l'espace de la table et l'espace 3D au-dessus de celle-ci est la manipulation d'une lentille (physique comme une tablette ou virtuelle) dans l'espace au-dessus de la table [75, 146, 147]. Par exemple, Spindler et Dachsel avec *PaperLens* [146, 147] implémentent une lentille magique avec une feuille rigide pour explorer l'espace 3D au-dessus de la table. Cette feuille agit comme une fenêtre sur le monde virtuel, permettant de visualiser différents espaces d'interaction : des espaces volumétriques, des espaces en couches, des espaces zoomables et des espaces temporels. Ces quatre espaces d'information sont illustrés à la Figure 7.6.

Dans le cas de la visualisation d'information en couches, l'espace au-dessus de la table est séparé en plusieurs couches horizontales, chacune permettant de visualiser une caractéristique unique d'un modèle. Spindler et Dachsel illustrent cette application en utilisant un modèle 2D d'un corps humain projeté sur la table. Les couches peuvent alors représenter différentes informations d'un corps humain (les os, les muscles par exemple). Il est donc possible de visualiser une couche spécifique sur la feuille en la plaçant horizontalement à une hauteur donnée. Pour changer de couche à visualiser, l'utilisatrice ajuste simplement la hauteur de la feuille.

Les auteurs proposent également l'utilisation de *PaperLens* pour gérer le niveau de zoom d'une information 2D présente sur la table (espace d'information zoomable). Une utilisatrice déplace la feuille au-dessus de la table et voit sur la feuille une vue zoomée de l'information présente dessous. La hauteur de la feuille permet alors de gérer le niveau de zoom : plus la feuille est haute (proche de ses yeux), plus le niveau de zoom est élevé.

Modélisation 3D : contrôle décomposé

En plus des exemples proposés par Marquardt et al. [100], cet espace hybride 2D/3D peut également être utilisé dans le contexte de la modélisation 3D [48, 177]. Par exemple, Xiao et al. [177] étendent l'espace d'interaction dans l'air de l'HoloLens en utilisant l'interaction tactile sur des surfaces 2D non instrumentées comme

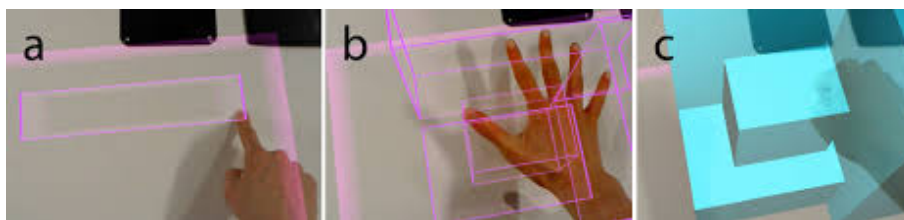


FIGURE 7.7 – Interaction hybride de Xiao et al. [177]. Images extraites de [177].

des tables ou des murs, sans infrastructure supplémentaire. Plusieurs scénarios d'interaction illustrent la transition fluide entre l'interaction tactile sur la table et l'interaction dans l'air. Par exemple après avoir dessiné le plan d'un bâtiment sur la table (interaction tactile sur une table), une utilisatrice monte ou descend sa main au-dessus du plan 2D de la table pour ajuster la hauteur du bâtiment. Cet exemple est illustré à la Figure 7.7.

Exploration multi-échelle rapide de grands espaces d'information 2D par des interactions dans l'air

Menzner et al. [109] proposent de combiner un téléphone et un casque de RV HTC Vive pour l'exploration de grands espaces d'information 2D. Contrairement à de nombreuses études en RA et RV [50, 105, 133, 156, 169, 170] qui se sont intéressées à l'utilisation de surfaces tactiles pour fournir essentiellement des interactions avec retour tangible, les auteurs combinent des interactions tactiles sur l'écran et des interactions dans l'air au-dessus du téléphone.

Dans leur étude, le téléphone est posé sur une table et une grande carte 2D virtuelle est affichée dans le casque. L'affichage de la carte dans le casque permet ainsi de la voir au-dessus de la table et au-delà de l'écran du téléphone. Combinant des interactions tactiles sur le téléphone et des interactions dans l'air au-dessus du téléphone, les auteurs proposent une technique efficace pour sélectionner rapidement des points d'intérêt dans une grande carte 2D. Pour explorer la carte (déplacement, changement d'échelle), l'utilisatrice interagit dans l'air au-dessus du téléphone. La position 3D de son doigt encode alors deux variables :

- *Hauteur du doigt sur l'axe Z pour régler le niveau de zoom.* L'utilisatrice déplace son doigt entre 0 et 5 cm au-dessus de l'écran. Cette zone est représentée à la Figure 7.8 par un cube en pointillé et est coupée en deux sur l'axe Z : la partie haute permet de dézoomer la carte, la partie basse permet de zoomer. La vitesse du zoom avant/arrière dépend de la hauteur du doigt dans ces parties. Ainsi, si le doigt est dans la partie haute du cube elle va pouvoir dézoomer la carte plus ou moins rapidement (Figure 7.8 a, b) : si son doigt est très haut

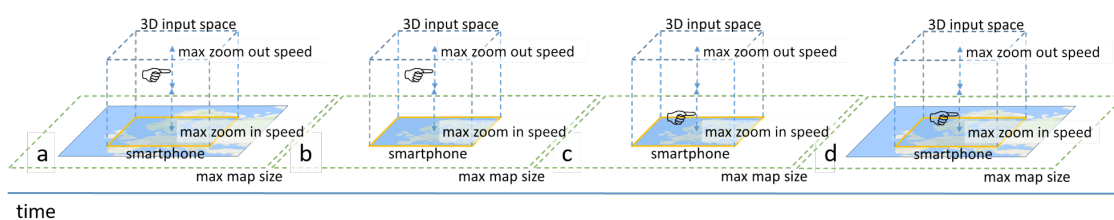


FIGURE 7.8 – Interaction hybride de Menzner et al. [109] pour l'exploration de cartes virtuelles. La position du doigt sur l'axe Z perpendiculaire à l'écran définit le niveau de zoom, la position sur le plan XY définit la position XY du curseur sur la carte. Images extraites de [109].

alors le zoom arrière sera rapide (ceci lui permet d'avoir rapidement une vue d'ensemble). En redescendant son doigt, toujours dans la partie supérieure du cube, la vitesse du zoom arrière ralentira. Au contraire, si elle place son doigt dans la partie basse du cube (Figure 7.8 c, d), l'utilisatrice pourra zoomer et la vitesse du zoom avant accélérera si elle se rapproche de l'écran. Le point de pivot du zoom est toujours la projection du doigt sur le plan XY de l'écran.

- *Position du doigt sur le plan horizontal XY de l'écran pour se déplacer.* La position du doigt sur le plan parallèle à l'écran permet de se déplacer sur la carte. La vitesse de déplacement est calculée à partir de la position relative du doigt par rapport au centre de l'écran. Par exemple, lorsque l'utilisatrice lève son doigt et le déplace vers la gauche par rapport au centre de l'écran, l'utilisatrice se déplace vers la partie gauche de la carte. Si son doigt est proche du centre du téléphone, le déplacement est lent et si elle s'en éloigne le déplacement accélère.

Enfin, la sélection d'un point d'intérêt est faite par interaction au toucher direct sur l'écran.

7.4 Études comparatives existantes

Comme nous l'avons illustré précédemment, la RA sur table offre trois espaces d'interaction : l'espace 3D au-dessus de la table, l'espace 2D sur la table et l'espace hybride combinant ces deux derniers. Dans cette section nous présentons plusieurs études de la littérature qui comparent des techniques reposant sur ces espaces d'interaction pour des tâches de manipulation 3D, de positionnement 3D et de sélection 2D et 3D.

Manipulation 3D : interaction tactile vs interaction dans l'air

Mendes et al. [106] comparent cinq techniques issues de la littérature afin de comprendre quelle approche est la plus adaptée pour des tâches de manipulation d'objets (déplacement 3D, rotation 3D, changement d'échelle) sur table stéréoscopique. Quatre d'entre elles considèrent des interactions dans l'air, la dernière n'utilise que l'interaction tactile (technique de référence). Avec la technique d'interaction tactile, l'utilisatrice touche la projection 2D d'un objet flottant pour le déplacer sur le plan XY de la table. En maintenant cette sélection, trois *widgets* apparaissent : un pour le contrôle de la hauteur de l'objet, et deux autres pour les rotations roulis et tangage. Le contrôle de la hauteur de l'objet est réalisé en bi-manuel et est inspiré de la technique *Balloon Selection* [19]. Si le deuxième doigt est posé sur la table *en dehors des widgets*, l'utilisatrice peut soit faire une rotation de l'objet autour de l'axe perpendiculaire à la surface de la table, soit changer la taille de l'objet en rapprochant ou en éloignant le deuxième doigt de la projection. Leurs résultats montrent que pour les tâches de translation 2D et les tâches de translation 3D avec changement d'échelle, l'interaction tactile est plus lente que les interactions dans l'air. Une explication proposée par les auteurs est qu'il y a beaucoup de temps perdu lorsque l'utilisatrice passe d'une interaction à l'autre. Par exemple, alterner entre une interaction bi-manuelle sans widget (pour changer la taille de l'objet) et une interaction avec *widgets* (pour la rotation roulis) est complexe et chronophage.

Sélection 2D : interaction tactile vs interaction hybride

Dans leur étude comparative, Menzner et al. [109] opposent deux techniques pour des tâches de recherche suivies de tâches de sélection 2D sur des petites et grandes cartes 2D. Cette étude compare :

- Une technique n'utilisant que des interactions tactiles pour le zoom, le déplacement et la sélection (*pinch-to-zoom*, *drag-to-pan*).
- Une technique hybride combinant des interactions dans l'air pour le zoom et le déplacement sur le plan 2D et une interaction au toucher direct pour la sélection de la cible (cette technique hybride est décrite à la fin de la Section 7.3).

Leurs résultats montrent que la technique hybride surpasse la technique de référence en termes de temps de réalisation des tâches, d'utilisabilité et de charge de travail (NASA-TLX). Ces différences entre les deux techniques sont encore plus importantes pour l'exploration des grandes cartes. Enfin, la technique hybride est aussi la technique préférée des participants.

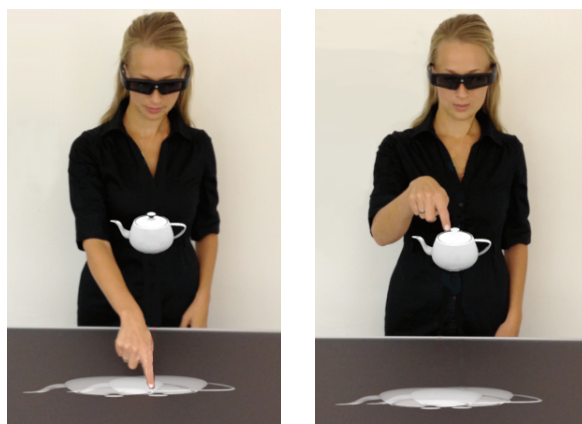


FIGURE 7.9 – Étude comparative de Bruder et al. [34, 35] et Colley et al. [47]. Images modifiées de [35].

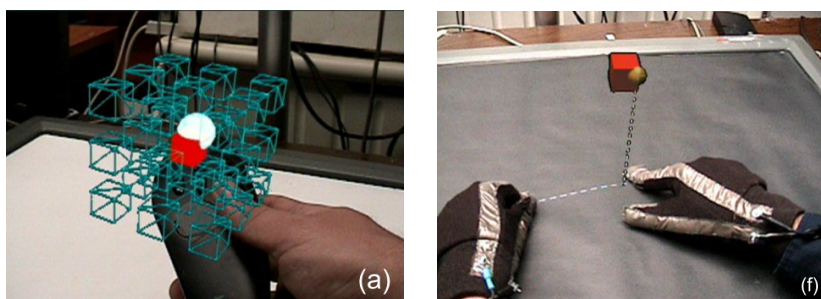


FIGURE 7.10 – Étude comparative de Benko et Feiner [19]. Images extraites de [19].

Sélection 3D : interaction tactile vs interaction dans l'air

Avec un système combinant une table interactive et un affichage 3D stéréoscopique (modèle 3D affiché par des lunettes 3D et créé à partir de deux images 2D affichées sur la table), Bruder et al. [34, 35] et Colley et al. [47] comparent le pointage 3D dans l'air avec le toucher 2D sur la table (Figure 7.9). Pour le toucher 2D, l'utilisatrice passe “à travers” la vue stéréoscopique de l'objet pour toucher son image 2D sur la table. Les résultats de leur expérimentation montrent que la position de la cible, et plus particulièrement sa hauteur par rapport à la table, impacte fortement la précision des techniques de sélection. En effet, leurs résultats montrent que la sélection 2D est meilleure que la sélection 3D dans l'air uniquement pour des objets proches de la surface (jusqu'à 10 cm de l'écran). Pour des cibles plus hautes (au-dessus de 10 cm de la table), la sélection 3D dans l'air est meilleure que le toucher 2D.

Benko et Feiner [19] comparent une technique de pointage directe (*WAND*) utilisant une manette suivie en temps réel dans l'espace 3D à la technique *Balloon Selection* (*BALLOON*) n'utilisant que l'interaction tactile (cette technique est décrite en

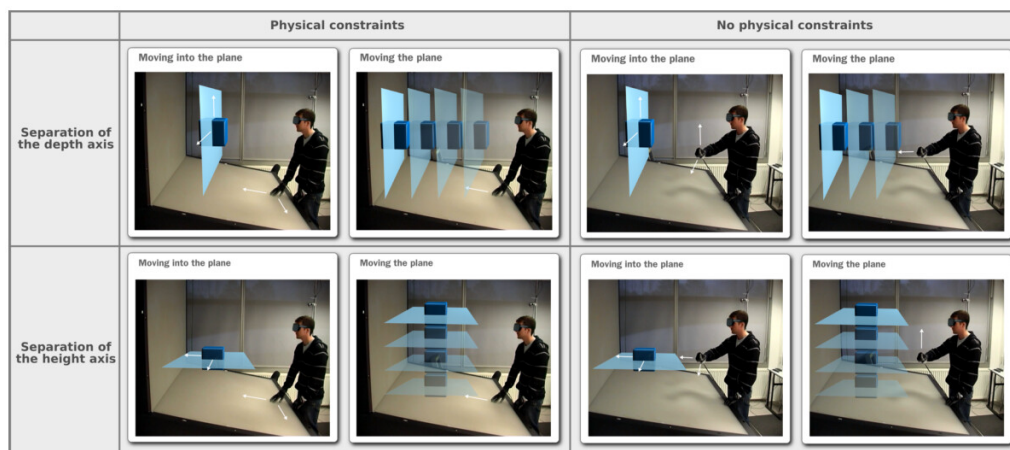


FIGURE 7.11 – Étude comparative de Veit et al. [165]. Image extraite de [165].

détail dans la Section 7.2). Les deux techniques sont illustrées à la Figure 7.10. Les résultats montrent que *BALLOON* et *WAND* présentent des temps de réalisation similaires pour les sélections 3D : la rapidité des utilisateurs à réaliser les tâches de sélection 3D ne semble donc pas pénalisée par la décomposition de la tâche induite par *BALLOON*. Par contre, *BALLOON* est la technique la plus précise de leur étude avec trois fois moins d'erreurs que *WAND*. Selon les auteurs, le faible taux d'erreurs obtenu par *BALLOON* peut être expliqué par le support physique fourni par la table pendant la sélection et la validation de la tâche. En effet, le fait de ne pas interagir dans l'air mais seulement sur la table rend *BALLOON* beaucoup moins fatigante et réduit considérablement voire supprime totalement les problèmes de tremblement de la main durant la validation de la sélection.

Positionnement 3D : interaction dans l'air avec ou sans décomposition vs interaction tactile

Veit et al. [165] comparent également des techniques d'interaction dans l'air avec des techniques d'interaction tactile pour une tâche de positionnement 3D. Le but de leur étude est d'évaluer l'impact de la table et l'impact de la décomposition de la tâche sur les performances des utilisateurs. Pour la décomposition de la tâche, ils s'intéressent plus particulièrement à deux manières de décomposer la tâche de positionnement 3D en une tâche 2D et une tâche 1D :

- *Décomposition 1 – Séparation de l'axe de profondeur Y* : positionnement de l'objet sur le plan vertical XZ puis ajustement de sa profondeur sur l'axe Y.
- *Décomposition 2 – Séparation de l'axe de hauteur Z* : positionnement de l'objet sur le plan horizontal XY puis ajustement de sa hauteur sur l'axe Z.

Pour évaluer l'impact de la table et de la décomposition, les auteurs comparent une technique de manipulation directe dans l'air (sans décomposition de la tâche) à quatre autres techniques décomposant la tâche 3D. Ces quatre techniques sont illustrées à la Figure 7.11. Parmi ces techniques, deux n'utilisent que des interactions tactiles, et les deux autres n'interagissent que dans l'air. Les deux techniques tactiles décomposent la tâche différemment : l'une sépare l'axe de profondeur Y (*Décomposition 1*), l'autre sépare l'axe de hauteur Z (*Décomposition 2*). Par exemple la technique d'interaction tactile utilisant la *Décomposition 1* (technique en haut à gauche dans le tableau Figure 7.11) déplace l'objet dans le plan XZ avec l'index puis déplace l'objet sur l'axe de profondeur Y avec le majeur.

De la même manière, les deux techniques d'interaction dans l'air décomposent la tâche en séparant soit l'axe de profondeur (*Décomposition 1*), soit l'axe de hauteur (*Décomposition 2*). Par exemple, avec la technique utilisant la *décomposition 1* (technique en haut à droite dans le tableau Figure 7.11), l'utilisatrice pince l'index et le pouce et maintient la position pour déplacer l'objet dans le plan XZ. Le déplacement en profondeur Y se fait en pinçant l'index et le majeur. Cette étude révèle trois principaux résultats :

- *Support de la table.* Tout d'abord, l'interaction tactile réduit la fatigue et n'a pas d'impact sur la précision par rapport à l'interaction dans l'air.
- *Décomposition de la tâche.* Les techniques décomposant la tâche sont plus lentes que celle intégrant les 3 dimensions simultanément. En termes de précision, les résultats montrent que les techniques (tactile ou dans l'air) utilisant la *décomposition 1* (séparation de l'axe de profondeur Y) sont les techniques les plus précises. Celles qui utilisent la *décomposition 2* (séparation de l'axe de hauteur Z) obtiennent un taux d'erreurs similaire à celui de la manipulation directe. Il semble donc que la manière dont la tâche est décomposée ait un impact direct sur la précision.
- *Évaluation empirique.* Les auteurs observent aussi que les participants intègrent les trois dimensions simultanément seulement lors la phase balistique. Pendant la phase de contrôle, ils décomposent généralement la tâche. Ce constat est aussi fait par Balakrishnan et al. [11].

7.5 Bilan

Alors que l'interaction dans l'air est couramment utilisée en RA, l'ajout d'une surface interactive en RA sur table élargit considérablement les possibilités d'interaction. Trois espaces d'interaction en entrée sont alors disponibles : le plan 2D de la table, l'espace 3D au-dessus ou l'espace hybride unifiant ces deux derniers. Nous avons présenté dans les Sections 7.1 à 7.3 plusieurs techniques d'interaction s'appuyant sur ces différents espaces.

Dans le contexte des tâches de sélection ou de positionnement d'objets 3D, plusieurs études ont comparé des interactions tactiles dans l'espace 2D de la table et des interactions 3D dans l'air. À partir de ces études comparatives (Section 7.4), nous concluons tout d'abord que :

1. Grâce au support de la table, l'interaction tactile réduit la fatigue par rapport à l'interaction dans l'air ;
2. Les résultats concernant la décomposition de la tâche sont contradictoires. Par exemple, dans l'étude de Benko et Feiner [19], la décomposition de la tâche ne semble pas impacter les performances (*Balloon selection* obtient des temps de réalisation similaires à ceux de l'interaction dans l'air et montre également une meilleure précision). Cependant, dans l'étude de Veit et al. [165], les techniques décomposant la tâche sont plus lentes que l'interaction 3D intégrant les trois dimensions à la fois.

De plus, aucune étude existante n'a comparé l'espace d'interaction 2D de la table et l'espace 3D au-dessus de la table avec un espace d'interaction hybride 2D+3D dans le contexte de la sélection d'objets 3D.

Enfin, nous formulons l'hypothèse que l'interaction sur table serait plus pertinente dans le cas d'objets ancrés à la table et non flottant au-dessus. Or, à l'exception de l'étude de Mendes et al. [106] qui affichent sur la table les ombres des objets flottants, les études comparatives menées se concentrent sur la sélection d'objets 3D flottant au-dessus et sans aucun lien visuel avec la table. Pourtant dans notre contexte de RA sur table, les objets sont principalement ancrés à la table. Par exemple dans le cas des cartes augmentées, les éléments 3D sélectionnables (les bâtiments par exemple) sont souvent placés directement sur la table. Pour les autres éléments qui flottent dans l'air au-dessus de la table, ils peuvent facilement être ancrés à la table à l'aide d'une ombre projetée. Pour la sélection d'objets ancrés à la table, aucune étude n'a comparé ces espaces d'interaction en entrée.

En conséquence, pour une tâche de sélection 3D d'objets ancrés à la table, des questions demeurent : quel espace d'interaction en entrée est le plus approprié (espace 2D, espace 3D, espace hybride) ? La décomposition de la tâche induite dans l'espace 2D et l'espace hybride 2D+3D est-elle pénalisante en termes de performances ou de retours utilisateurs ? La section suivante (Section 8) est consacrée à ces questions.

8. Contribution : comparaison des trois espaces d'interaction en entrée disponibles en RA sur table

La section précédente (Section 7) présente plusieurs techniques d'interaction en réalité augmentée (RA) sur table pour manipuler des objets à portée de main. Cette section a mis en évidence que ces techniques reposent sur trois espaces d'interaction en entrée : 2D (sur la table), 3D (au-dessus de la table) et hybride (2D+3D). Dans cette section nous présentons la première étude comparant ces trois espaces d'interaction en entrée pour des tâches de sélection d'objets 3D ancrés à la table. Nous considérons des objets 3D ancrés car l'interaction tactile a plus de sens lorsque les objets sont ancrés à la table. De plus ces objets ancrés sont fréquents dans de nombreuses applications, comme illustré dans les trois scénarios suivants :

- *Scénario 1¹ - Carte augmentée.* Une promotrice immobilier présente un nouveau quartier et les appartements à vendre dans ce quartier. Un modèle virtuel 3D du quartier est placé sur la table et permet à une acheteuse de visualiser les bâtiments, les routes et les commerces environnants. La promotrice sélectionne des appartements dans des immeubles afin d'afficher une visite virtuelle des appartements à vendre sur une tablette.
- *Scénario 2² - Jeu.* Dans des jeux tels que Minecraft (le jeu de construction populaire), les joueuses peuvent construire des structures dans un monde composé de blocs. Le monde 3D est affiché sur une table et les joueuses construisent des châteaux. Tout au long du processus de création, elles peuvent sélectionner un bloc dans la structure du château (pour le supprimer par exemple).
- *Scénario 3 - Modélisation 3D.* Une designer utilise un poste de travail augmenté pour faire de la modélisation 3D comme dans les travaux de Reipschläger et Dachsel [130] (présenté en Section 1.2.1). Elle porte un casque de RA pour visualiser les objets 3D placés sur un écran interactif horizontal. Elle peut sélectionner une partie du modèle pour l'extruder.

Au-delà des objets 3D directement placés sur la table comme dans les scénarios

1. Utilisation de la RA en immobilier <https://youtu.be/LgBCkIDQjb0>
2. Minecraft en RA <https://youtu.be/xgakdcEzVwg>

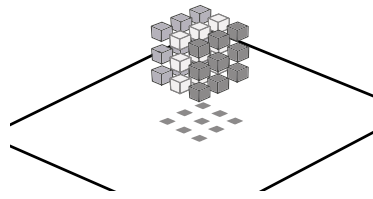


FIGURE 8.1 – Exemple d’objets flottants 3D étudié dans [19] : Ajout d’ombres aux objets afin de les ancrer à la table.

ci-dessus, des liens virtuels [67] ou des projections 2D d’objets 3D flottants (des ombres [106] par exemple) peuvent être affichées sur la table, comme illustré à la Figure 8.1. Les objets 3D ancrés à la table sont donc omniprésents en RA sur table et il est important de comprendre quel espace d’interaction en entrée (2D, 3D ou un espace hybride 2D+3D) est le plus efficace pour les sélectionner.

Pour mener cette étude sur les trois espaces d’interaction et identifier les avantages et limitations de chacun, nous avons implémenté trois techniques, chacune interagissant dans un des trois espaces d’interaction. Ces techniques sont présentées en Section 8.1. Nous avons également réalisé deux études expérimentales afin d’étudier l’effet des espaces d’interaction et de la décomposition de la tâche pour une tâche de sélection 3D d’objets ancrés à la table. La première étude compare ces techniques pour une tâche de sélection simple avec des cibles visibles dans un environnement épars. Le protocole ainsi que les résultats obtenus sont présentés dans les Sections 8.2 à 8.3. La seconde étude compare ces mêmes techniques dans un scénario plus réaliste (par exemple, le scénario 1 avec une carte augmentée d’un quartier dense avec plusieurs bâtiments), avec un environnement dense et des cibles cachées. Le protocole ainsi que les résultats obtenus sont présentés dans les Sections 8.4 à 8.5.

8.1 Éléments de conception : trois techniques étudiées

8.1.1 Espace de conception

La Figure 8.2 classe les trois techniques étudiées en fonction de leur espace d’interaction en entrée (espace 3D dans l’air au-dessus de la table, espace 2D sur la table, espace 2D sur la table + espace 3D dans l’air) et de l’implication, ou non, d’une décomposition de la tâche.

Technique DirectTouch - DT. Nous considérons tout d’abord une technique de

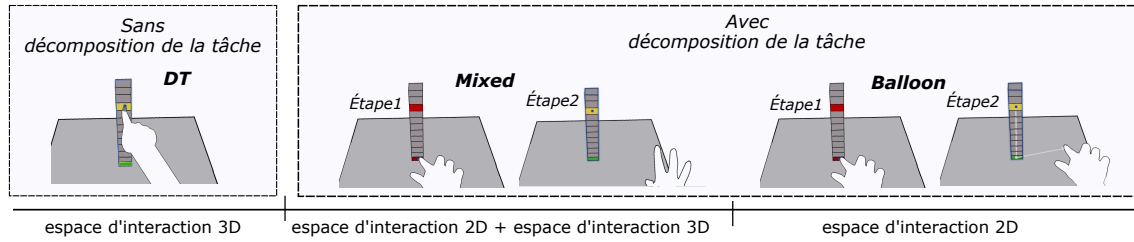


FIGURE 8.2 – Les trois techniques comparées, classées selon i) leurs espaces d'interaction en entrée, et ii) si elles impliquent ou non une décomposition de la tâche.

main virtuelle “standard” [127] (*DirectTouch* - *DT*) comme technique de référence. Selon Pouyrev et al. [127], la main virtuelle (Chapitre III) et le pointeur virtuel (Chapitre II) ont des performances comparables pour la sélection d'objets 3D placés à portée de main. Nous avons alors choisi la technique de la main virtuelle car elle est intuitive pour des interactions avec des objets proches (elle simule la façon dont nous manipulons les objets proches dans le monde réel [89]). Cette technique n'implique pas de décomposition de la tâche, les trois dimensions sont contrôlées simultanément en interagissant dans l'espace 3D au-dessus de la table. Elle est illustrée à gauche à la Figure 8.2.

Technique Balloon. Nous comparons la technique d'interaction 3D *DT* à une technique d'interaction tactile (*Balloon*). Contrairement à *DT*, *Balloon* implique une décomposition de la tâche de sélection 3D en une tâche 2D suivie d'une tâche 1D (Figure 8.3), toutes deux effectuées par interaction tactile sur la table. Cette seconde technique est largement inspirée de *Balloon Selection* [19], une technique détaillée à la Section 7.2. Pour rappel, la technique repose sur la métaphore d'un ballon d'hélium attaché à une corde. La ficelle est maintenue par deux points de contrôle situés dans le plan de la table et le ballon se déplace verticalement, le long de l'axe perpendiculaire à la table, en réduisant ou en augmentant la distance entre ces deux points. La Section 7 décrit plusieurs systèmes [67, 102, 106] utilisant la technique *Balloon Selection*. Nous l'avons aussi choisie car elle permet une interaction directe avec l'objet ciblé (sans contrôle intermédiaire comme un *widget*). De plus, contrairement à *Triangle Cursor* [155] décrite à la Section 7.2, elle permet une interaction à côté de l'objet ciblé, ce qui réduit l'occultation causée par les mains. La technique *Balloon* permet donc d'interagir seulement dans l'espace d'interaction 2D de la table. Elle est illustrée à droite à la Figure 8.2.

Technique Mixed. La troisième technique, *Mixed*, repose sur l'espace d'interaction en entrée hybride. Elle décompose la tâche de sélection 3D de la même manière que la

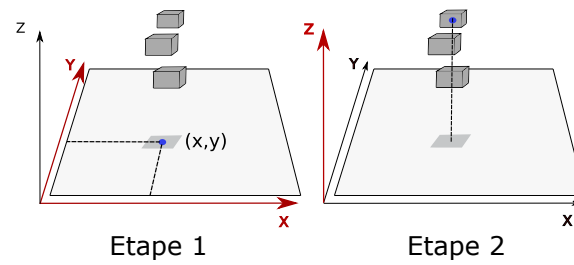


FIGURE 8.3 – Décomposition d'une tâche de sélection 3D en une tâche 2D (étape 1) suivie d'une tâche 1D (étape 2).

technique *Balloon* (Figure 8.3), mais combine de manière transparente l'interaction au toucher sur la table et l'interaction dans l'air. Avec la technique *Mixed*, la tâche 2D (étape 1) est réalisée par une interaction au toucher direct sur la table comme la technique *Balloon*. Cependant, contrairement à *Balloon*, la tâche 1D (étape 2) est réalisée par une interaction dans l'air. La technique *Mixed* interagit donc dans l'espace d'interaction hybride (2D+3D). Elle est illustrée au centre à la Figure 8.2.

Pour les deux techniques *Balloon* et *Mixed* qui impliquent de décomposer la tâche, le choix de conception déterminant l'axe à séparer (tâche 1D) est un facteur clé, comme l'expliquent Veit et al [165]. Nous supposons que le choix de séparer l'axe vertical Z (axe perpendiculaire au plan de la table) comme représenté à la Figure 8.3 est conforme à la structure perceptive de la tâche. En effet, réaliser la tâche 2D fait sens intuitivement en permettant de sélectionner un objet ou de présélectionner un groupe d'objets dans le plan XY (plan de la table). Si les objets sont empilés (voir les trois scénarios présentés en introduction de cette section), une tâche 1D complémentaire est alors effectuée afin de sélectionner avec précision l'objet visé. Cette étape 2 de la Figure 8.3 s'apparente à une phase de désambiguïsation comme expliqué à la Section 4.1.1.2 pour les extensions du raycasting.

8.1.2 Conception des techniques

Nous détaillons la conception de ces trois techniques d'interaction pour sélectionner une boîte 3D spécifique dans une pile de boîtes 3D disposée sur la table (objets 3D ancrés à la table). La base de la pile est en contact avec la table et est identifiée par une boîte rouge (puis verte lorsqu'elle est sélectionnée). Les trois techniques n'utilisent qu'une seule main pour la tâche de sélection.

DirectTouch - DT. L'utilisatrice touche directement la boîte ciblée avec son index (tâche de sélection 3D, Figure 8.4a). Pour cela, une sphère bleue augmentée suit

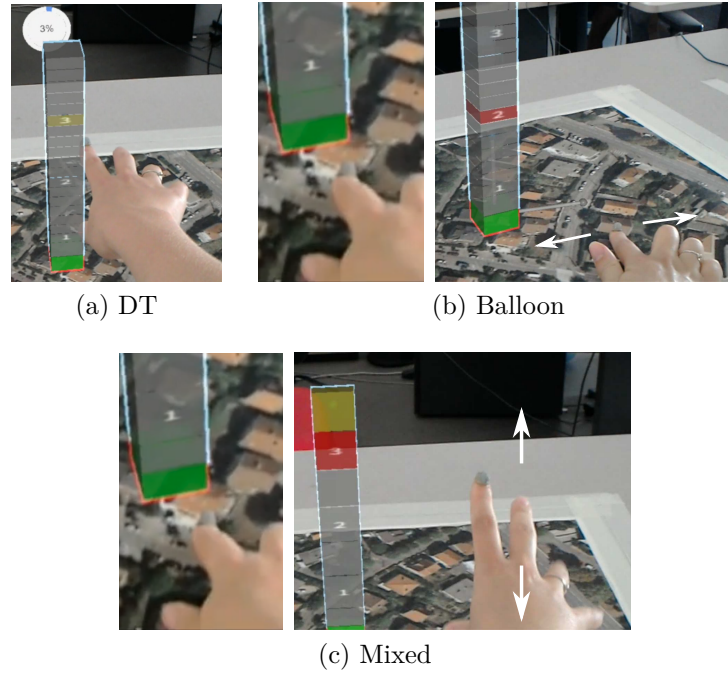


FIGURE 8.4 – Réalisation d'une tâche de sélection avec chacune des trois techniques.

le bout de son doigt et est utilisée à la fois comme collisionneur et comme retour visuel. La sélection se produit lorsque cette sphère est à l'intérieur d'une boîte.

Balloon. L'interaction se décompose en deux étapes. Étape 1, l'utilisatrice sélectionne la base de la pile en touchant cette base sur la table (tâche 2D par interaction tactile), et place ainsi un premier point de contrôle au centre de la base. Étape 2, l'utilisatrice place le même doigt sur la table (où elle veut) afin de placer le deuxième point de contrôle, faisant ainsi apparaître la ficelle virtuelle. Le ballon est la même sphère bleue qu'avec *DT*. L'utilisatrice déplace son doigt sur la table pour ajuster la hauteur de la sphère (tâche 1D par interaction tactile) : la sphère monte ou descend en réduisant ou en augmentant la distance entre les deux points de contrôle (Figure 8.4b). Une sélection se produit lorsque la sphère est à l'intérieur d'une boîte.

Mixed. Comme *Balloon*, l'interaction se décompose en deux étapes. Étape 1, l'utilisatrice sélectionne la base de la pile en touchant cette base sur la table (tâche 2D par interaction tactile) : la sphère bleue qui suit le bout de son doigt doit se trouver totalement à l'intérieur de la base. Cette sphère est ensuite placée au centre de la base et ne peut se déplacer que le long de l'axe vertical Z (tâche 1D³ dans l'air). Étape 2, l'utilisatrice déplace sa main dans l'air et la hauteur de la sphère correspond

3. Il s'agit d'une tâche 1D, et non d'une tâche de pilotage 2D à travers un tunnel 2D comme modélisé par la loi de pilotage *Steering law* [1].

à la hauteur du bout du doigt de l'utilisatrice (Figure 8.4c). Seule la composante Z de la position 3D du doigt est prise en compte. La sélection se produit lorsque cette sphère est à l'intérieur d'une boîte.

8.2 Étude expérimentale 1 : sélection d'objets 3D visibles ancrés à la table dans un environnement épars

8.2.1 Objectif de l'étude et hypothèses

Cette première étude a pour objectif de comparer les performances des trois techniques présentées en Section 8.1 pour une tâche de sélection 3D d'objets visibles. Pour cette étude comparative, nous formulons deux hypothèses :

- *H1.* La technique *DT* est moins précise que les autres techniques en raison des limites de précision du pointage humain, du tremblement de la main et de la fatigue [19, 165].
- *H2.* La décomposition de la tâche n'a pas d'incidence sur les performances des techniques *Balloon* et *Mixed* : elles sont plus performantes (temps de réalisation, précision) que la technique *DT*. La Section 7.5 met en évidence des résultats contradictoires quant à la décomposition de la tâche. Cependant nous pensons que :
 1. L'ancrage des objets donne du sens à la première étape de la tâche de sélection : la tâche 2D sur la table sélectionne la base de la pile. De plus, l'utilisation de la table comme support facilite cette sélection. Pour ces raisons, cette première tâche 2D est intuitive, simple et rapide.
 2. Pour la seconde étape, la précision nécessaire à la réalisation de la sélection est moindre que celle nécessaire avec *DT*.

8.2.2 Système expérimental

Le système expérimental de cette étude, présenté à la Figure 8.5, reprend le système de RA sur table décrit dans le Chapitre I, Section 2.2. Les participants sont debout devant la table (qui est non-interactive). L'image suivie par Vuforia est collée sur la table. Un carré bleu est placé devant l'image : ce carré représente la position initiale du doigt des participants (les participants remplaçant leur doigt dans

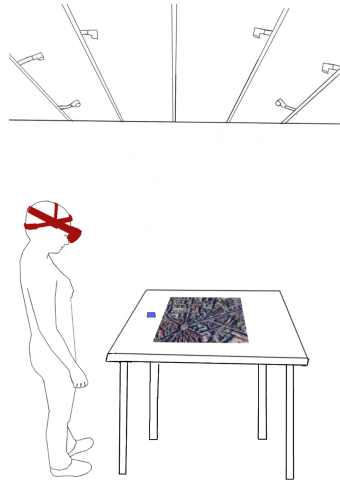


FIGURE 8.5 – Système expérimental.

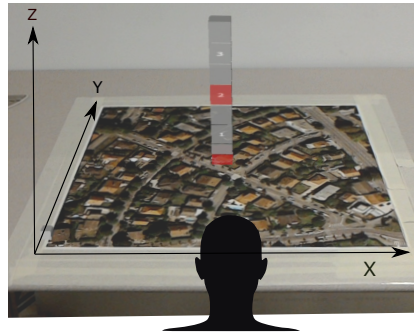


FIGURE 8.6 – [Étude 1] Exemple d'une scène 3D avec une pile de boîtes.

ce carré avant chaque sélection). L'implémentation des trois techniques repose sur les mécanismes décrits dans le Chapitre I, Section 2.2 : le bout du doigt de l'utilisatrice est suivi en temps réel par les caméras OptiTrack. La sphère bleue, utilisée comme collisionneur pour la sélection avec les trois techniques, mesure 8 mm de diamètre.

8.2.3 Tâche

Dans cette étude expérimentale, les participants doivent sélectionner une boîte parmi une pile de boîtes placée sur la table, au centre de l'image cible. Chaque scène 3D est composée d'une seule pile, comme illustrée à la Figure 8.6. La boîte à sélectionner est rouge, toutes les autres sont grises. Toutes les boîtes sont semi-transparentes afin de voir la sphère bleue (le collisionneur) lorsqu'elle est à l'intérieur de celles-ci. Une boîte passe du gris au jaune pour indiquer une sélection. Lorsqu'une utilisatrice touche la base d'une pile sur la table lors de l'étape 1 (pour les techniques *Balloon*, *Mixed*), cette base passe du rouge au vert et la pile de boîtes est mise en évidence en colorant les bordures en bleu (Figure 8.7). Lorsque l'utilisatrice utilise



FIGURE 8.7 – Jauge affichée pour indiquer le temps restant pour valider la sélection.

la technique *DT*, la pile est dès le début mise en évidence et la base est verte.

La confirmation de la sélection n'est pas l'objet de l'étude. Pour les trois techniques, nous utilisons un mécanisme de temporisation (ou *dwell time*) : la sphère bleue doit se trouver entièrement à l'intérieur de la boîte cible pendant 1 seconde [121, 174]. Comme pour les techniques décrites dans la Section 2.1, lorsque le minuteur démarre, une jauge est affichée pour indiquer le temps restant (Figure 8.7). Si la sphère sort de la cible pendant cette seconde, le minuteur est réinitialisé. Comme l'indiquent Ozacar et al. [179], cette confirmation de sélection est la plus utilisée pour l'interaction gestuelle. De plus, d'autres mécanismes de sélection tels que l'utilisation de la main non dominante (cliquer avec un appareil supplémentaire) ou l'utilisation d'une pédale, compliqueraient la tâche ou introduiraient des problèmes de division bi-manuelle de la tâche. Néanmoins, pour une comparaison valide, il est important que la confirmation de la sélection ne pénalise pas une technique plus que les autres. Pour le pointage dans l'air, Ozacar et al. [179] ont comparé expérimentalement deux mécanismes de confirmation : le mécanisme de temporisation et le clic avec un dispositif supplémentaire tenu par la main non dominante. Les auteurs n'ont pas constaté de différence significative entre les deux techniques. Nous concluons que le mécanisme de temporisation pour la confirmation de la sélection ne pénalise pas plus les techniques qui interagissent dans l'air (*DT* et *Mixed*) que la technique *Balloon*.

Chaque tâche de sélection comprend les étapes suivantes :

1. L'utilisatrice place son doigt en position initiale, dans le carré bleu illustré à la Figure 8.5.
2. Pour les techniques *Balloon* et *Mixed*, l'utilisatrice doit tout d'abord sélectionner la base de la pile. La sélection de la base est automatiquement validée lorsque la sphère qui suit son doigt est entièrement à l'intérieur de la base.

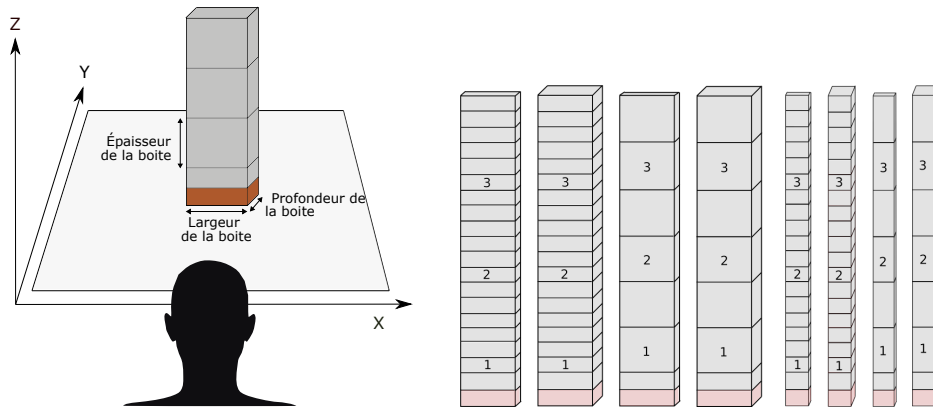


FIGURE 8.8 – Dimensions et formes des cibles.

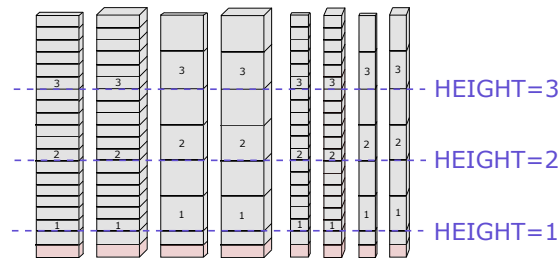


FIGURE 8.9 – Hauteur de la boîte dans la pile.

3. L'utilisatrice doit ensuite sélectionner la boîte cible rouge le plus rapidement possible. Pour valider la sélection, l'utilisatrice doit garder la sphère entièrement dans la cible pendant 1 seconde.

8.2.4 Protocole

L'étude expérimentale est conçue comme une étude intra-sujet, avec les trois variables indépendantes suivantes :

- *La technique d'interaction utilisée (TECH) : DT, Balloon et Mixed.*
- *La forme de la boîte ciblée (SHAPE) : 8 formes de cibles résultant de la combinaison de petites (1,2 cm) et de grandes (3,6 cm) mesures de profondeur, de largeur et d'épaisseur. Ces formes sont illustrées à la Figure 8.8. Notons que la cible et la base de la pile ont toujours la même largeur et profondeur.*
- *La position/hauteur de la boîte dans la pile (HEIGHT) : 3 hauteurs de la cible dans la pile : 3 cm de la table (HEIGHT 1), 10 cm de la table (HEIGHT 2) et 17 cm de la table (HEIGHT 3). Les hauteurs sont illustrées à la Figure 8.9.*

L'ordre des techniques est contrebalancé par un carré latin. Les participants doivent se tenir debout devant la table pendant l'étude. Aucune distance spécifique

entre les participants et la table n'est imposée. S'ils le souhaitent, ils peuvent se pencher vers l'avant ou sur le côté pour mieux voir la pile.

Nous commençons par une brève introduction de l'étude et une présentation de l'HoloLens. Les participants réalisent une phase d'entraînement sur un ensemble de 12 essais avant chaque technique. Ils effectuent ensuite 72 tâches de sélection par technique groupées en 3 blocs d'essais couvrant les 24 combinaisons de SHAPE \times HEIGHT. Ces blocs permettront d'étudier un potentiel effet d'apprentissage ou de fatigue. Les participants remplissent un questionnaire après chaque technique (Annexe H) et un questionnaire de synthèse à la fin de l'étude (Annexe I) pour comparer les trois techniques. L'étude se termine par un entretien. L'étude a duré environ une heure par participant.

Nous avons recruté 15 participants, âgés de 20 à 39 ans (moyenne = $27,67 \pm 4,97$). Aucun d'entre eux n'était un expert en RA. 10 des 15 participants n'avaient jamais utilisé de casque de RA, les autres en avaient utilisé un seulement une fois ou deux.

Au regard de nos hypothèses, nous considérons deux métriques objectives (temps de réalisation, nombre d'erreurs), et deux métriques subjectives (charge de travail, préférence des utilisateurs). Le temps de réalisation de la tâche est le temps écoulé entre le moment où le doigt des participants quitte la position initiale et la fin de la validation de la sélection (fin du mécanisme de temporisation). Le nombre d'erreurs est le nombre de fois où le minuteur est réinitialisé, c'est-à-dire le nombre de fois où la sphère sort de la cible pendant la phase de validation.

Pour les métriques subjectives, le premier questionnaire rempli à la fin de chaque technique reprend le questionnaire RTLX (Annexe H). Le questionnaire final (Annexe I) permet de déterminer la technique la moins exigeante mentalement et physiquement, la technique la plus performante, celle qui demande le moins d'effort et la moins frustrante. Les participants choisissent également la technique la plus précise et la plus rapide. Enfin, les participants classent les techniques par ordre de préférence.

8.3 Étude expérimentale 1 : résultats

Nous présentons dans cette section les résultats quantitatifs et qualitatifs de cette étude. La normalité des données est testée à l'aide du test Shapiro-Wilk. Pour les données normalement distribuées (temps de réalisation) nous utilisons un test statistique par analyse de variance (ANOVA) et des tests de Student (ou tests t) avec ajustement de Bonferroni pour les comparaisons par paires. Les données qui ne

suivent pas une loi normale (nombre d'erreurs) sont pré-traitées par une transformation de rang aligné (ART) [175] avant d'utiliser une ANOVA avec une correction de Bonferroni pour les comparaisons par paires. Dans les graphiques suivants, nous représentons les valeurs moyennes (m) et les intervalles de confiance à 95% calculés avec la méthode bootstrap Bca (la fonction du bootstrap étant la moyenne) [51].

8.3.1 Résultats quantitatifs

Pour l'analyse quantitative, nous calculons :

- T_{Etape1} : le temps pour sélectionner la base de la pile (étape 1 pour les techniques *Balloon* et *Mixed*).
- $T_{\text{Atteindre}}$: le temps pour atteindre la cible pour la première fois après le mouvement balistique (les *overshoots* ne sont pas comptés). Ce temps comprend donc le temps de sélection de la base pour les techniques *Balloon* et *Mixed*.
- T_{Total} : le temps total pour valider la sélection de la cible (temps pour atteindre la cible + temps pour valider la sélection). Le temps total comprend donc les éventuelles entrées et sorties de la boîte ciblée lors de la validation de la sélection par le mécanisme de temporisation.

Pour toutes les techniques, les participants passent 1 seconde dans la boîte pour valider la sélection, nous retirons donc 1 seconde du temps total de réalisation de la tâche de sélection.

Temps de réalisation moyens

En comparant les performances des participants pendant les 3 blocs, nous n'observons pas d'effet d'apprentissage ni de fatigue.

La Figure 8.10 présente le temps total moyen de réalisation d'une tâche de sélection pour chacune des trois techniques. Cette figure présente également le temps moyen nécessaire pour atteindre une cible pour la première fois, ainsi que le temps moyen de sélection de la base de la pile.

Temps moyens de sélection d'une boîte. Nous observons un effet statistiquement significatif de la technique TECH [$F_{2,28} = 9,11, p = 0,0009$], de la forme de la cible SHAPE [$F_{7,98} = 131, p < 0,0001$] et de la hauteur de la cible dans la pile HEIGHT [$F_{2,28} = 94, p < 0,0001$] sur le temps total de réalisation de la tâche de sélection. Le temps total moyen de réalisation de la tâche est de 3,8 s pour *Mixed*, 4,6 s pour *Balloon* et 4,7 s pour *DT*. Les comparaisons par paires montrent qu'il faut sta-

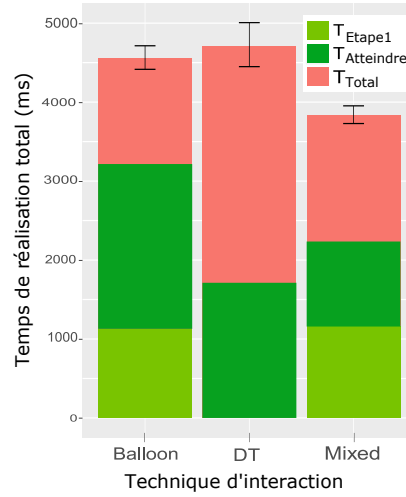


FIGURE 8.10 – Temps total de réalisation de la tâche (T_{Total}), temps pour atteindre la cible pour la première fois ($T_{Atteindre}$), et temps pour sélectionner la base (T_{Etape1}).

tistiquement moins de temps pour sélectionner une boîte avec *Mixed* qu'avec *DT* ($p = 0,03$) et *Balloon* ($p = 0,041$). Nous n'observons pas de différence statistiquement significative entre *DT* et *Balloon*. Le temps moyen pour sélectionner la base est de 1,1 s, soit 25% du temps total de sélection avec *Balloon*, et 30% de temps total de sélection avec *Mixed*.

Temps moyens pour atteindre une boîte. Nous observons un effet statistiquement significatif de la technique TECH [$F_{2,28} = 62,1$, $p < 0,0001$], de la forme de la cible SHAPE [$F_{7,98} = 46$, $p < 0,0001$] et de la hauteur de la cible dans la pile HEIGHT [$F_{2,28} = 101$, $p < 0,0001$] sur le temps nécessaire pour atteindre la cible. En moyenne, les participants atteignent la cible pour la première fois en 1,9 s avec *DT*, en 2,2 s avec *Mixed* et en 3,2 s avec *Balloon*. Les comparaisons par paires montrent qu'il faut statistiquement plus de temps pour atteindre la cible avec *Balloon* qu'avec *Mixed* ($p = 0,0007$) et *DT* ($p < 0,0001$). *DT* est aussi légèrement plus rapide que *Mixed* mais nous n'observons pas de différence statistiquement significative entre ces deux techniques.

Effet d'interaction SHAPE \times TECH. Nous observons un effet d'interaction statistiquement significatif entre SHAPE et TECH [$F_{14,196} = 26,5$, $p < 0,0001$] sur le temps total de réalisation. Les résultats sont présentés à la Figure 8.11a.

Pour *DT*, les résultats montrent de grandes variations entre les formes. En particulier, nous observons une différence importante pour le temps total de sélection entre deux groupes de formes : les quatre formes avec 0 ou 1 petite dimension ($m_{0-1petite} =$

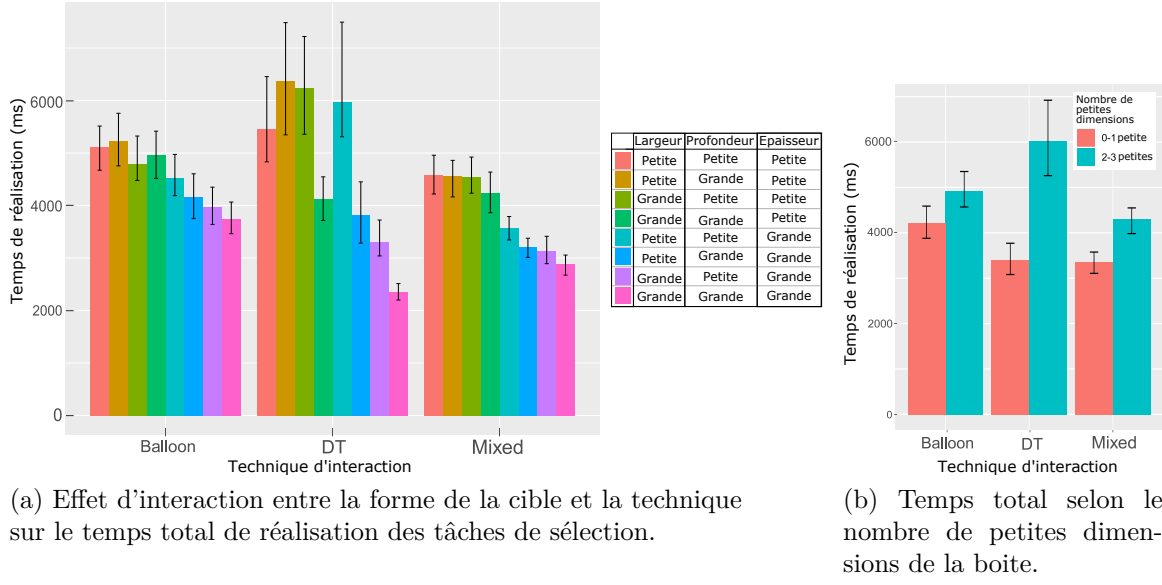


FIGURE 8.11 – Temps total de sélection selon la forme de la boîte.

3,4 s) et les quatre formes avec 2 ou 3 petites dimensions ($m_{2-3petites} = 6$ s). La Figure 8.11b présente le temps moyen de réalisation de la sélection en fonction du nombre de petites dimensions (largeur, épaisseur, profondeur) de la boîte ciblée. Les techniques *DT* et *Mixed* ($m_{0-1petite} = 3,4$ s, $m_{2-3petites} = 4,3$ s) sont statistiquement plus lentes avec des formes ayant au moins 2 petites dimensions ($p < 0,008$). Nous n'observons pas de différence significative entre ces deux groupes pour *Balloon* ($m_{0-1petite} = 4,2$ s, $m_{2-3petites} = 4,9$ s).

Pour les formes ayant au plus une petite dimension, *Mixed* et *DT* sont statistiquement plus rapides que *Balloon* ($p < 0,03$). Aucune différence statistiquement significative n'est observée entre *Mixed* et *DT*. Pour les formes ayant au moins 2 petites dimensions, *Mixed* est statistiquement plus rapide ($p = 0,0004$) que *DT*. Aucune différence statistiquement significative n'est observée entre *Balloon* et les deux autres techniques.

En agrégeant les formes par épaisseur (Figure 8.12a), nous observons que seules les deux techniques *Mixed* ($m_{petite} = 4,5$ s et $m_{grande} = 3,2$ s) et *DT* ($m_{petite} = 5,5$ s et $m_{grande} = 3,9$ s) utilisant l'interaction dans l'air sont statistiquement impactées par l'épaisseur de la cible (pour *Balloon* : $m_{petite} = 5$ s et $m_{grande} = 4,1$ s). Les comparaisons par paires montrent qu'elles sont statistiquement plus rapides quand les cibles sont épaisses ($p < 0,0001$).

Enfin, en agrégeant les formes par largeur et par profondeur (Figures 8.12b & 8.12c), nous observons un effet statistiquement significatif de ces deux dimensions pour *DT*

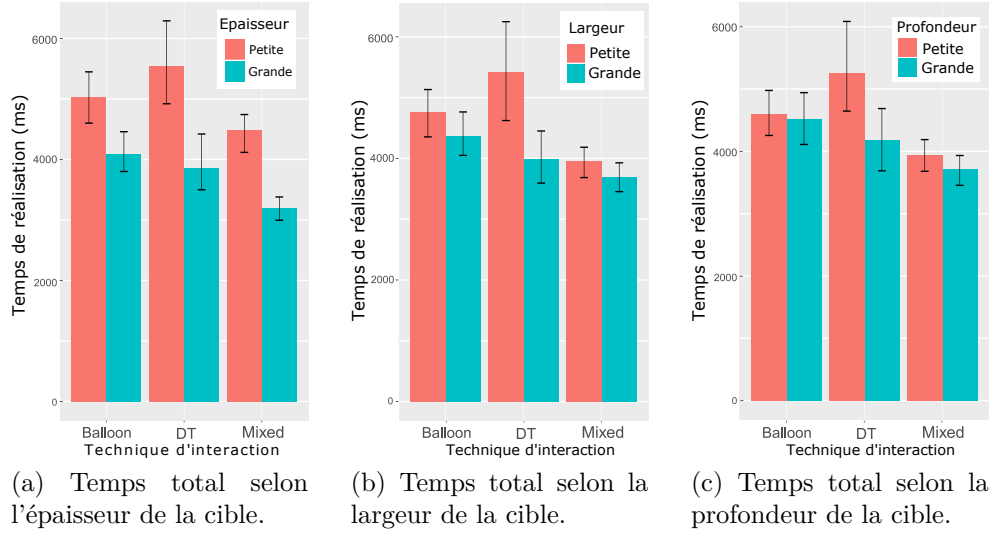


FIGURE 8.12 – Temps total selon les trois dimensions de la cible.

(pour la largeur : $p = 0,001$, $m_{petite} = 5,4$ s et $m_{grande} = 4$ s ; pour la profondeur : $p = 0,022$, $m_{petite} = 5,2$ s et $m_{grande} = 4,2$ s). Nous n'observons pas d'effet significatif de ces dimensions sur *Mixed* (pour la largeur : $m_{petite} = 4$ s et $m_{grande} = 3,7$ s ; pour la profondeur : $m_{petite} = 3,9$ s et $m_{grande} = 3,7$ s), ni sur *Balloon* (pour la largeur : $m_{petite} = 4,8$ s et $m_{grande} = 4,4$ s ; pour la profondeur : $m_{petite} = 4,6$ s et $m_{grande} = 4,5$ s).

Effet d'interaction HEIGHT \times TECH. Nous observons un effet d'interaction statistiquement significatif entre HEIGHT et TECH [$F_{4,56} = 3,26$, $p = 0,018$] sur le temps total de réalisation. Les résultats sont présentés à la Figure 8.13.

Le temps nécessaire pour atteindre une boîte dans la pile pour la première fois avec *DT* est presque le même quelle que soit la hauteur de la boîte ($m_{H1} = 1,9$ s, $m_{H2} = 2,0$ s, $m_{H3} = 1,8$ s). Bien que les deux techniques impliquant la décomposition de la tâche, *Mixed* ($m_{H1} = 1,9$ s, $m_{H2} = 2,4$ s, $m_{H3} = 2,5$ s) et *Balloon* ($m_{H1} = 2,4$ s, $m_{H2} = 3,3$ s, $m_{H3} = 4$ s), semblent être plus rapides pour atteindre une boîte à la hauteur 1 que les autres hauteurs, nous observons une différence statistiquement significative seulement pour *Balloon* et entre les hauteurs 1 et 3 ($p < 0,0001$).

Pour la hauteur 1, nous n'observons pas de différence significative entre les techniques. Pour la hauteur 2, nous observons une différence significative uniquement entre *Balloon* et *DT* ($p = 0,0008$). Enfin pour la hauteur 3, *Balloon* est plus lent pour atteindre la cible que les deux autres techniques ($p < 0,0005$).

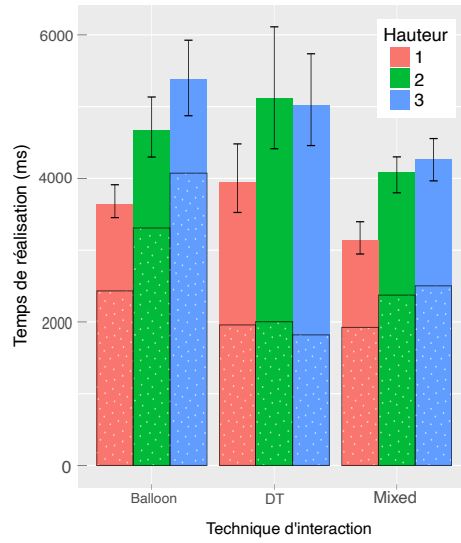


FIGURE 8.13 – Effet d'interaction entre la hauteur de la cible et la technique sur le temps total de réalisation des tâches de sélection. Le temps moyen pour atteindre les cibles pour la première fois est indiqué en pointillé.

Pour le temps total de réalisation, les cibles en hauteur 1 sont celles sélectionnées le plus rapidement pour toutes les techniques (*DT* : $m_{H1} = 4$ s, $m_{H2} = 5,1$ s, $m_{H3} = 5$ s ; *Mixed* : $m_{H1} = 3,1$ s, $m_{H2} = 4,1$ s, $m_{H3} = 4,3$ s ; *Balloon* : $m_{H1} = 3,6$ s, $m_{H2} = 4,7$ s, $m_{H3} = 5,4$ s). Les comparaisons par paires montrent que les cibles en hauteur 1 sont sélectionnées plus rapidement que celles en hauteur 2 ($p < 0,05$) quelle que soit la technique utilisée. Enfin, les cibles en hauteur 1 sont sélectionnées plus rapidement que celles en hauteur 3 ($p \leq 0,003$) pour les techniques utilisant la décomposition de la tâche (*Mixed* et *Balloon*). Nous n'observons pas de différence statistiquement significative entre les hauteurs 2 et 3 pour aucune des trois techniques. Quelle que soit la hauteur de la cible, nous n'observons pas de différence statistiquement significative entre les techniques sur le temps total de sélection.

Précision de la tâche de sélection

Le nombre d'erreurs est le nombre de fois où le minuteur est remis à zéro pendant la phase de validation de la sélection (c'est-à-dire le nombre de fois où la cible est perdue pendant le temps de validation de la sélection).

Nous observons un effet statistiquement significatif de la technique TECH [$F_{2,994} = 308,3$, $p < 0,0001$], de la forme de la cible SHAPE [$F_{7,994} = 150$, $p < 0,0001$] et de la hauteur de la cible dans la pile HEIGHT [$F_{2,994} = 122$, $p < 0,0001$] sur le nombre moyen d'erreurs. Le nombre moyen d'erreurs est de 0,5 erreur pour *Balloon*, 0,9 erreur pour *Mixed*, et 2,2 erreurs pour *DT*. Les comparaisons par paires montrent que les participants font statistiquement plus d'erreurs avec *DT* qu'avec *Balloon*

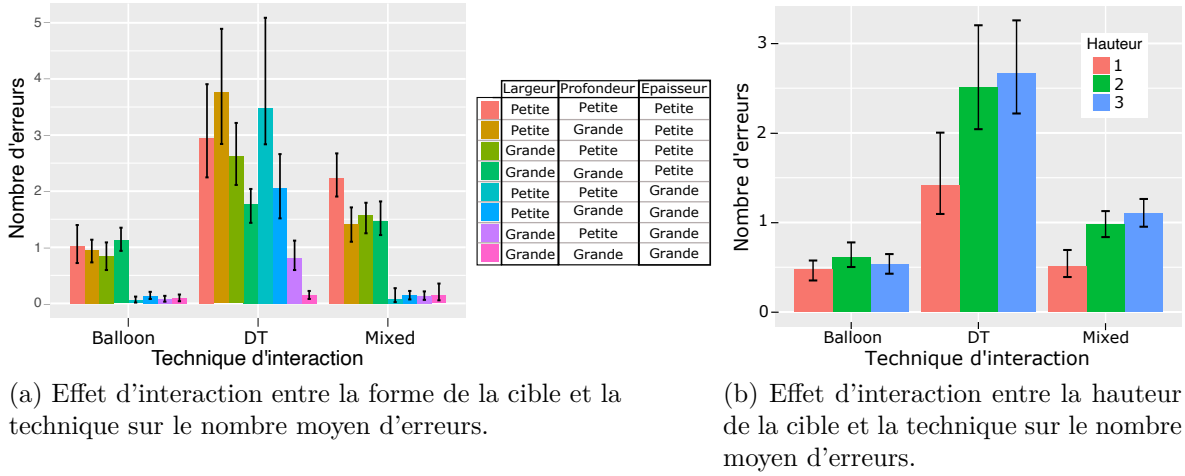


FIGURE 8.14 – Effet d'interaction entre la hauteur et la forme de la cible et la technique sur le nombre moyen d'erreurs avant validation de la cible.

($p = 0,0002$) et *Mixed* ($p = 0,003$). Aucune différence statistiquement significative n'est observée entre *Balloon* et *Mixed*.

Effet d'interaction SHAPE \times TECH. Nous observons un effet d'interaction statistiquement significatif entre SHAPE et TECH [$F_{14,994} = 31,1$, $p < 0,0001$] sur le nombre d'erreurs. Les résultats sont présentés à la Figure 8.14a.

Pour *DT*, les résultats montrent de grandes variations entre les formes. En particulier, nous observons une grande différence sur le nombre moyen d'erreurs entre deux groupes de formes pour cette technique ($p = 0,0009$) : les quatre formes avec 0 ou 1 petite dimension ($m_{0-1petite} = 1,2$ erreur) et les quatre formes avec 2 ou 3 petites dimensions ($m_{2-3petites} = 3,2$ erreurs). Nous observons aussi une différence statistiquement significative entre ces deux groupes pour *Mixed* ($p = 0,0009$, $m_{0-1petite} = 0,5$ erreur, $m_{2-3petites} = 1,3$ erreur) et *Balloon* ($p = 0,02$, $m_{0-1petite} = 0,4$ erreur, $m_{2-3petites} = 0,7$ erreur).

Grâce à la décomposition de la tâche, seule l'épaisseur de la boîte impacte les techniques *Mixed* ($p = 0,003$, $m_{petite} = 1,7$ erreur, $m_{grande} = 0,1$ erreur) et *Balloon* ($p = 0,003$, $m_{petite} = 1$ erreur, $m_{grande} = 0,09$ erreur).

Effet d'interaction HEIGHT \times TECH. Nous observons un effet d'interaction statistiquement significatif entre HEIGHT et TECH [$F_{4,994} = 35,6$, $p < 0,0001$] sur le nombre d'erreurs. Les résultats sont présentés à la Figure 8.14b. Nous n'observons pas de différence statistiquement significative entre les hauteurs pour *Balloon* ($m_{H1} = 0,5$ erreur, $m_{H2} = 0,6$ erreur, $m_{H3} = 0,5$ erreur). Les comparaisons par

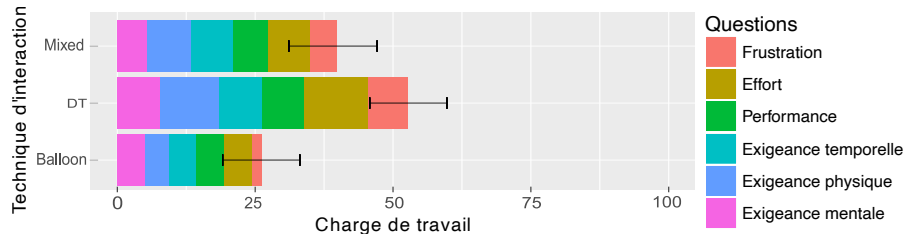


FIGURE 8.15 – [RTLX] Charge de travail totale entre 0 et 100.

paires montrent que les techniques *DT* ($m_{H1} = 1,4$ erreur, $m_{H2} = 2,5$ erreurs, $m_{H3} = 2,7$ erreurs) et *Mixed* ($m_{H1} = 0,5$ erreur, $m_{H2} = 1$ erreur, $m_{H3} = 1,1$ erreur) font statistiquement moins d'erreurs lorsque les cibles sont placées en hauteur 1 que pour les autres hauteurs ($p < 0,03$). Ceci peut être expliqué par le fait qu'en hauteur 1, les participants peuvent poser leur main sur la table et sont donc plus stables pour valider la sélection. Nous n'observons pas de différence statistiquement significative entre les hauteurs 2 et 3 pour ces deux techniques.

Lorsque les cibles sont à la hauteur 1, les participants font statistiquement moins d'erreurs avec *Mixed* et *Balloon* qu'avec *DT* ($p < 0,03$). Nous n'observons pas de différence significative entre *Mixed* et *Balloon* pour cette hauteur. Pour les hauteurs 2 et 3, les participants font statistiquement moins d'erreurs avec *Balloon* qu'avec *Mixed* et *DT* ($p = 0,002$). Ils font également moins d'erreurs avec *Mixed* qu'avec *DT* ($p < 0,004$).

8.3.2 Résultats qualitatifs

Charge de travail

Les résultats obtenus à partir du questionnaire RTLX (Annexe H) sont reportés à la Figure 8.15. Les participants estiment que la technique *Balloon* nécessite la charge de travail la plus faible ($26,1/100 \pm 7,9$). La technique *Mixed* présente une charge de travail de $39,1/100 \pm 8,8$. Enfin, la technique *DT* présente la charge de travail la plus élevée avec $52,8/100 \pm 7,8$.

En particulier, les participants estiment que la technique *Balloon* est beaucoup moins frustrante (4,3 fois moins), qu'elle demande moins d'efforts (2,3 fois moins) et qu'elle est moins exigeante physiquement (2,3 fois moins) que la technique *DT*. Les participants estiment également que la technique *Mixed* est moins frustrante (1,5 fois moins), moins exigeante physiquement (1,3 fois moins) et qu'elle demande moins d'efforts (1,5 fois moins) que la technique *DT*.

Classements des techniques

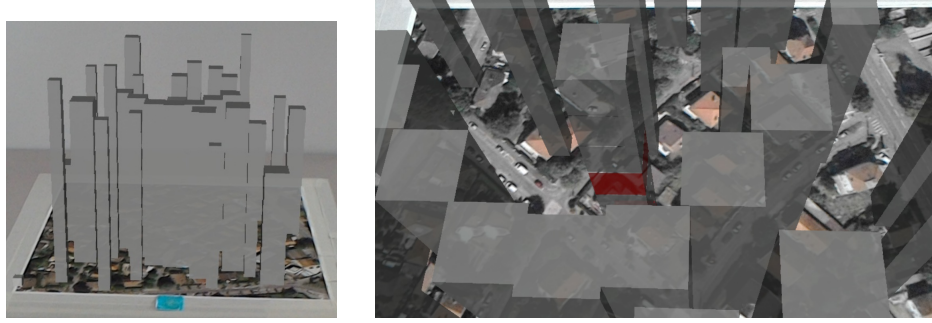
Le questionnaire final (Annexe I) confirme les résultats du questionnaire RTLX : les participants choisissent *Balloon* comme étant la technique la moins exigeante mentalement (*Balloon* : 8/11 participants ; *DT* : 4/15 participants ; *Mixed* : 3/15 participants), la moins exigeante physiquement (*Balloon* : 14/15 participants ; *Mixed* : 1/15 participants), la plus performante (*Balloon* : 13/15 participants ; *Mixed* : 2/15 participants), celle demandant le moins d'effort (*Balloon* : 14/15 participants ; *Mixed* : 1/15 participants) et enfin la moins frustrante (*Balloon* : 10/15 participants ; *Mixed* : 3/15 participants ; *DT* : 2/15 participants). *Balloon* est également choisit comme étant la technique la plus précise (*Balloon* : 15/15 participants) et la plus rapide (*Balloon* : 9/15 participants ; *Mixed* : 4/15 participants ; *DT* : 2/15 participants). *Balloon* est la technique préférée de tous les participants, et *DT* est la technique la moins préférée pour 11/15 participants.

Au cours de l'entretien, la majorité des participants expliquent préférer les techniques *Mixed* et *Balloon* à la technique *DT* car la décomposition de la tâche réduit selon eux considérablement la fatigue (mentale et physique), et surtout la frustration lorsque la tâche demande de la précision. De plus, ils choisissent *Balloon* plutôt que *Mixed* comme leur technique préférée en raison de la stabilité de ses performances quelle que soit la forme de la boîte (contrairement à *Balloon*, *Mixed* restant assez difficile à utiliser pour des boîtes peu épaisses).

Il est également intéressant de noter que, bien que les résultats montrent que *Mixed* est la technique la plus rapide en moyenne, la majorité des participants estiment que *Balloon* est la technique la plus rapide de l'étude.

Retours utilisateurs

Pendant l'entretien, nous avons demandé aux participants s'ils avaient ressenti un problème de perception de profondeur et d'estimation de la distance de la cible, et si oui, avec quelles techniques. Sur les 15 participants, 2 déclarent avoir eu ce type de problèmes avec les trois techniques, et 9 autres déclarent avoir eu des problèmes uniquement avec la technique *DT*. Pour les techniques décomposant la tâche, l'estimation de la profondeur n'est nécessaire que pour la sélection de la base de la pile. Pour sélectionner les petites bases (petite largeur et petite profondeur de boîtes), nous avons observé que la plupart des participants s'aidaient de la table : ils placent tout d'abord leur doigt sur la table près de la base et le font ensuite glisser jusqu'à ce qu'ils atteignent la pile.



(a) Densité de l'environnement autour de la pile cible.

(b) Vue au-dessus du mur.

FIGURE 8.16 – [Étude 2] Exemple d'une scène 3D densément peuplée avec occultation de la boîte cible.

8.4 Étude expérimentale 2 : sélection d'objets 3D occultés ancrés à la table dans un environnement dense

8.4.1 Objectif de l'étude

Cette seconde étude expérimentale complète la précédente en comparant les trois techniques dans un environnement dense et avec occultation des boîtes cibles. La Figure 8.16 montre un exemple d'une scène 3D présentée lors de cette étude.

Les principaux objectifs sont d'étudier l'impact de la densité de l'environnement et de l'occultation de la cible sur les préférences des utilisateurs. Nous voulons également observer comment les participants utilisent ces techniques dans des scènes denses. En effet, le fait de passer à travers des objets virtuels avec *Balloon* pendant l'étape 2 pourrait par exemple perturber les utilisateurs.

8.4.2 Système expérimental

Le système expérimental et les techniques d'interaction sont les mêmes que pour l'étude précédente (Section 8.2.2). Seules les scènes 3D diffèrent par leurs densités.

8.4.3 Tâche

Contrairement à la précédente étude, un mur virtuel composé de hauts distracteurs est inséré devant la pile de boîtes ciblée. La boîte cible étant occultée, les

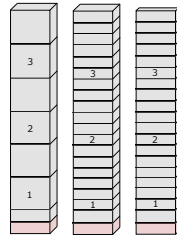


FIGURE 8.17 – Formes des cibles pour la seconde étude.

participants sont obligés de se pencher vers l'avant, au-dessus de ce mur, pour la voir (Figure 8.16b). Trente autres distracteurs sont placés sur la table (au-dessus de l'image cible imprimée). Tous les distracteurs sont totalement opaques. Les cibles sont toujours à la hauteur 2 afin d'éviter que les participants ne posent leur main sur la table pendant la phase de validation, comme cela a été observé lors de la précédente étude.

La tâche de sélection comprend les mêmes étapes que pour la précédente étude :

1. L'utilisatrice place son doigt en position initiale, dans le carré bleu illustré à la Figure 8.5.
2. Pour les techniques *Balloon* et *Mixed*, l'utilisatrice doit tout d'abord sélectionner la base de la pile. La sélection de la base est automatiquement validée lorsque la sphère qui suit son doigt est entièrement à l'intérieur de la base.
3. L'utilisatrice doit ensuite sélectionner la boîte cible rouge le plus rapidement possible. Pour valider la sélection, l'utilisatrice doit garder la sphère entièrement dans la cible pendant 1 seconde.

8.4.4 Protocole

L'étude expérimentale est conçue comme une étude intra-sujet, avec les deux variables indépendantes suivantes :

- *La technique d'interaction utilisée (TECH)* : *DT*, *Balloon* et *Mixed* ;
- *La forme de la boîte ciblée (SHAPE)* : dans le but de réduire le nombre de variables, nous n'utilisons que 3 formes différentes pour cette étude, comme illustré à la Figure 8.17. En tenant compte des résultats de l'étude précédente sur l'impact des petites dimensions sur la performance des techniques, en particulier pour la technique *DT*, nous choisissons les 3 formes suivantes : 1 avec trois grandes dimensions (grande boîte), 1 avec seulement une petite dimension (boîte moyenne) et enfin 1 avec deux petites dimensions (petite boîte). Pour la forme moyenne, nous choisissons une petite épaisseur afin d'influencer

les trois techniques. Pour la petite boîte, nous choisissons une petite épaisseur et une petite profondeur.

L'ordre des techniques est contrebalancé par un carré latin.

Les participants réalisent une phase d'entraînement sur un ensemble de 6 essais avant chaque technique. Ils effectuent ensuite 36 tâches de sélection de cibles cachées en environnements denses par technique (12 essais pour chacune des 3 formes). Les participants remplissent également un questionnaire RTLX après chaque technique. Après avoir utilisé les trois techniques, les participants les classent par ordre de préférence. Enfin, l'étude se termine par une session qualitative. Tout d'abord, les utilisateurs sont invités à utiliser les mêmes techniques mais sans occultation de la cible : la densité de l'environnement est la même, seul le distracteur occultant la pile est enlevé (scènes avec une forte densité mais sans occultation de la cible). Les utilisateurs sont alors interrogés sur :

1. L'impact perçu de la densité de l'environnement lorsque la cible est visible ;
2. L'ordre de préférence des techniques dans le cas où l'environnement est dense mais que la cible est visible ;
3. Leur utilisation des techniques en fonction des conditions de visibilité.

L'étude a duré environ 40 minutes par participant.

Nous avons recruté 9 participants, âgés de 24 à 37 ans (moyenne = $29 \pm 3,57$). Tous ont participé à la première étude expérimentale (Section 8.2), avec un délai d'au moins 6 mois entre les deux études.

Nous considérons les mêmes mesures que dans l'étude précédente : deux mesures objectives (temps de réalisation, nombre d'erreurs), et deux mesures subjectives (charge de travail, préférence des utilisateurs).

8.5 Étude expérimentale 2 : résultats

8.5.1 Résultats quantitatifs

Temps de réalisation moyens

La Figure 8.18 présente le temps total moyen de réalisation d'une tâche de sélection pour chacune des trois techniques. Cette figure présente également le temps moyen nécessaire pour atteindre une cible pour la première fois, ainsi que le temps moyen de sélection de la base de la pile.

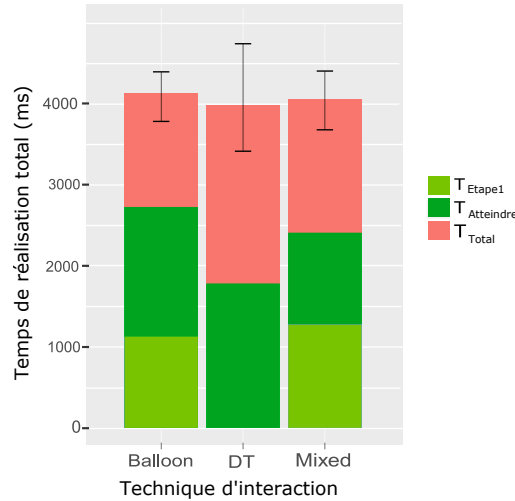


FIGURE 8.18 – Temps total de réalisation de la tâche (T_{Total}), temps pour atteindre la cible pour la première fois ($T_{Atteindre}$) et temps pour sélectionner la base (T_{Etape1}).

Temps moyens de sélection d'une boîte. Nous observons un effet statistiquement significatif de la forme de la cible SHAPE [$F_{2,16} = 137,1$, $p = 0,0001$] sur le temps total de réalisation de la tâche de sélection. Cependant, nous n'observons pas d'effet statistiquement significatif de la technique TECH. Le temps total moyen de réalisation de la tâche est de 4,08 s pour *Mixed*, 4,12 s pour *Balloon* et 3,8 s pour *DT*. Contrairement à la précédente étude, la technique *DT* présente des temps de réalisation similaires aux autres techniques. Une explication est que nous ne prenons pas en compte les boîtes ayant une petite largeur ou une petite profondeur, ce qui a eu un impact significatif sur les performances de la technique *DT* uniquement.

Le temps moyen pour sélectionner la base est de 1,2 s : 29% du temps total pour *Balloon* et *Mixed*. En comparant les temps de sélection de la base de la pile lors de l'étude précédente pour les trois formes étudiées dans cette seconde étude, nous observons que ce temps n'est que légèrement impacté par l'occultation de la cible : (sans occultation : $T_{Etape1} = 1,1$ s ; avec occultation : $T_{Etape1} = 1,2$ s).

Temps moyens pour atteindre une boîte. Nous observons un effet statistiquement significatif de la technique TECH [$F_{2,16} = 17,36$, $p < 0,0001$] et de la forme de la cible SHAPE [$F_{2,16} = 70,02$, $p < 0,0001$] sur le temps nécessaire pour atteindre la cible. En moyenne, les participants atteignent la cible pour la première fois en 1,8 s avec *DT*, en 2,4 s avec *Mixed* et en 2,7 s avec *Balloon*. Les comparaisons par paires montrent qu'il faut statistiquement moins de temps pour atteindre la cible avec *DT* qu'avec *Balloon* ($p = 0,001$) et *Mixed* ($p = 0,04$). Nous n'observons pas de différence statistiquement significative entre les deux techniques *Balloon* et *Mixed*.

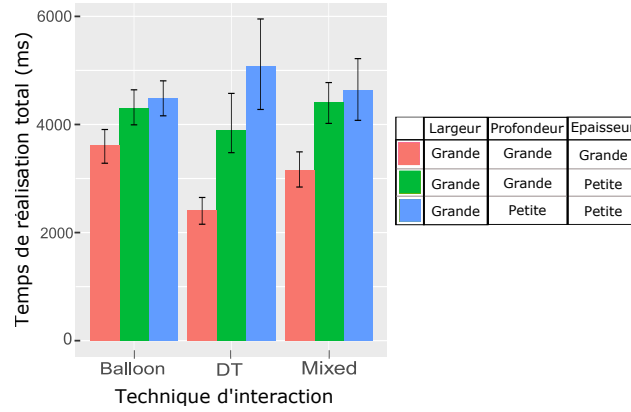


FIGURE 8.19 – Effet d'interaction entre la forme de la cible et la technique sur le temps total de réalisation des tâches de sélection.

Effet d'interaction SHAPE \times TECH. Nous observons un effet d'interaction statistiquement significatif entre SHAPE et TECH [$F_{4,32} = 15,49$, $p < 0,0001$] sur le temps total de réalisation. Les résultats sont présentés à la Figure 8.19.

Comme dans la précédente étude, nous observons de grandes variations sur les temps de réalisation de *DT* ($m_{petite} = 4,9$ s, $m_{moyenne} = 3,9$ s, $m_{grande} = 2,4$ s). Les techniques *DT* et *Mixed* ($m_{petite} = 4,7$ s, $m_{moyenne} = 4,5$ s, $m_{grande} = 3,2$ s) sont statistiquement plus rapides pour les grandes boîtes que pour les moyennes et petites boîtes ($p < 0,003$). Ce résultat confirme l'impact important de l'épaisseur de la boîte sur les deux techniques interagissant dans l'air. Nous n'observons pas de différence statistiquement significative entre les moyennes et les petites boîtes pour *DT* et *Mixed*. Pour *Balloon* ($m_{petite} = 4,5$ s, $m_{moyenne} = 4,3$ s, $m_{grande} = 3,6$ s), nous n'observons pas de différence statistiquement significative entre les 3 formes de boîtes.

Pour les grandes boîtes, *DT* est statistiquement plus rapide que *Balloon* ($p = 0,0002$) et *Mixed* ($p = 0,005$). Nous n'observons pas de différence statistiquement significative entre les trois techniques pour les moyennes et petites boîtes.

Précision de la tâche de sélection

Nous observons un effet statistiquement significatif de la technique TECH [$F_{2,64} = 26,2$, $p < 0,0001$] et de la forme de la cible SHAPE [$F_{2,64} = 72,7$, $p < 0,0001$] sur le nombre moyen d'erreurs. Le nombre moyen d'erreurs est de 0,57 erreur pour *Balloon*, 1,1 erreur pour *Mixed*, et 1,5 erreur pour *DT*. Les comparaisons par paires montrent que les participants font statistiquement moins d'erreurs avec *Balloon* qu'avec *DT* et *Mixed* ($p = 0,01$).

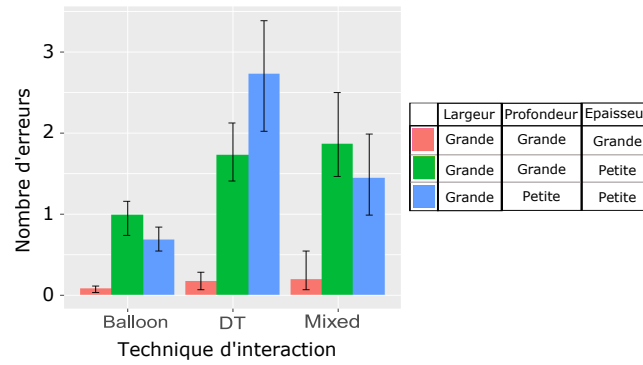


FIGURE 8.20 – Effet d'interaction entre la forme de la cible et la technique sur le nombre d'erreurs moyen de sélection.

Effet d'interaction SHAPE \times TECH. Nous observons un effet d'interaction statistiquement significatif entre SHAPE et TECH [$F_{4,64} = 7,75$, $p < 0,0001$] sur le nombre d'erreurs. Les résultats sont présentés à la Figure 8.20.

Les 3 techniques font statistiquement moins d'erreurs ($p < 0,02$) avec de grandes boîtes ($< 0,18$ erreur) qu'avec les autres boîtes (*Balloon* : $m_{petite} = 0,7$ erreur, $m_{moyenne} = 0,9$ erreur, $m_{grande} = 0$ erreur ; *Mixed* : $m_{petite} = 1,4$ erreur, $m_{moyenne} = 1,8$ erreur, $m_{grande} = 0,2$ erreur ; *DT* : $m_{petite} = 2,7$ erreurs, $m_{moyenne} = 1,7$ erreur, $m_{grande} = 0,2$ erreur). Entre les petites et les moyennes boîtes, nous observons une différence statistiquement significative seulement pour *DT* ($p = 0,01$).

Pour les grandes cibles, nous n'observons pas de différence significative entre les trois techniques. Pour les petites et moyennes cibles, *Balloon* est statistiquement plus précis que *DT* ($p < 0,04$). Enfin, nous n'observons pas de différence significative entre *Mixed* et les autres techniques.

8.5.2 Résultats qualitatifs

Charge de travail

Les résultats obtenus à partir du questionnaire RTLX (Annexe H) sont reportés à la Figure 8.21. Les participants estiment que la technique *Balloon* nécessite la charge de travail la plus faible ($20,8/100 \pm 10,1$). La technique *Mixed* présente une charge de travail de $33,7/100 \pm 6,6$. Enfin, la technique *DT* présente la charge de travail la plus élevée avec $51/100 \pm 13$ (Figure 8.21). *Balloon* est la technique préférée pour 7/9 participants, les 2 autres choisissent *Mixed*. *DT* est la technique la moins préférée pour 7/9 participants (les 2 autres choisissent *Mixed* et *Balloon*).

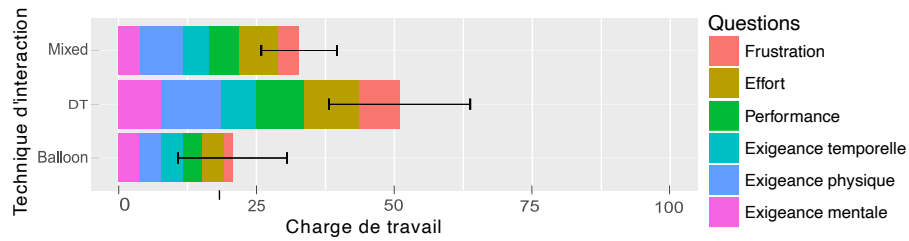


FIGURE 8.21 – [RTLX] Charge de travail totale entre 0 et 100.

Retours utilisateurs

Pour les trois techniques, les participants doivent se pencher en avant et changer de point de vue afin de voir la pile (Figure 8.16a). Aucun consensus ne se dégage sur la technique la moins impactée par l'occultation (*Balloon* : 5/9 participants, *DT* : 4/9 participants) et celle étant la plus impactée par l'occultation (*DT* : 4/9 participants, *Mixed* : 3/9 participants, *Balloon* : 2/9 participants). Selon les participants, l'occultation a un faible impact sur la difficulté de la tâche. Pour *Balloon* et *Mixed*, l'occultation augmente légèrement le temps de sélection de la base de 0,12 seconde. À l'exception des changements de point de vue, les participants rapportent utiliser toutes les techniques de la même manière que les cibles soient visibles ou occultées. Par exemple, ils passent directement à travers les obstacles pour atteindre les boîtes au lieu de passer au-dessus du mur de distracteurs (comportement aussi observé dans l'expérimentation décrite au Chapitre I, Section 2.5). Avec une forte densité et des cibles visibles, 8/9 participants rapportent un impact nul ou très faible de la densité sur la façon d'utiliser les techniques. Un seul participant a déclaré avoir été gêné par les distracteurs positionnés à côté de la cible pour l'interaction tactile avec *Balloon*. Ce participant rapporte une gêne causée par l'occultation de sa main pendant l'étape 2. Pour une grande majorité des participants, les techniques de décomposition des tâches sont préférées, quelles que soient les conditions de visibilité et de densité.

8.6 Études expérimentales 1 et 2 : discussion

8.6.1 Confirmation de résultats existants

La technique *DT* est performante pour des grandes cibles, bien moins précise que les deux autres techniques pour les plus petites cibles.

La technique de sélection dans l'air *DT* est une technique intuitive et facile à

comprendre. Les résultats de cette étude montrent que cette technique est rapide pour atteindre la cible quelle que soit sa taille. Elle est également efficace pour sélectionner des grandes cibles et des cibles proches de la table. En effet, une utilisatrice peut poser sa main sur la table lors de la phase de validation : ceci lui apporte alors une meilleure stabilité.

Cette étude met également en évidence les limitations de la technique *DT* en termes de précision : elle est peu performante pour la sélection de petites cibles, c'est-à-dire des cibles avec au moins deux petites dimensions. Le manque de précision de cette technique est particulièrement visible pour les cibles situées en hauteur par rapport à la table car l'utilisatrice ne peut alors plus utiliser la table comme support. La faible précision de *DT* confirme donc la première hypothèse *H1*.

Enfin, 9 des 15 participants rapportent avoir eu des problèmes d'estimation de la distance de la cible avec cette technique alors que seulement 2 participants rapportent ce type de problème avec les techniques décomposant la tâche.

La technique *Balloon* utilisant uniquement l'interaction tactile est moins fatigante et plus précise que la technique *DT*.

L'étude de Benko et al. [19] montre une meilleure précision (taux d'erreurs plus faible) des participants avec la technique *Balloon* qu'avec la technique d'interaction 3D dans l'air *Wand*. De plus, l'étude de Veit et al. [165] ainsi que l'étude de [19], ont constaté que les techniques d'interaction tactile étaient moins fatigantes que les techniques d'interaction dans l'air grâce au support de la table.

Notre étude confirme ces deux résultats. Tout d'abord, les participants font en moyenne 4 fois moins d'erreurs avec la technique *Balloon* qu'avec la technique *DT* (confirmation des résultats de Benko et al. [19]). De plus, les participants estiment que la technique *Balloon* est beaucoup moins frustrante, demande moins d'efforts et est moins exigeante physiquement que la technique *DT*. Les participants estiment aussi qu'elle demande une charge de travail bien moins importante que la technique *DT* (confirmation des résultats de Veit et al. [165] et de Benko et al. [19]).

8.6.2 Nouveaux résultats : avantages et limitations de la décomposition de la tâche

Les techniques décomposant la tâche sont plus précises sans être systématiquement plus rapides que la technique *DT*.

Bien que la technique *Balloon* soit plus lente que la technique *DT* pour atteindre

une boîte pour la première fois, sa précision permet à une utilisatrice de sélectionner une boîte avec un temps total moyen similaire à celui de la technique *DT*. De plus, nous avons recueilli de nombreux commentaires positifs sur la technique *Balloon*. Par exemple, les participants préfèrent largement cette technique (contrairement aux résultats de Benko et al. [19] qui exprimaient une préférence des participants pour la technique 3D). Elle est également la moins fatigante des techniques comparées et est toujours efficace et appréciée dans un environnement dense alors que la technique impose d'interagir à travers des obstacles. Ces avantages font de la technique *Balloon* un bon candidat pour des tâches de sélection d'objets 3D ancrés à la table, particulièrement lorsque la précision est requise et pour un usage prolongé.

Contredisant les résultats des études [106] et [165] sur la décomposition de la tâche, la technique *Mixed* est en moyenne la technique de sélection la plus rapide de notre étude : son temps pour atteindre les cibles est similaire à celui de la technique *DT*, tandis que la décomposition de la tâche impliquée par *Mixed* permet à une utilisatrice de contrôler une seule dimension dans l'air, ce qui la rend bien plus précise que *DT* (2,75 fois moins d'erreurs en moyenne).

Notre seconde hypothèse *H2* supposait que les deux techniques décomposant la tâche étaient plus performantes que la technique impliquant un contrôle simultané des trois dimensions. Les résultats de cette étude confirment partiellement cette hypothèse car seule la technique *Mixed* est plus rapide que la technique *DT*.

La décomposition de la tâche avec interaction dans un espace hybride constitue un bon compromis.

Tout d'abord, notre étude révèle des résultats prometteurs des deux techniques utilisant l'interaction tactile pour la sélection d'objets ancrés à la table. Nous concluons que la sélection 2D par interaction au toucher direct sur la table est efficace. Avec des objets flottant dans l'air, sans lien avec la table, la tâche 2D n'est pas évidente, l'utilisatrice doit placer son doigt sur la table et sous la cible, mais elle ne sait pas exactement où. Au contraire, quand les objets sont ancrés à la table, la tâche 2D est évidente (par exemple, sélection de l'ombre projetée sur la table ou, dans notre cas, de la base de la pile). Cette tâche 2D est également rapide, quelles que soient la taille de la base de pile, la visibilité des cibles et la densité de l'environnement.

Nous avons également observé l'avantage du support tangible qu'offre la table lors de la sélection de petites bases. Les participants plaçaient rapidement leur doigt proche de la base puis ajustaient sa position pour atteindre la base par glissement du doigt sur la table. Ce comportement pourrait aussi expliquer pourquoi les par-

ticipants ressentaient peu de problèmes de perception de profondeur (estimation de la distance de la pile de boîtes) avec les techniques décomposant la tâche.

Enfin, pour la deuxième étape de la décomposition de la tâche, nous observons un ajustement rapide de la hauteur de la sélection via l'interaction dans l'air avec la technique *Mixed*. En outre, cette technique est légèrement impactée par la hauteur des cibles, contrairement à la technique *Balloon*.

9. Résumé de l'étude sur les trois espaces d'interaction pour la sélection d'objets proches de l'utilisatrice

Dans ce chapitre, nous avons mis en évidence trois espaces d'interaction en entrée permis par les systèmes de RA sur table :

- L'espace 3D dans l'air au-dessus de la table.
- L'espace 2D de la table.
- L'espace hybride combinant les deux premiers.

Nous avons tout d'abord présenté l'état de l'art (Section 7) des techniques de sélection et de positionnement 3D interagissant dans chacun de ces trois espaces d'interaction. Alors que les techniques d'interaction évoluant dans l'espace 3D contrôlent les trois dimensions simultanément, les deux autres espaces impliquent généralement de décomposer la tâche. Le choix de l'espace d'interaction en entrée est donc un choix de conception important car il impacte considérablement la manière de sélectionner un objet.

L'état de l'art a également présenté des études comparant ces différents espaces. Dans le cas de la sélection d'objets 3D, nous avons constaté que ces études comparatives considéraient uniquement des objets flottant dans l'air, sans lien avec la table. De plus, nous avons constaté qu'aucune étude n'avait comparé l'interaction 3D dans l'air et l'interaction tactile sur la table à une interaction évoluant dans un espace hybride.

Ainsi, l'étude présentée à la Section 8 vise à comprendre les avantages et limitations de ces trois espaces d'interaction en entrée et l'impact de la décomposition de la tâche pour la sélection d'objets 3D ancrés à la table. Pour cela, nous avons comparé trois techniques de sélection (*DT*, *Balloon* et *Mixed*), chacune interagissant dans un de ces trois espaces. Nous avons alors étudié les effets de l'espace d'interaction et de la décomposition de la tâche sur les performances par deux études expérimentales : une étude avec des environnements épars (Section 8.2) et l'autre avec des environnements denses (Section 8.4). Le Tableau 9.1 résume les résultats obtenus, mettant en évidence les avantages et les limites des trois techniques.

Nous rappelons les résultats clés à considérer lors de la conception d'un système de RA pour la sélection d'objets 3D à portée de main :

- Lorsque la tâche nécessite une très grande précision ou que l'usage anticipé est long, nous proposons d'utiliser des techniques d'interaction tactile telles que *Balloon* ;
- Les bonnes performances de la technique *Mixed* et les retours des participants nous encouragent à unifier les espaces d'interaction 2D et 3D. En effet, des techniques de sélection comme *Mixed* combinent avec succès la précision de l'interaction tactile et la rapidité de l'interaction dans l'air.

	Bénéfices	Limitations
<i>Espace d'interaction 2D</i> <i>Interaction au toucher (Balloon)</i>	<ul style="list-style-type: none"> • Technique la plus précise ✓ • Temps de réalisation moyen similaire à celui de <i>DT</i> ✓ • Sélection 2D sur la table rapide, quelle que soit la condition de visibilité • Technique peu impactée par la taille de la cible • Technique préférée • Impression de contrôle • Support physique constant • Faible charge de travail perçue • Peu de fatigue ressentie ✓ 	<ul style="list-style-type: none"> • Technique plus lente que <i>Mixed</i> pour sélectionner une cible • Technique la plus lente pour atteindre une cible • Technique impactée par la hauteur de la cible
<i>Espace d'interaction 2D+3D</i> <i>Interaction hybride (Mixed)</i>	<ul style="list-style-type: none"> • Technique la plus rapide en moyenne ✗ • Technique plus rapide que <i>Balloon</i> pour sélectionner des grandes cibles • Technique plus rapide que <i>DT</i> pour sélectionner des petites cibles • Technique plus précise que <i>DT</i> • Sélection 2D sur la table rapide, quelle que soit la condition de visibilité • Technique précise et rapide pour sélectionner des cibles proches de la table (main reposant sur la table) 	<ul style="list-style-type: none"> • Technique moins précise que <i>Balloon</i> • Technique impactée par l'épaisseur de la cible (dimension selon l'axe Z)
<i>Espace d'interaction 3D</i> <i>Interaction dans l'air (DT)</i>	<ul style="list-style-type: none"> • Technique rapide pour atteindre une cible, quelle que soit sa hauteur • Technique plus rapide que <i>Balloon</i> pour sélectionner des grandes cibles • Technique intuitive • Technique précise et rapide pour sélectionner des cibles proches de la table (main reposant sur la table) 	<ul style="list-style-type: none"> • Technique la moins précise ✓ • Technique plus lente en moyenne que <i>Mixed</i> pour sélectionner des cibles • Technique impactée par les 3 dimensions de la cible • Forte charge de travail perçue • Fatigue physique ✓ • Frustration

Tableau 9.1 – Bénéfices et limitations des techniques. Nous indiquons aussi si un résultat confirme ✓ ou contredit ✗ ceux obtenus dans d'autres études de la littérature.

10. Mise en contexte : vers une installation artistique exploitant les trois espaces d'interaction disponibles en RA sur table

Dans le cadre de notre collaboration avec l'École Supérieure d'Art et Design (ESAD) de Valence, nous avons commencé à questionner le rapport entre l'Interaction Humain-Machine et l'Art. En particulier nous avons étudié l'usage de technologies interactives dans le domaine artistique et leurs appropriations par le grand public dans des lieux comme les galeries ou les musées. Par exemple, au milieu des années 2000, Bowers et al. [28] développent cette thématique à propos de l'usage de la réalité augmentée dans le cadre d'expositions.

Dans ce contexte, une expérimentation liant les aspects scientifiques et esthétiques a été conçue pour d'une part évaluer qualitativement les trois techniques d'interaction *DT*, *Balloon* et *Mixed* (Section 8) en dehors d'une expérimentation contrôlée en laboratoire, et d'autre part de porter une idée, un propos, et devenir ainsi une proposition esthétique. Nous avons réalisé plusieurs séances de travail avec un designer graphique et d'interaction et ancien étudiant de l'Esad • Valence. Nos séances avaient pour objectif de concevoir cette nouvelle expérimentation entre IHM, art et design interactif.

Présentation générale

Le point de départ et source d'inspiration est le livre *Flatland* publié par Edwin Abbott Abbott en 1884. Ce livre raconte l'histoire d'un monde en deux dimensions où évoluent différentes formes géométriques, telles que les lignes droites, les triangles, etc. Il y est raconté que des classes sociales composent ce monde : les femmes sont représentées par des lignes droites, les hommes de la classe populaire sont des triangles isocèles et ceux de la classe moyenne sont des carrés. La noblesse est caractérisée par le nombre de côtés de la forme géométrique, allant jusqu'au cercle qui représente la plus haute classe sociale de ce monde.

Nous avons imaginé une installation ludique (un jeu) avec pour objectif de peupler un monde 3D par différentes formes géométriques. À l'issue de ces sessions de travail, nous avons produit le *storyboard* illustré aux Figures 10.8 et 10.9. Ce storyboard présente les étapes clés du jeu interactif conçu. L'installation doit pouvoir

être utilisée sans aucune aide extérieure, le jeu débute donc par une présentation des techniques de sélection qui seront nécessaires au cours de la partie. Dans cette première partie du jeu, nous présentons les trois techniques étudiées à la Section 8. Dans la seconde partie du jeu, la joueuse est libre de choisir la technique qu'elle souhaite utiliser.

Déroulement du jeu

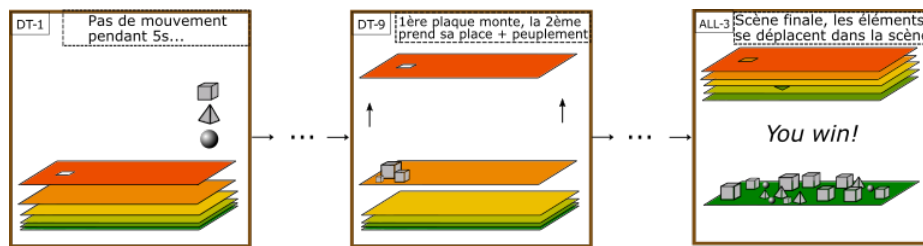


FIGURE 10.1 – Déroulement du jeu.

La scène virtuelle est composée d'un ensemble de plaques colorées placées les unes sur les autres. Chaque plaque est associée à un niveau. Toutes les plaques virtuelles sont au départ empilées en-dessous d'une table physique, la plaque du dessus étant à la surface de la table pour permettre l'interaction tactile (Figure 10.1 – vignette *DT-1*¹ : toutes les plaques sont empilées, la plaque rouge est la plaque en cours et est placée sur la table). La joueuse interagit toujours sur celle présente à la surface de la table.

Lorsque le niveau en cours est réussi, la plaque actuelle monte (Figure 10.1 – vignette *DT-9*) et une nouvelle plaque prend sa place sur la table : le niveau suivant commence alors. Les plaques, et donc les niveaux, se succèdent ainsi jusqu'à atteindre la dernière plaque (Figure 10.1 – vignette *ALL-3*).

Déroulement d'un niveau

À chaque nouveau niveau, trois éléments 3D (un cube, un tétraèdre et une sphère) sont placés en colonne sur le côté droit de la plaque (Figure 10.1 – vignette *DT-1*). Une ouverture est également présente dans la plaque et permet à un seul de ces trois éléments de passer à travers la plaque pour atteindre la suivante, située en-dessous. Pendant la présentation des techniques, la taille de ces trois éléments est fixe. Ces éléments sont de grande taille afin d'en faciliter la sélection. Dans la seconde partie du jeu, lorsque la joueuse est libre de choisir la technique qu'elle sou-

1. Pour chaque vignette, son nom est indiqué en haut à gauche.

haite utiliser, la taille des éléments varie à chaque niveau (variation de la difficulté) (Figure 10.9 – vignettes *ALL-1* et *ALL-2*).

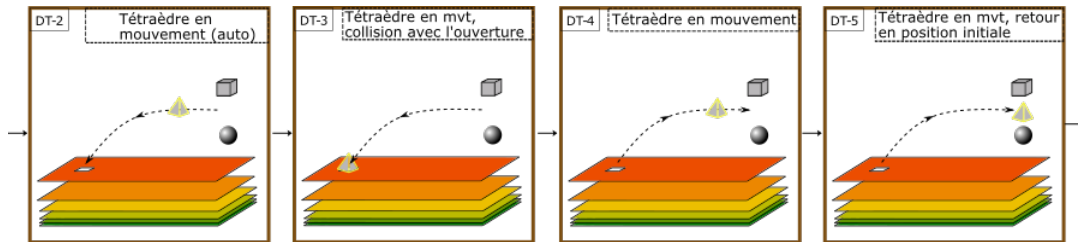


FIGURE 10.2 – Animation de l'élément sélectionné : mouvement automatique de l'élément vers l'ouverture puis remplacement de l'élément à sa position initiale.

Au début du niveau, un élément, autre que celui correspondant à l'ouverture, est présélectionné et effectue un mouvement automatique entre sa position initiale et l'ouverture. Cette animation, illustrée à la Figure 10.2, vise à expliquer la tâche à la joueuse. La joueuse doit ensuite sélectionner un élément. Lorsqu'un élément dont la forme ne correspond pas à celle de l'ouverture est sélectionné, il entre en collision avec l'ouverture (Figure 10.2 – vignette *DT-3*) et se replace automatiquement à sa position initiale (Figure 10.2 – vignette *DT-5*). Cette animation se répète avec cet élément tant que la joueuse n'a pas sélectionné un autre élément.

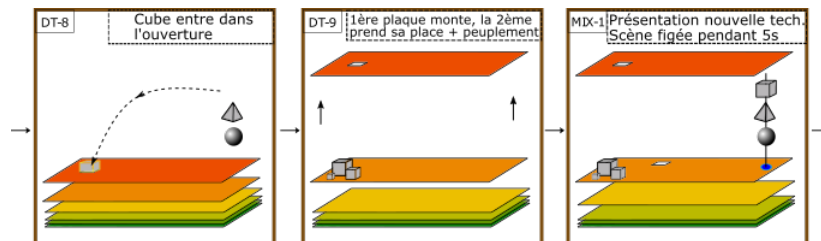


FIGURE 10.3 – Niveau complété avec succès : peuplement de la scène et passage au niveau suivant.

Pour passer au niveau suivant (Figure 10.3), la joueuse doit donc sélectionner l'élément parmi les trois proposés dont la forme correspond à l'ouverture de la plaque. Lorsque le bon élément tombe sur la plaque suivante (Figure 10.3 – vignette *DT-8*), la scène se peuple autour de l'élément : des éléments de la même forme que l'élément sélectionné sont placés autour de l'élément sur la plaque inférieure (Figure 10.3 – vignette *DT-9*).

Sélection d'un élément : explication des techniques d'interaction

Au début de chaque niveau, une sphère grise suit le bout de l'index de la joueuse.

Les trois techniques *DT*, *Mixed* et *Balloon* sont introduites au début du jeu comme suit :

- Technique *DT* (Storyboard Figure 10.8 – vignettes *DT-1* à *DT-9*) : Pour sélectionner un élément, la joueuse doit simplement placer son doigt dans celui-ci (Figure 10.4). Nous pensons que la joueuse va tenter d’approcher et de toucher les éléments dès le début du jeu. Mise à part la sphère grise qui suit le doigt de la joueuse, nous n’implémentons donc pas de retour visuel particulier pour cette technique.

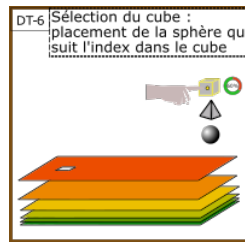


FIGURE 10.4 – Technique *DT*.

- Technique *Mixed* (Storyboard Figure 10.8 – vignettes *MIX-1* à *MIX-9*) : Une petite zone circulaire est affichée sur la table, sous les trois éléments. Un lien virtuel vertical (comme une ficelle) lie également cette zone aux éléments. Tout d’abord, la zone “pulse” afin d’attirer l’attention de la joueuse et la motiver à interagir avec. Une fois cette zone sélectionnée (la joueuse la touchant avec son doigt), elle change de couleur et devient rouge. De plus, la sphère qui suivait son doigt est placée au-dessus de cette zone et le lien virtuel est mis en évidence (Figure 10.5 – vignette *MIX-6*). La sphère ne peut alors que se déplacer le long du lien virtuel, sa hauteur étant la hauteur du doigt de la joueuse (Figure 10.5 – vignette *MIX-7*). Si la joueuse replace son doigt sur la zone (annulation du mode *Mixed*), la zone repasse en bleu, le lien n’est plus mis en évidence et la sphère se remet à suivre le doigt de la joueuse.

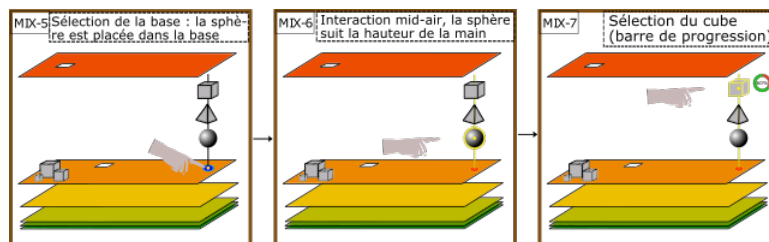
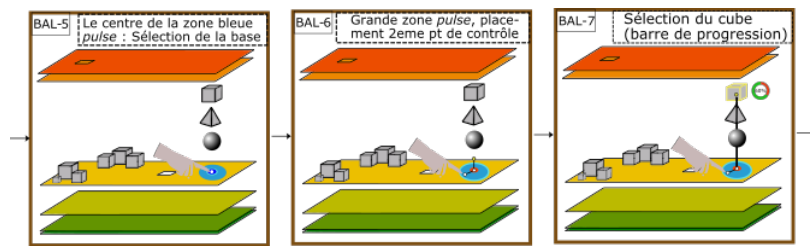


FIGURE 10.5 – Technique *Mixed*.

- Technique *Balloon* (Storyboard Figure 10.9 – vignettes *BAL-1* à *BAL-9*) : Une zone bleu foncé comme celle utilisée pour *Mixed* est affichée sur la table, sous les trois éléments. Nous ajoutons une seconde zone, bleu clair, autour de la précédente et contrairement à *Mixed*, il n'y a pas de lien virtuel entre les éléments et la zone. Cette seconde zone permet : (1) de différencier *Balloon* de *Mixed*, (2) de maintenir l'attention visuelle de la joueuse sur la table, après la sélection de la zone bleu foncé. La première étape de l'interaction est donc la même que pour *Mixed* : la petite zone pulse et la joueuse doit la sélectionner (Figure 10.6 – vignette *BAL-5*). Une fois sélectionnée, elle passe en rouge et une sphère grise est placée au-dessus. Une sphère suit toujours son doigt. C'est ensuite la seconde zone, la plus grande, qui pulse. Si la joueuse place son doigt sur la table, la ficelle virtuelle ainsi que le ballon apparaissent (Figure 10.6 – vignette *BAL-6*). Les pulsations de cette zone continuent tant que la joueuse n'a pas sélectionné le bon élément. La joueuse peut, comme avec *Mixed*, annuler le mode *Balloon* en sélectionnant à nouveau la petite zone.

FIGURE 10.6 – Technique *Balloon*.

La validation de la sélection d'un élément est faite par un mécanisme de temporisation (*dwel time*) de 1 seconde. Une barre de progression (Figure 10.6 – vignette *BAL-7*) est affichée près de l'élément et indique le temps restant. Lorsque les trois techniques ont été introduites, toutes les techniques et les retours visuels associés sont affichés sur le côté pour chaque plaque (Figure 10.7 – vignette *ALL-1*). La joueuse a le choix de la technique pour sélectionner un élément.

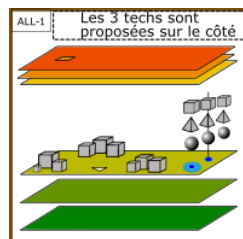


FIGURE 10.7 – Niveau libre : la joueuse a le choix de la technique pour la sélection d'un élément.

Étude expérimentale

Le mini-jeu n'a pas encore été implémenté. Lorsque l'installation sera réalisée elle fera l'objet d'une étude expérimentale. Contrairement aux études expérimentales présentées en Section 8, cette étude n'a pas pour objectif d'étudier les performances des joueurs. Il s'agit d'étudier la prise en main des techniques (apprentissage, émergence d'usages imprévus) et d'observer les comportements et réactions des joueurs lorsqu'ils les utilisent (émotions, intérêt, etc.). L'objectif est également d'observer si tous les niveaux sont réalisés ou si le jeu est rapidement abandonné par les joueurs et quelles sont les techniques privilégiées pendant les niveaux libres.

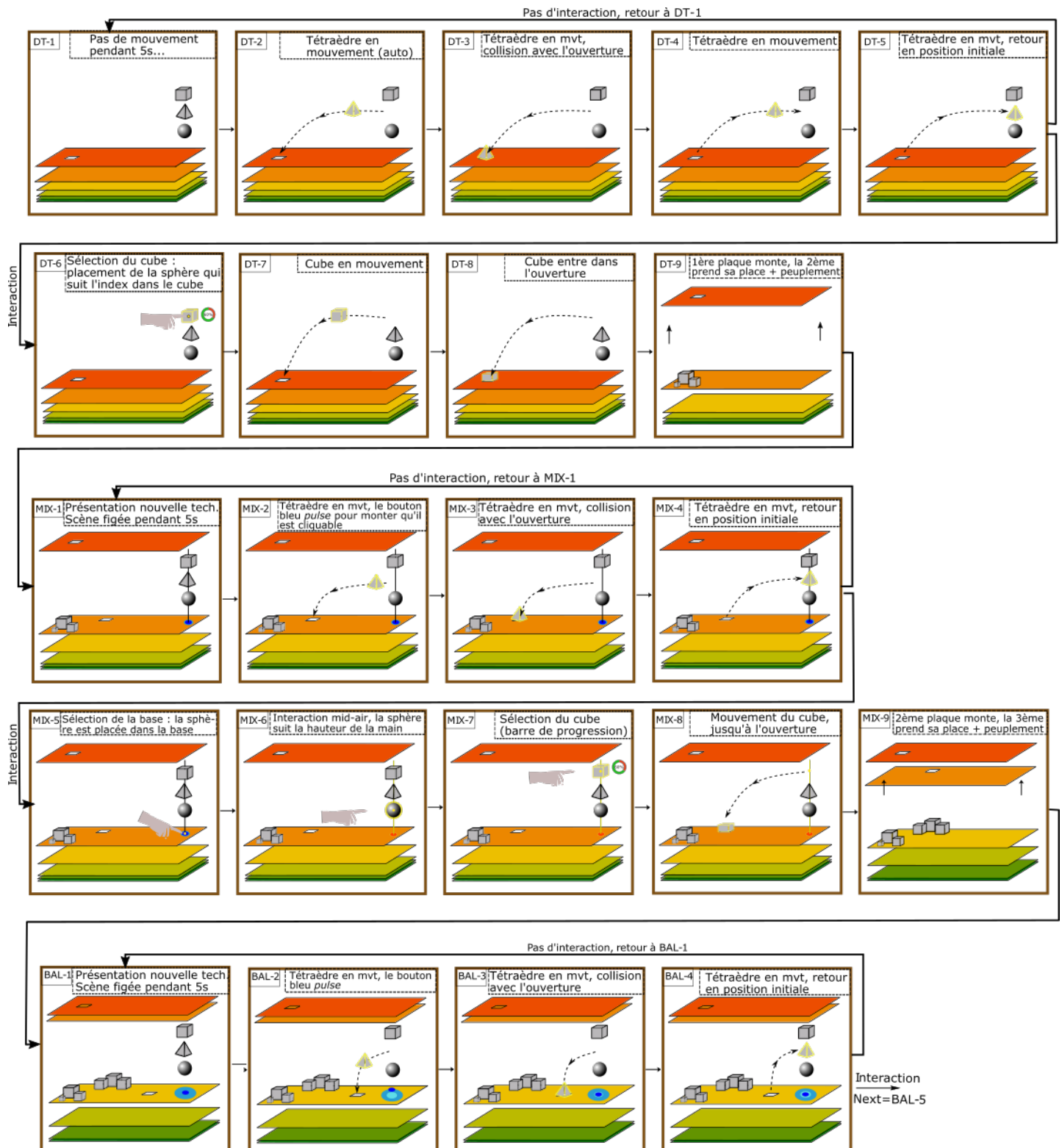


FIGURE 10.8 – Storyboard complet du jeu interactif inspiré du livre Flatland.

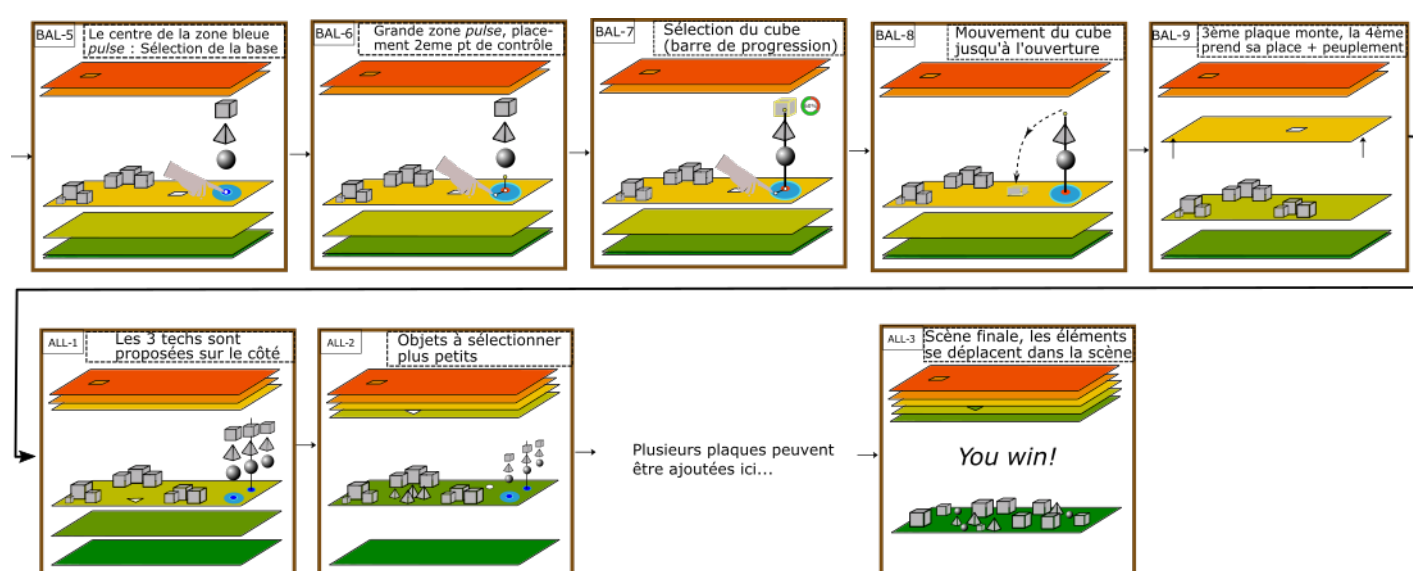


FIGURE 10.9 – Storyboard complet du jeu interactif inspiré du livre Flatland (suite).

Chapitre IV

Conclusion et perspectives

11. Contributions

Nos travaux sont dédiés à la réalité augmentée (RA) sur table. Celle-ci permet de visualiser des informations virtuelles 3D liées à une table qui sert de support physique à l'interaction. Les systèmes de RA sur table sont considérés dans de nombreux domaines tels que la modélisation 3D, la CAO et l'urbanisme. Avec ces systèmes, l'utilisatrice est régulièrement amenée à interagir avec des scènes virtuelles denses (un modèle 3D d'une ville par exemple) et avec des cibles plus ou moins distantes et de tailles variées (une maison, un étage de bâtiment). De plus, dans un contexte professionnel par exemple, ces systèmes peuvent impliquer une utilisation prolongée. Il est alors important de proposer aux utilisateurs des techniques d'interaction performantes, peu fatigantes et faciles à comprendre et à utiliser.

Nous nous sommes concentrés sur la tâche fondamentale de pointage d'objets 3D en RA sur table. La question de recherche de nos travaux concerne la conception de techniques de sélection d'objets 3D en RA sur table qui répondent à ces critères : précision, fatigue limitée, simplicité de prise en main et d'utilisation.

Cette thèse aborde cette question de recherche en explorant le vaste espace de possibilités offert par la RA sur table. Dans un premier temps (Chapitre I), nous avons exploré cet espace en nous concentrant sur le dispositif de RA qui prend en charge l'augmentation (dispositifs portables, casques de RA). La suite de nos travaux porte sur la sélection précise d'objets 3D en considérant la distance entre l'utilisatrice et les objets à sélectionner. Nous avons alors étudié la sélection d'objets distants d'une part (Chapitre II), et la sélection d'objets à portée de main d'autre part (Chapitre III).

Dans ce qui suit, nous résumons les trois principales contributions résultant de cette exploration des techniques d'interaction en RA sur table.

Étude sur les dispositifs de RA : dispositifs portables et casques de RA.

Nous avons tout d'abord mis en exergue deux classes de RA sur table : l'une basée sur un dispositif portable tel qu'un téléphone ou une tablette, et l'autre basée sur un casque de RA. Le choix du dispositif de RA utilisé pour l'affichage de l'augmentation 3D est important dans la conception d'un système de RA sur table car

il va impacter considérablement la manière de visualiser et d'interagir avec la scène virtuelle (interaction tactile sur l'écran du dispositif portable ou interaction dans l'air avec le casque par exemple). Notre première contribution a alors été d'identifier les avantages et inconvénients de ces deux dispositifs (tablette, casque de RA) et ainsi identifier le dispositif le plus approprié pour la visualisation et la sélection d'objets virtuels 3D en RA sur table. Pour cela, nous avons comparé expérimentalement trois techniques d'interaction couramment utilisées avec ces dispositifs : une technique d'interaction au toucher direct avec une tablette, et deux techniques avec un casque de RA implémentant soit la métaphore du pointeur virtuel soit la métaphore de la main virtuelle [5]. Deux études comparatives ont été menées sur des tâches de recherche et de sélection en environnements épars d'une part, et en environnements denses d'autre part. Les résultats de ces études montrent de bien meilleures performances pour le casque de RA, que ce soit pour la recherche ou la sélection d'objets 3D. Les participants préfèrent également les deux techniques avec casque qui sont, selon eux, plus simples et moins fatigantes que de déplacer et porter la tablette.

Compte-tenu des enseignements de cette étude introductive, nous nous sommes concentrés sur le casque de RA dans la suite de nos travaux. Nous avons alors étudié plus en profondeur les techniques d'interaction de RA sur table pour ce dispositif, en considérant : (1) les techniques dédiées à la sélection d'objets distants, et (2) les techniques dédiées à la sélection d'objets à portée de main de l'utilisatrice.

Conception d'une technique performante basée sur le raycasting pour la sélection d'objets distants.

Notre seconde étude porte sur la sélection d'objets 3D hors de portée de l'utilisatrice. Pour sélectionner de tels objets, les techniques d'interaction s'appuient généralement sur la métaphore du pointeur virtuel (rayon virtuel ou raycasting). Cependant, une implémentation standard du rayon virtuel se révèle inadaptée pour la sélection de petits objets, d'objets cachés et pour la sélection d'objets en environnements denses.

Ainsi, nous proposons *RayLens* une nouvelle technique efficace pour la sélection précise d'objets 3D distants. Cette technique bi-manuelle combine un rayon virtuel tenu par la main dominante et une lentille virtuelle grossissante déplaçable dans l'espace 3D par l'autre main. L'association de ces deux mécanismes (raycasting, lentille) offre de nombreux avantages. Tout d'abord, l'effet de zoom de la lentille

est indépendant des cibles (*target-agnostic*) et de la densité de l'environnement. Deuxièmement, le zoom facilite la sélection de petits objets. Enfin, le déplacement à distance de la lentille et le filtre de transparence appliqué sur tous les objets présents entre la lentille et l'utilisatrice permettent de pointer des objets cachés sans que l'utilisatrice n'ait besoin de se déplacer.

Cette technique se révèle performante en environnements denses, peu fatigante et particulièrement appréciée par les participants grâce à sa simplicité de prise en main et d'utilisation.

Étude des espaces d'interaction disponibles en RA sur table pour une sélection efficace d'objets à portée de main.

Notre troisième étude concerne la sélection d'objets à portée de main de l'utilisatrice et les améliorations possibles de la métaphore de la main virtuelle. La proximité des objets permet également l'utilisation de la table comme espace d'interaction en entrée et comme aide à la sélection. Ainsi, les techniques de sélection d'objets proches peuvent impliquer trois espaces d'interaction en entrée : l'espace 3D au-dessus de la table, l'espace 2D de la table, et un espace hybride combinant ces deux derniers. Nous avons mené une étude expérimentale comparative pour mieux comprendre les avantages et inconvénients de chacun de ces espaces.

Pour cette étude, nous avons considéré une technique de main virtuelle "standard". Cette technique, impliquant des mouvements dans l'espace 3D, est intuitive et directe mais les tremblements de la main la rendent peu précise et l'interaction dans l'air est fatigante lors d'un usage prolongé. Nous avons également considéré deux alternatives s'appuyant sur la présence de la table : une technique d'interaction tactile et une technique hybride combinant l'interaction tactile et l'interaction dans l'air.

Les résultats montrent tout d'abord l'importance du support physique (la table) pour réaliser des tâches de sélection d'objets 3D car il améliore la précision de l'interaction et réduit la fatigue. En effet les résultats montrent l'efficacité des techniques uniquement tactiles (espace 2D de la table) lorsque l'utilisatrice a besoin d'une très grande précision et pour une utilisation prolongée. Enfin, nous mettons en évidence l'apport de techniques hybrides combinant l'interaction tactile et l'interaction dans l'air : ces techniques définissent en effet un bon compromis entre précision et rapidité.

Publications

Les contributions de cette thèse ont fait l'objet des publications suivantes :

- Plasson, C., Cunin, D., Laurillau, Y., & Nigay, L. (2019, November). Tabletop AR with HMD and Tablet : A Comparative Study for 3D Selection. In Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces (pp. 409-414).
- Plasson, C., Cunin, D., Laurillau, Y., & Nigay, L. (2020, September). 3D Tabletop AR : A Comparison of Mid-Air, Touch and Touch+ Mid-Air Interaction. In Proceedings of the International Conference on Advanced Visual Interfaces (pp. 1-5).
- Un article décrivant une partie des travaux du Chapitre II est en cours d'évaluation.

12. Perspectives

Nos travaux sur les techniques d'interaction en RA sur table ouvrent la voie à plusieurs perspectives de recherche à court terme et à long terme.

12.1 Perspectives à court terme

12.1.1 Perspectives à court terme sur la technique *RayLens*

Mener une étude en contexte plus écologique. Cette étude a pour but d'étudier *RayLens* avec une scène non abstraite et avec des utilisateurs aux profils plus variés que ceux de notre étude expérimentale présentée en Section 5. Nous pourrions ainsi confronter les résultats de cette étude en contexte à ceux obtenus lors de l'étude expérimentale et quantifier la facilité de prise en main et d'utilisation de *RayLens*. La scène virtuelle a déjà été créée et est présentée à la Figure 12.1. Cette scène représente une ville composée de plusieurs bâtiments, de tailles variées et parfois occultés. Une utilisatrice devra sélectionner plusieurs types de bâtiments (le bâtiment cible est en rouge) : (1) de grands bâtiments non cachés pouvant être sélectionnés directement avec le rayon sans utiliser la lentille ; (2) des bâtiments cachés devant tout d'abord être trouvés avec la lentille pour être ensuite sélectionnés ; et enfin (3) de petits bâtiments non cachés pouvant être sélectionnés soit avec le rayon seul, soit en utilisant la lentille.

Nous devons réaliser cette étude à l'occasion des portes ouvertes 2021 de l'École

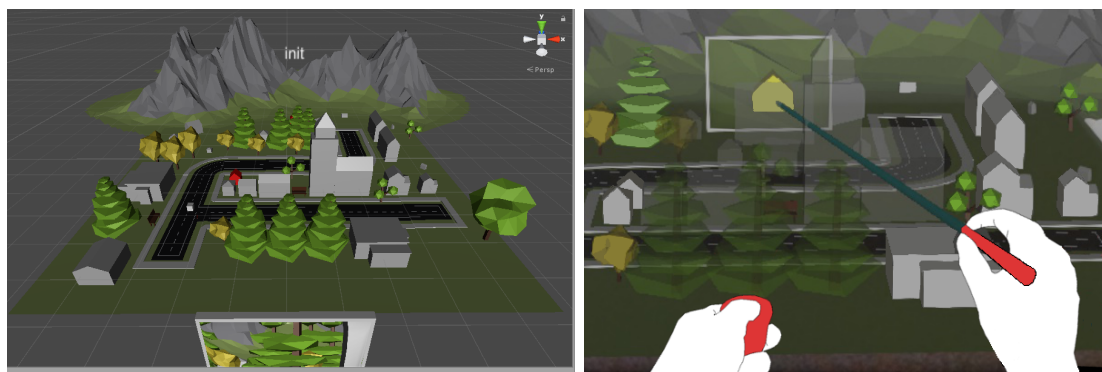


FIGURE 12.1 – Scène non abstraite et évaluation de *RayLens* dans un contexte grand public avec des utilisateurs aux profils variés.

Supérieure d'Art et Design de Valence. Malheureusement, l'évènement a été annulé à cause de la crise sanitaire.

Étude approfondie de l'interaction à deux mains. La technique *RayLens* est "asymétrique et dépendante" car les deux mains ont des rôles différents à jouer qui dépendent l'une de l'autre : l'action de la main dominante (qui contrôle le rayon) dépend de celle de la main non-dominante (qui contrôle la lentille) [78]. Selon Guiard [64], la main non-dominante pose le contexte pour la main dominante, son action précéderait donc celle de la main non-dominante. Pourtant, au cours de notre étude, plusieurs participants ont rapporté l'intérêt de pouvoir déplacer simultanément la lentille et le rayon.

Nous souhaitons explorer plus en profondeur l'apport de la conception bi-manuelle de *RayLens* et savoir si les mouvements du rayon et de la lentille sont réalisés en parallèle ou en séquentiel. Si les mouvements sont parfois réalisés en parallèle, à quelle fréquence et dans quelles situations le sont-ils ?

Améliorations de la lentille de *RayLens*. La lentille implémentée est rectangulaire et est inspirée de la forme et de la taille d'une tablette physique. D'autres formes et tailles de lentille sont possibles. Une lentille 2D mais circulaire rappelant la métaphore de la loupe aurait-elle un impact sur l'utilisabilité de la technique ?

D'autre part, si la lentille doit être déplacée très loin de l'utilisatrice (pour atteindre des objets très distants), la lentille sera petite et il sera donc plus difficile de sélectionner un objet dessus. Une solution serait l'utilisation d'une lentille de taille adaptative : la taille de la lentille varierait en fonction de la distance entre la lentille et l'utilisatrice (plus la lentille est loin, plus elle est grande).

Enfin, l'implémentation actuelle de *RayLens* ne permet pas de tourner la lentille. Ainsi, si l'utilisatrice la déplace loin vers la droite, l'écran de la lentille sera peu visible et il sera difficile voire impossible de pointer dessus. Une amélioration importante est donc la possibilité de faire des rotations de cette lentille mais cela complexifie beaucoup la tâche avec des degrés de liberté supplémentaires à gérer. L'objectif est alors étudier le contrôle de ces rotations tout en préservant la simplicité de la technique. Dans la version actuelle de *RayLens*, la position de la lentille dépend du mouvement du *clicker* tenu par l'utilisatrice. Une solution immédiate serait donc de faire correspondre l'orientation de la lentille à celle du *clicker* (Figure 12.2). Les techniques considérant la métaphore de l'œil dans la main [154, 171] implémentent souvent cette correspondance entre contrôleur et lentille.

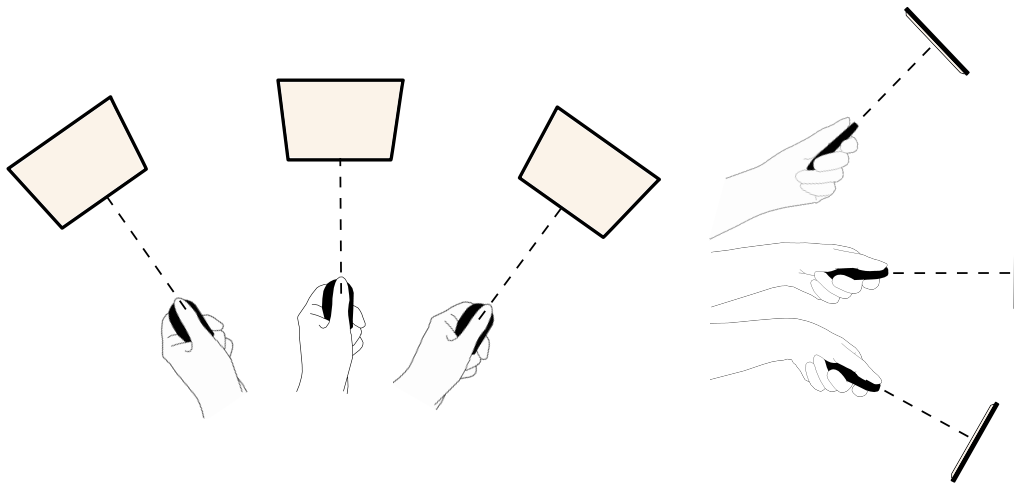


FIGURE 12.2 – Rotations de la lentille.

12.1.2 Interaction hybride combinant interaction tactile et interaction dans l'air (technique *Mixed*)

Réaliser l'installation artistique exploitant les trois techniques : DT, Balloon, Mixed. Un travail à venir est la réalisation de l'installation artistique et de l'étude présentées au Chapitre III Section 10. Comme l'exposition qui avait été planifiée pour expérimenter *RayLens* en contexte grand public (Section 12.1.1), cette installation artistique permettrait de tester les trois techniques auprès d'un public varié et de confronter les résultats obtenus lors des deux expérimentations contrôlées en laboratoire. De plus, cette installation permet d'explorer le choix de la technique d'interaction par les utilisateurs selon la précision requise pour la tâche. Les propriétés respectives des trois techniques sont-elles comprises par les utilisateurs ? Sont-elles les raisons principales de leurs choix ?

Étude avec *Mixed* et *Balloon* en bi-manuel. Pour les deux techniques *Mixed* et *Ballon* qui impliquent une décomposition de la tâche, une perspective de recherche est de considérer l'interaction bi-manuelle. La main dominante réaliserait la tâche 2D sur le plan de la table, l'autre main gèrerait la hauteur de la sélection. Une étude approfondie nous permettrait de savoir si les participants parallélisent les mouvements des mains avec ces techniques. S'il y a parallélisation, est-elle plus simple lorsque les deux mains interagissent dans le même espace comme avec *Balloon* ?

12.2 Perspectives à long terme

12.2.1 Unification des travaux : techniques *RayLens* et *Mixed*

Dans nos travaux, nous avons considéré séparément la sélection d'objets distants (Chapitre II) et la sélection d'objets à portée de main de l'utilisatrice (Chapitre III). Nous avons alors proposé et étudié la technique *RayLens* uniquement dans le contexte de la sélection d'objets distants. De même, la technique *Mixed* n'a été étudiée que pour la sélection d'objets proches. Avec l'objectif d'unifier nos travaux, ces deux techniques pourraient être utilisées à la fois pour la sélection d'objets proches et distants :

- Extension de *RayLens* pour l'interaction avec des objets à portée de main : Si l'utilisatrice interagit avec des objets situés à proximité, la lentille reste à portée de main. Il est tout à fait possible de garder la sélection par raycasting même si la lentille est proche [127]. Dans ce cas, la technique *RayLens* ne change pas : le rayon prolonge le bâton tenu par l'utilisatrice quelle que soit la distance de l'objet.

Cependant, lorsque la lentille est à portée de main, la sélection sur la lentille pourrait aussi être réalisée avec le doigt par toucher direct dans l'air. Dans ce cas, lorsque la lentille s'éloignerait (par exemple à plus de 5 cm du bout du doigt), un rayon prolongeant le doigt pourrait apparaître, et la sélection se ferait alors par raycasting.

- Extension de *Mixed* pour l'interaction avec des objets hors de portée. La première étape (tâche 2D sur la table) pourrait être effectuée par des interactions tactiles indirectes sur la table pour contrôler un curseur sur le plan (interactions indirectes semblables à celles réalisées sur un pavé tactile). La deuxième étape, l'ajustement de la hauteur de la sélection par interaction dans l'air, reste inchangée. Notons qu'avec cette implémentation, *Mixed* ne peut sélectionner que des objets *au-dessus de la table*.

Il s'agirait ensuite de comparer ces deux nouvelles techniques.

12.2.2 Combinaison de la RA sur table et des micro-gestes

Nous avons considéré dans le Chapitre III uniquement l'espace hybride permettant la combinaison de l'interaction tactile sur la table et de l'interaction dans l'air. D'autres espaces hybrides peuvent être explorés. Une possibilité est de combiner l'in-

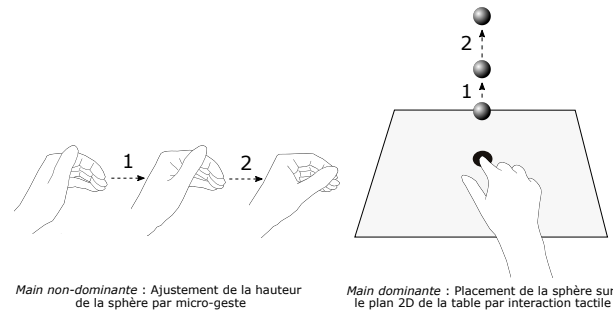


FIGURE 12.3 – Espace hybride combinant l’espace 2D de la table et l’interaction par micro-gestes : exemple de décomposition d’une tâche de positionnement ou de sélection 3D.

teraction tactile et les micro-gestes comme avec la technique *gunslinger* [93]. Ainsi, au lieu d’ajuster la hauteur de la sélection par l’interaction dans l’air comme *Mixed* (interaction qui est peu précise), l’utilisatrice pourrait faire un geste de glissement (*swipe*) du pouce sur l’index. Cette interaction hybride est illustrée à la Figure 12.3.

12.2.3 Pointage : vers de nouvelles taxonomies

Que ce soit pour le pointage en 2D ou en 3D, il existe de nombreuses taxonomies bien établies [5, 10, 29, 30, 126]. Au cours de nos travaux présentés au Chapitre II, nous nous sommes intéressés aux lentilles et avons observé qu’elles étaient principalement utilisées pour des tâches d’exploration et rarement pour des tâches de pointage. Pourtant, les résultats obtenus pour la technique *RayLens* montrent que l’utilisation de ce mécanisme comme une aide au pointage peut être performante.

Il nous semble donc important de réconcilier le mécanisme de lentille et les techniques de facilitation de pointage en retravaillant les taxonomies existantes en 2D et en 3D pour y inclure les techniques utilisant des lentilles.

Nous avons vu au cours de nos travaux que l’espace de conception en RA sur table est vaste. Les facilitations de pointage utilisables pour ces systèmes sont nombreuses :

- Des techniques basées sur des curseurs étendus (par exemple *RayCursor* [15]),
- Des techniques avec mécanismes de désambiguïsation, héritées souvent de la RV (par exemple *SQUAD* [84]),
- Des techniques considérant des lentilles (par exemple *RayLens*, *Slicing volume* [111]),
- Des techniques basées sur la décomposition de la tâche 3D (par exemple *Balloon* [19]),
- etc.

Ces mécanismes de facilitation de pointage sont ensuite à instancier en considérant les modalités d'interaction et donc les espaces d'interaction en entrée (espace d'interaction 2D – table, 3D – au-dessus de la table, ou autre comme avec les micro-gestes) ainsi que la multiplicité des espaces ou non lors de l'interaction à une ou deux mains.

Les techniques de sélection 3D en RA sur table ont besoin d'être répertoriées, classées, et comparées. Une telle taxonomie permettra à la fois de guider la conception de techniques de RA sur table et permettra également l'identification de nouvelles techniques d'interaction au travers de son exploration.

Bibliographie

- [1] J. Accot and S. Zhai. Performance evaluation of input devices in trajectory-based tasks : An application of the steering law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 466–472, New York, NY, USA, 1999. ACM.
- [2] B. Agarwal and W. Stuerzlinger. Widgetlens : A system for adaptive content magnification of widgets. In *Proceedings of the 27th International BCS Human Computer Interaction Conference*, BCS-HCI '13, Swindon, GBR, 2013.
- [3] L. Angelini, E. Mugellini, O. Abou Khaled, and N. Couture. Internet of tangible things (iott) : Challenges and opportunities for tangible interaction with iot. *Informatics*, 5(1), 2018.
- [4] F. Argelaguet and C. Andujar. Efficient 3d pointing selection in cluttered virtual environments. *IEEE Computer Graphics and Applications*, 29(6) :34–43, 2009.
- [5] F. Argelaguet and C. Andujar. A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3) :121–136, 2013.
- [6] B. Bach, R. Sicat, J. Beyer, M. Cordeil, and H. Pfister. The hologram in my hand : How effective is interactive exploration of 3d visualizations in immersive tangible augmented reality? *IEEE Transactions on Visualization & Computer Graphics*, pages 457–467, Jan 2018.
- [7] F. Bacim, R. Kopper, and D. A. Bowman. Design and evaluation of 3d selection techniques based on progressive refinement. *Int. J. Hum.-Comput. Stud.*, 71(7–8) :785–802, July 2013.
- [8] F. Bacim de Araujo e Silva. *Increasing Selection Accuracy and Speed through Progressive Refinement*. PhD thesis, Virginia Tech, 2015.
- [9] H. Bai, G. A. Lee, M. Ramakrishnan, and M. Billinghurst. 3d gesture interaction for handheld augmented reality. In *SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications*, SA '14, New York, NY, USA, 2014. Association for Computing Machinery.
- [10] R. Balakrishnan. "beating" fitts' law : Virtual enhancements for pointing facilitation. *Int. J. Hum.-Comput. Stud.*, 61(6) :857–874, Dec. 2004.

- [11] R. Balakrishnan, T. Baudel, G. Kurtenbach, and G. Fitzmaurice. The rockin'mouse : Integral 3d manipulation on a plane. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, page 311–318, New York, NY, USA, 1997. Association for Computing Machinery.
- [12] R. Balakrishnan and G. Kurtenbach. Exploring bimanual camera control and object manipulation in 3d graphics interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, page 56–62, New York, NY, USA, 1999. Association for Computing Machinery.
- [13] M. Baloup, V. Oudjail, T. Pietrzak, and G. Casiez. Pointing techniques for distant targets in virtual reality. In *Proceedings of the 30th Conference on l'Interaction Homme-Machine*, IHM '18, page 100–107, New York, NY, USA, 2018. Association for Computing Machinery.
- [14] M. Baloup, T. Pietrzak, and G. Casiez. Amélioration du Raycasting par utilisation de la sélection par proximité et du filtrage. *Journal d'Interaction Personne-Système*, 8(1) :61 – 83, Dec. 2019. Special issue : the best of IHM'2018.
- [15] M. Baloup, T. Pietrzak, and G. Casiez. Raycursor : A 3d pointing facilitation technique based on raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.
- [16] R. Bane and T. Hollerer. Interactive tools for virtual x-ray vision in mobile augmented reality. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, page 231–239, USA, 2004. IEEE Computer Society.
- [17] A. Banerjee, J. Burstyn, A. Girouard, and R. Vertegaal. Pointable : An in-air pointing technique to manipulate out-of-reach targets on tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, page 11–20, New York, NY, USA, 2011. Association for Computing Machinery.
- [18] M. Bellgardt, S. Pick, D. Zielasko, B. Weyers, T. Vierjahn, and T. Kuhlen. Utilizing immersive virtual reality in everyday work. 03 2017.
- [19] H. Benko and S. Feiner. Balloon selection : A multi-finger technique for accurate low-fatigue 3d selection. *3D User Interfaces*, 0, 01 2007.
- [20] H. Benko, E. Ishak, and S. Feiner. Cross-dimensional gestural interaction techniques for hybrid immersive environments. pages 209 – 216, 04 2005.

- [21] J. Besharat, A. Komninos, G. Papadimitriou, E. Lagiou, and J. Garofalakis. Augmented paper maps : Design of poi markers and effects on group navigation. *Journal of Ambient Intelligence and Smart Environments*, 8(5) :515–530, 2016.
- [22] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses : The see-through interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, page 73–80, New York, NY, USA, 1993. Association for Computing Machinery.
- [23] R. Blanch and M. Ortega. Benchmarking pointing techniques with distractors : Adding a density factor to fitts' pointing paradigm. In *Proceedings of CHI '11*, page 1629–1638, 2011.
- [24] J. Bobrich. An immersive environment based on paper maps. In *Proc. ICC*, 2003.
- [25] J. Bobrich and S. Otto. Augmented maps. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4) :502–505, 2002.
- [26] S. Boring, M. Jurmu, and A. Butz. Scroll, tilt or move it : Using mobile phones to continuously control pointers on large public displays. volume 411, pages 161–168, 01 2009.
- [27] C. W. Borst, J. Tiesel, and C. M. Best. Real-time rendering method and performance evaluation of composable 3d lenses for interactive vr. *IEEE Transactions on Visualization and Computer Graphics*, 16(3) :394–410, 2010.
- [28] J. Bowers, L. Bannon, M. Fraser, J. Hindmarsh, S. Benford, C. Heath, G. Taxén, and L. Ciolfi. *From the Disappearing Computer to Living Exhibitions : Shaping Interactivity in Museum Settings*, page 30–49. Springer-Verlag, Berlin, Heidelberg, 2007.
- [29] D. A. Bowman and L. F. Hodges. Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, 10(1) :37–53, 1999.
- [30] D. A. Bowman, D. B. Johnson, and L. F. Hodges. Testbed evaluation of virtual environment interaction techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '99, page 26–33, New York, NY, USA, 1999. Association for Computing Machinery.
- [31] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces : Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., USA, 2004.

- [32] L. D. Brown and H. Hua. Magic lenses for augmented virtual environments. *IEEE Computer Graphics and Applications*, 26(4) :64–73, 2006.
- [33] L. D. Brown, H. Hua, and C. Gao. A widget framework for augmented interaction in scape. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*, UIST '03, page 1–10, New York, NY, USA, 2003. Association for Computing Machinery.
- [34] G. Bruder, F. Steinicke, and W. Stuerzlinger. Touching the void revisited : Analyses of touch behavior on and above tabletop surfaces. volume 8117, pages 278–296, 09 2013.
- [35] G. Bruder, F. Steinicke, and W. Sturzlinger. To touch or not to touch ? : Comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop surfaces. In *Proceedings of the 1st Symposium on Spatial User Interaction*, SUI '13, pages 9–16, New York, NY, USA, 2013. ACM.
- [36] V. Buchmann, S. Violich, M. Billinghamurst, and A. Cockburn. Fingartips : Gesture based direct manipulation in augmented reality. In *Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, GRAPHITE '04, page 212–221, New York, NY, USA, 2004. Association for Computing Machinery.
- [37] W. Büschel, J. Chen, R. Dachzelt, S. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, and W. Stuerzlinger. Interaction for immersive analytics. In *Immersive Analytics*, pages 95–138. Springer, 2018.
- [38] W. Büschel, P. Reipschläger, R. Langner, and R. Dachzelt. Investigating the use of spatial interaction for 3d data visualization on mobile devices. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, page 62–71, New York, NY, USA, 2017. Association for Computing Machinery.
- [39] J. Cashion, C. Wingrave, and J. J. LaViola Jr. Dense and dynamic 3d selection for game-based virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 18(4) :634–642, 2012.
- [40] G. Casiez, N. Roussel, and D. Vogel. 1 € filter : A simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 2527–2530, New York, NY, USA, 2012. Association for Computing Machinery.
- [41] O. Chapuis and P. Dragicevic. Effects of motor scale, visual scale, and quantization on small target acquisition difficulty. *ACM Trans. Comput.-Hum. Interact.*, 18(3), Aug. 2011.

- [42] X. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. E. Hudson. Air+touch : Interweaving touch & in-air gestures. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, page 519–525, New York, NY, USA, 2014. Association for Computing Machinery.
- [43] Z. Chen, Y. Su, Y. Wang, Q. Wang, H. Qu, and Y. Wu. Marvist : Authoring glyph-based visualization in mobile augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 26(8) :2645–2658, 2020.
- [44] W. H. Chun and T. Höllerer. Real-time hand interaction for augmented reality on mobile phones. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, IUI '13, page 307–314, New York, NY, USA, 2013. Association for Computing Machinery.
- [45] P. Cignoni, C. Montani, and R. Scopigno. Magicsphere : an insight tool for 3d data visualization. *Computer Graphics Forum*, 13 :317 – 328, 02 2003.
- [46] A. Cockburn, A. Karlson, and B. B. Bederson. A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, 41(1), Jan. 2009.
- [47] A. Colley, J. Hakkikla, J. Schöning, F. Daiber, F. Steinicke, and A. Kruger. Touch the 3rd dimension ! understanding stereoscopic 3d touchscreen interaction. 05 2015.
- [48] B. R. De Araùjo, G. Casiez, and J. A. Jorge. Mockup builder : Direct 3d modeling on and above the surface in a continuous interaction space. In *Proceedings of Graphics Interface 2012*, GI '12, page 173–180, CAN, 2012. Canadian Information Processing Society.
- [49] N. J. Dedual, O. Oda, and S. K. Feiner. Creating hybrid user interfaces with a 2d multi-touch tabletop and a 3d see-through head-worn display. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pages 231–232, 2011.
- [50] P. Dias, L. Afonso, S. Eliseu, and B. S. Santos. Mobile devices for interaction in immersive virtual environments. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [51] P. Dragicevic. *Fair Statistical Communication in HCI*, pages 291–330. Springer International Publishing, Cham, 2016.
- [52] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Henry Riche, F. Schreiber, W. Stuerzlinger, and B. Thomas. *Immersive Analytics : An Introduction*. 01 2018.

- [53] A. O. S. Feiner. The flexible pointer : An interaction technique for selection in augmented and virtual reality. In *Proc. UIST*, volume 3, pages 81–82, 2003.
- [54] J. A. W. Filho, W. Stuerzlinger, and L. Nedel. Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1) :514–524, 2020.
- [55] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6) :381, 1954.
- [56] A. Forsberg, K. Herndon, and R. Zeleznik. Aperture based selection for immersive virtual environments. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, UIST '96, page 95–96, New York, NY, USA, 1996. Association for Computing Machinery.
- [57] A. Fuhrmann and E. Gröller. Real-time techniques for 3d flow visualization. In *Proceedings of the Conference on Visualization '98*, VIS '98, page 305–312, Washington, DC, USA, 1998. IEEE Computer Society Press.
- [58] R. Gasteiger, M. Neugebauer, O. Beuing, and B. Preim. The flowlens : A focus-and-context visualization approach for exploration of blood flow in cerebral aneurysms. *IEEE Transactions on Visualization and Computer Graphics*, 17(12) :2183–2192, Dec. 2011.
- [59] P. Gatalsky, N. Andrienko, and G. Andrienko. Interactive analysis of event data using space-time cube. In *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, pages 145–152, 2004.
- [60] R. Grasset and J.-D. Gascuel. Mare : Multiuser augmented reality environment on table setup. In *ACM SIGGRAPH 2002 Conference Abstracts and Applications*, SIGGRAPH '02, pages 213–213, New York, NY, USA, 2002. ACM.
- [61] T. Grossman and R. Balakrishnan. Pointing at trivariate targets in 3d environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 447–454, New York, NY, USA, 2004. ACM.
- [62] T. Grossman and R. Balakrishnan. The bubble cursor : Enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, page 281–290, New York, NY, USA, 2005. Association for Computing Machinery.
- [63] T. Grossman and R. Balakrishnan. 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, page 3–12, New York, NY, USA, 2006. Association for Computing Machinery.

- [64] Y. Guiard. Asymmetric division of labor in human skilled bimanual action. *Journal of Motor Behavior*, 19(4) :486–517, 1987. PMID : 15136274.
- [65] M. Guillon, F. Leitner, and L. Nigay. Investigating visual feedforward for target expansion techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 2777–2786, New York, NY, USA, 2015. Association for Computing Machinery.
- [66] C. Gutwin. Improving focus targeting in interactive fisheye views. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, page 267–274, New York, NY, USA, 2002. Association for Computing Machinery.
- [67] M. Hachet, B. Bossavit, A. Cohé, and J.-B. de la Rivière. Toucheo : Multitouch and stereo combined in a seamless workspace. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 587–592, New York, NY, USA, 2011. ACM.
- [68] C. Harrison, S. Ramamurthy, and S. Hudson. On-body interaction : Armed and dangerous. pages 69–76, 02 2012.
- [69] S. G. Hart. Nasa-task load index (nasa-tlx) ; 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA : Los Angeles, CA, 2006.
- [70] N. R. Hedley, M. Billinghamurst, L. Postner, R. May, and H. Kato. Explorations in the use of augmented reality for geographic visualization. *Presence*, 11(2) :119–133, 2002.
- [71] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance : A metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 1063–1072, New York, NY, USA, 2014. Association for Computing Machinery.
- [72] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. A survey of design issues in spatial input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology*, UIST '94, page 213–222, New York, NY, USA, 1994. Association for Computing Machinery.
- [73] W. Hürst and C. Van Wezel. Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications*, 62(1) :233–258, 2013.

- [74] W. Hürst and K. Vriens. Multimodal feedback for finger-based interaction in mobile augmented reality. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ICMI '16, page 302–306, New York, NY, USA, 2016. Association for Computing Machinery.
- [75] S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display : A surface technology with an electronically switchable diffuser. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, page 269–278, New York, NY, USA, 2008. Association for Computing Machinery.
- [76] R. J. K. Jacob and L. E. Sibert. The perceptual structure of multidimensional input device selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, page 211–218, New York, NY, USA, 1992. Association for Computing Machinery.
- [77] R. J. K. Jacob, L. E. Sibert, D. C. McFarlane, and M. P. Mullen, Jr. Integrality and separability of input devices. *ACM Trans. Comput.-Hum. Interact.*, 1(1) :3–26, Mar. 1994.
- [78] P. Kabbash, W. Buxton, and A. Sellen. Two-handed input in a compound task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, page 417–423, New York, NY, USA, 1994. Association for Computing Machinery.
- [79] C. M. Kim and T.-J. Nam. G-raff : An elevating tangible block for spatial tabletop interaction. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, page 171, New York, NY, USA, 2015. Association for Computing Machinery.
- [80] M. Kim and J. Y. Lee. Touch and hand gesture-based interactions for directly manipulating 3d virtual objects in mobile augmented reality. *Multimedia Tools Appl.*, 75(23) :16529–16550, Dec. 2016.
- [81] S. Kim, H. Kim, S. Eom, N. P. Mahalik, and B. Ahn. A reliable new 2-stage distributed interactive tgs system based on gis database and augmented reality. *IEICE - Trans. Inf. Syst.*, E89-D(1) :98–105, Jan. 2006.
- [82] U. Kister, P. Reipschläger, and R. Dachsel. Multilens : Fluent interaction with multi-functional multi-touch lenses for information visualization. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, ISS '16, page 139–148, New York, NY, USA, 2016. Association for Computing Machinery.

- [83] S. Kluge, S. Gladisch, U. Freiherr von Lukas, O. Staadt, and C. Tominski. Virtual lenses as embodied tools for immersive analytics. In B. Weyers, C. Lürig, and D. Zielasko, editors, *GI VR / AR Workshop*. Gesellschaft für Informatik e.V., 2020.
- [84] R. Kopper, F. Bacim, and D. A. Bowman. Rapid and accurate 3d selection by progressive refinement. In *Proceedings of the 2011 IEEE Symposium on 3D User Interfaces*, 3DUI '11, page 67–74, USA, 2011. IEEE Computer Society.
- [85] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2010.
- [86] E. Kruijff, J. E. Swan, and S. K. Feiner. Perceptual issues in augmented reality revisited. *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2010.
- [87] G. Kurtenbach and W. Buxton. The limits of expert performance using hierarchic marking menus. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, page 482–487, New York, NY, USA, 1993. Association for Computing Machinery.
- [88] J. Laukkanen, P. Isokoski, and K.-J. Rähä. The cone and the lazy bubble : Two efficient alternatives between the point cursor and the bubble cursor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 309–312, New York, NY, USA, 2008. Association for Computing Machinery.
- [89] J. J. LaViola Jr., E. Kruijff, D. Bowman, I. P. Poupyrev, and R. P. McMahan. *3D User Interfaces : Theory and Practice (second edition)*. Addison-Wesley, 2017.
- [90] J. J. Lee and J. M. Park. 3d mirrored object selection for occluded objects in virtual environments. *IEEE Access*, 8 :200259–200274, 2020.
- [91] J. Liang and M. Green. Geometric modeling using six degrees of freedom input devices. In *3rd Int'l Conference on CAD and Computer Graphics*, pages 217–222. Citeseer, 1993.
- [92] J. Liang and M. Green. Jdcad : A highly interactive 3d modeling system. *Computers & graphics*, 18(4) :499–506, 1994.
- [93] M. Liu, M. Nancel, and D. Vogel. Gunslinger : Subtle arms-down mid-air interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, page 63–71, New York, NY, USA, 2015. Association for Computing Machinery.

- [94] J. Looser, M. Billinghurst, and A. Cockburn. Through the looking glass : The use of lenses as an interface tool for augmented reality interfaces. In *Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, GRAPHITE '04, page 204–211, New York, NY, USA, 2004. Association for Computing Machinery.
- [95] J. Looser, M. Billinghurst, R. Grasset, and A. Cockburn. An evaluation of virtual lenses for object selection in augmented reality. In *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*, GRAPHITE '07, page 203–210, New York, NY, USA, 2007. Association for Computing Machinery.
- [96] J. Looser, R. Grasset, and M. Billinghurst. A 3d flexible and tangible magic lens in augmented reality. pages 51 – 54, 12 2007.
- [97] Y. Lu, C. Yu, and Y. Shi. Investigating bubble mechanism for ray-casting to improve 3d target acquisition in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 35–43, 2020.
- [98] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Hum.-Comput. Interact.*, 7(1) :91–139, Mar. 1992.
- [99] S. Malacria, E. Lecolinet, and Y. Guiard. Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces : The cyclostar approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, page 2615–2624, New York, NY, USA, 2010. Association for Computing Machinery.
- [100] N. Marquardt, R. Jota, S. Greenberg, and J. A. Jorge. The continuous interaction space : Interaction techniques unifying touch and gesture on and above a digital surface. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part III*, INTERACT'11, page 461–476, Berlin, Heidelberg, 2011. Springer-Verlag.
- [101] A. Martinet, G. Casiez, and L. Grisoni. The design and evaluation of 3d positioning techniques for multi-touch displays. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 115–118, 2010.
- [102] A. Martinet, G. Casiez, and L. Grisoni. The design and evaluation of 3d positioning techniques for multi-touch displays. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 115–118, March 2010.
- [103] A. Martinet, G. Casiez, and L. Grisoni. Integrality and separability of multitouch interaction techniques in 3d manipulation tasks. *IEEE Transactions on Visualization and Computer Graphics*, 18(3) :369–380, 2012.

- [104] S. Mayer, V. Schwind, R. Schweigert, and N. Henze. The effect of offset correction and cursor on mid-air pointing in real and virtual environments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 653 :1–653 :13, New York, NY, USA, 2018. ACM.
- [105] D. Medeiros, L. Teixeira, F. Carvalho, I. Santos, and A. Raposo. 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*, VRCAI '13, page 211–218, New York, NY, USA, 2013. Association for Computing Machinery.
- [106] D. Mendes, F. Fonseca, B. Araujo, A. Ferreira, and J. Jorge. Mid-air interactions above stereoscopic interactive tables. 03 2014.
- [107] D. Mendes, D. Medeiros, M. Sousa, E. Cordeiro, A. Ferreira, and J. A. Jorge. Design and evaluation of a novel out-of-reach selection technique for vr using iterative refinement. *Comput. Graph.*, 67(C) :95–102, Oct. 2017.
- [108] E. Mendez, D. Kalkofen, and D. Schmalstieg. Interactive context-driven visualization tools for augmented reality. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, page 209–218, USA, 2006. IEEE Computer Society.
- [109] T. Menzner, T. Gesslein, A. Otte, and J. Grubert. Above surface interaction for multiscale navigation in mobile virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 372–381. IEEE, 03 2020.
- [110] Microsoft. Microsoft hololens, 2017.
- [111] R. Montano, C. Nguyen, R. Kazi, S. Subramanian, S. DiVerdi, and D. Martinez Plascencia. Slicing-volume : Hybrid 3d/2d multi-target selection technique for dense virtual environments. pages 53–62, 03 2020.
- [112] T. Moscovich, F. Chevalier, N. Henry, E. Pietriga, and J.-D. Fekete. Topology-aware navigation in large networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, page 2319–2328, New York, NY, USA, 2009. Association for Computing Machinery.
- [113] A. Mossel, B. Venditti, and H. Kaufmann. Drillsample : Precise selection in dense handheld augmented reality environments. In *Proceedings of the Virtual Reality International Conference : Laval Virtual*, VRIC '13, New York, NY, USA, 2013. Association for Computing Machinery.

- [114] R. C. R. Mota, A. Rocha, J. D. Silva, U. Alim, and E. Sharlin. 3de interactive lenses for visualization in virtual environments. In *2018 IEEE Scientific Visualization Conference (SciVis)*, pages 21–25, 2018.
- [115] M. E. Mott and J. O. Wobbrock. Beating the bubble : Using kinematic triggering in the bubble lens for acquiring small, dense targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, page 733–742, New York, NY, USA, 2014. Association for Computing Machinery.
- [116] S. Nilsson, B. J. E. Johansson, and A. Jonsson. Cross-organizational collaboration supported by augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 17(10) :1380–1392, 2011.
- [117] A. Olwal, H. Benko, and S. Feiner. Senseshapes : Using statistical geometry for object selection in a multimodal augmented reality system. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR '03*, page 300, USA, 2003. IEEE Computer Society.
- [118] A. Olwal and A. Henrysson. Lumar : A hybrid spatial display system for 2d and 3d handheld augmented reality. In *17th International Conference on Artificial Reality and Telexistence (ICAT 2007)*, pages 63–70, 2007.
- [119] L. Pavanatto, C. North, D. Bowman, R. Stoakley, and C. Badea. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2021.
- [120] A. R. Payne, B. Plimmer, A. McDaid, A. Luxton-Reilly, and T. C. Davies. Expansion cursor : A zoom lens that can be voluntarily activated by the user at every individual click. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction, OzCHI '16*, page 81–90, New York, NY, USA, 2016. Association for Computing Machinery.
- [121] C. H. Peck. Useful parameters for the design of laser pointer interaction techniques. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems, CHI EA '01*, page 461–462, New York, NY, USA, 2001. Association for Computing Machinery.
- [122] E. Peillard, F. Argelaguet, J. Normand, A. Lécuyer, and G. Moreau. Studying exocentric distance perception in optical see-through augmented reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 115–122, 2019.

- [123] E. Pietriga and C. Appert. Sigma lenses : Focus-context transitions combining space, time and translucence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 1343–1352, New York, NY, USA, 2008. Association for Computing Machinery.
- [124] C. Pindat, E. Pietriga, O. Chapuis, and C. Puech. Drilling into complex 3d models with gimlenses. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, VRST '13, page 223–230, New York, NY, USA, 2013. Association for Computing Machinery.
- [125] M. E. Porter and J. E. Heppelmann. Why every organization needs an augmented reality strategy. 2017.
- [126] I. Poupyrev and T. Ichikawa. Manipulating objects in virtual worlds : Categorization and empirical evaluation of interaction techniques. *Journal of Visual Languages & Computing*, 10(1) :19–35, 1999.
- [127] I. Poupyrev, T. Ichikawa, S. Weghorst, and M. Billinghurst. Egocentric object manipulation in virtual environments : empirical evaluation of interaction techniques. In *Computer graphics forum*, volume 17, pages 41–52. Wiley Online Library, 1998.
- [128] G. Ramos, A. Cockburn, R. Balakrishnan, and M. Beaudouin-Lafon. Pointing lenses : Facilitating stylus input through visual-and motor-space magnification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, page 757–766, New York, NY, USA, 2007. Association for Computing Machinery.
- [129] R. C. Ramos Mota, S. Cartwright, E. Sharlin, H. Hamdi, M. Costa Sousa, and Z. Chen. Exploring immersive interfaces for well placement optimization in reservoir models. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, SUI '16, page 121–130, New York, NY, USA, 2016. Association for Computing Machinery.
- [130] P. Reipschläger and R. Dachzelt. Designar : Immersive 3d-modeling combining augmented reality with interactive displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, ISS '19, page 29–41, New York, NY, USA, 2019. Association for Computing Machinery.
- [131] P. Reipschläger, S. Engert, and R. Dachzelt. Augmented displays : Seamlessly extending interactive surfaces with head-mounted augmented reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–4, New York, NY, USA, 2020. Association for Computing Machinery.

- [132] P. Reipschlag, T. Flemisch, and R. Dachsel. Personal augmented reality for information visualization on large interactive displays. *IEEE Transactions on Visualization and Computer Graphics*, page 1–1, 2020.
- [133] G. Reitmayr and D. Schmalstieg. Mobile collaborative augmented reality. pages 114 – 123, 02 2001.
- [134] G. Ren and E. O’Neill. 3d selection with freehand gesture. *Computers & Graphics*, 37(3) :101–120, 2013.
- [135] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments - a review. *ACM Comput. Surv.*, 46(2) :23 :1–23 :40, Dec. 2013.
- [136] H. Ro, S. Chae, I. Kim, J. Byun, Y. Yang, Y. Park, and T. Han. A dynamic depth-variable ray-casting interface for object manipulation in ar environments. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2873–2878, 2017.
- [137] A. Rocha, J. D. Silva, U. R. Alim, S. Carpendale, and M. C. Sousa. Decal-lenses : Interactive lenses on surfaces for multivariate visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(8) :2568–2582, 2019.
- [138] K. A. Satriadi, B. Ens, M. Cordeil, B. Jenny, T. Czauderna, and W. Willett. Augmented reality map navigation with freehand gestures. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 593–603, 2019.
- [139] D. Schmalstieg, L. M. Encarnação, and Z. Szalavári. Using transparent props for interaction with the virtual table. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics*, I3D ’99, page 147–153, New York, NY, USA, 1999. Association for Computing Machinery.
- [140] D. Schmalstieg and T. Höllerer. *Augmented Reality : Principles and Practice*. Addison-Wesley usability and HCI series. Addison-Wesley, 2016.
- [141] G. Schmidt, Y. Baillot, D. G. Brown, E. B. Tomlin, and J. E. Swan. Toward disambiguating multiple selections for frustum-based pointing. In *3D User Interfaces (3DUI’06)*, pages 87–94, 2006.
- [142] J. Seifert, S. Boring, C. Winkler, F. Schaub, F. Schwab, S. Herrdum, F. Maier, D. Mayer, and E. Rukzio. Hover pad : Interacting with autonomous and self-actuated displays in space. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST ’14, page 139–147, New York, NY, USA, 2014. Association for Computing Machinery.

- [143] L. Shao, A. Mahajan, T. Schreck, and D. J. Lehmann. Interactive regression lens for exploring scatter plots. *Comput. Graph. Forum*, 36(3) :157–166, June 2017.
- [144] C. D. Shaw. Pain and fatigue in desktop vr : Initial results. In *Graphics Interface*, volume 98, pages 18–20, 1998.
- [145] M. Sousa, D. Mendes, S. Paulo, N. Matela, J. Jorge, and D. S. o. Lopes. Vrrrroom : Virtual reality for radiologists in the reading room. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 4057–4062, New York, NY, USA, 2017. Association for Computing Machinery.
- [146] M. Spindler and R. Dachsel. Paperlens : Advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, New York, NY, USA, 2009. Association for Computing Machinery.
- [147] M. Spindler and R. Dachsel. Exploring information spaces by using tangible magic lenses in a tabletop environment. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, page 4771–4776, New York, NY, USA, 2010. Association for Computing Machinery.
- [148] M. Spindler, C. Tominski, H. Schumann, and R. Dachsel. Tangible views for information visualization. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, page 157–166, New York, NY, USA, 2010. Association for Computing Machinery.
- [149] S. Y. Ssin, J. A. Walsh, R. T. Smith, A. Cunningham, and B. H. Thomas. Geogate : Correlating geo-temporal datasets using an augmented reality space-time cube and tangible interactions. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 210–219, 2019.
- [150] A. Steed. Towards a general model for selection in virtual environments. In *3D User Interfaces (3DUI'06)*, pages 103–110, 2006.
- [151] F. Steinicke, T. Ropinski, and K. Hinrichs. Object selection in virtual environments using an improved virtual pointer metaphor. In *Computer vision and graphics*, pages 320–326. Springer, 2006.
- [152] S. Stoev, M. Feurer, and M. Ruckaberle. Exploring the past : A toolset for visualization of historical events in virtual environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '01, page 63–70, New York, NY, USA, 2001. Association for Computing Machinery.

- [153] S. Stoev, D. Schmalstieg, and W. Straßer. Two-handed through-the-lens-techniques for navigation in virtual environments. In *Proceedings of the 7th Eurographics Conference on Virtual Environments & 5th Immersive Projection Technology*, EGVE'01, page 51–60, Goslar, DEU, 2001. Eurographics Association.
- [154] S. Stoev, D. Schmalstieg, and W. Straßer. The through-the-lens metaphor : taxonomy and application. In *Proceedings IEEE Virtual Reality 2002*, pages 285–286, 2002.
- [155] S. Strothoff, D. Valkov, and K. Hinrichs. Triangle cursor : Interactions with objects above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pages 111–119, New York, NY, USA, 2011. ACM.
- [156] H. B. Surale, A. Gupta, M. Hancock, and D. Vogel. Tabletinvr : Exploring the design space for using a multi-touch tablet in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.
- [157] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 13(3) :429–442, 2007.
- [158] C. Tominski, S. Gladisch, U. Kister, R. Dachsel, and H. Schumann. A survey on interactive lenses in visualization. In *EuroVis (STARs)*. Citeseer, 2014.
- [159] C. Tominski, S. Gladisch, U. Kister, R. Dachsel, and H. Schumann. Interactive lenses for visualization : An extended survey. In *Computer Graphics Forum*, volume 36, pages 173–200. Wiley Online Library, 2017.
- [160] X. Tong, C. Li, and H. Shen. Glyphlens : View-dependent occlusion management in the interactive glyph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1) :891–900, 2017.
- [161] M. Traoré, C. Hurter, and A. Telea. Interactive obstruction-free lensing for volumetric data visualization. *IEEE Transactions on Visualization and Computer Graphics*, PP :1–1, 09 2018.
- [162] B. Ullmer and H. Ishii. The metadesk : Models and prototypes for tangible user interfaces. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, page 223–232, New York, NY, USA, 1997. Association for Computing Machinery.

- [163] L. Vanacken, T. Grossman, and K. Coninx. Exploring the effects of environment density and target visibility on object selection in 3d virtual environments. In *2007 IEEE Symposium on 3D User Interfaces*, 2007.
- [164] L. Vanacken, T. Grossman, and K. Coninx. Multimodal selection techniques for dense and occluded 3d virtual environments. *International Journal of Human-Computer Studies*, 67(3) :237–255, 2009.
- [165] M. Veit, A. Capobianco, and D. Bechmann. An experimental analysis of the impact of touch screen interaction techniques for 3-d positioning tasks. pages 75 – 82, 04 2011.
- [166] D. L. Vickers. *Sorcerer’s Apprentices : Head-Mounted Display and Wand*. PhD thesis, 1974.
- [167] J. Viega, M. J. Conway, G. Williams, and R. Pausch. 3d magic lenses. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, UIST ’96, page 51–58, New York, NY, USA, 1996. Association for Computing Machinery.
- [168] D. Vogel and P. Baudisch. Shift : A technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’07, page 657–666, New York, NY, USA, 2007. Association for Computing Machinery.
- [169] J. Wang and R. Lindeman. Coordinated 3d interaction in tablet- and hmd-based hybrid virtual environments. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction*, SUI ’14, page 70–79, New York, NY, USA, 2014. Association for Computing Machinery.
- [170] J. Wang and R. Lindeman. Object impersonation : Towards effective interaction in tablet- and hmd-based hybrid virtual environments. pages 111–118, 03 2015.
- [171] C. Ware and S. Osborne. Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the 1990 Symposium on Interactive 3D Graphics*, I3D ’90, page 175–183, New York, NY, USA, 1990. Association for Computing Machinery.
- [172] F. Wiehr, F. Daiber, F. Kosmalla, and A. Krüger. Artopos : Augmented reality terrain map visualization for collaborative route planning. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, UbiComp ’17, page 1047–1050, New York, NY, USA, 2017. Association for Computing Machinery.

- [173] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence : Teleoper. Virtual Environ.*, 17(1) :91–101, Feb. 2008.
- [174] A. Wilson and S. Shafer. Xwand : Ui for intelligent spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, page 545–552, New York, NY, USA, 2003. Association for Computing Machinery.
- [175] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 143–146, New York, NY, USA, 2011. ACM.
- [176] A. Worden, N. Walker, K. Bharat, and S. Hudson. Making computers easier for older adults to use : Area cursors and sticky icons. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, page 266–271, New York, NY, USA, 1997. Association for Computing Machinery.
- [177] R. Xiao, J. Schwarz, N. Throm, A. D. Wilson, and H. Benko. Mrtouch : Adding touch input to head-mounted mixed reality. *IEEE Transactions on Visualization and Computer Graphics*, 24(4) :1653–1660, Apr. 2018.
- [178] J. S. Yi, Y. a. Kang, J. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6) :1224–1231, 2007.
- [179] K. Özacar, J. D. Hincapié-Ramos, K. Takashima, and Y. Kitamura. 3d selection techniques for mobile augmented reality head-mounted displays. *Interacting with Computers*, 29(4) :579–591, July 2017.

Annexes

A. Chapitre I – Étude expérimentale 1 :

Questionnaire intermédiaire

Q1. I found the selection technique easy to understand

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2. I found it easy to select the target

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q3. I feel that I found quickly the target with this technique

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q4. I feel that I selected quickly the target with this technique

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q5. I found using this technique physically demanding

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q6. I found using this technique mentally demanding

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

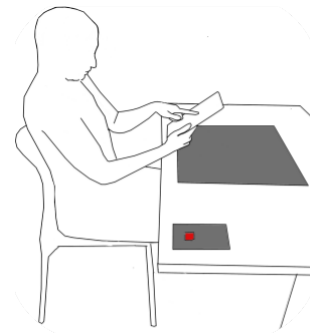
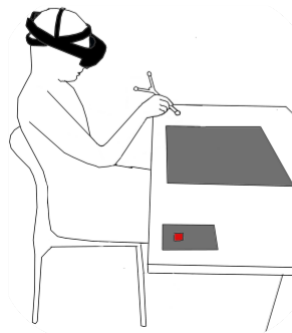
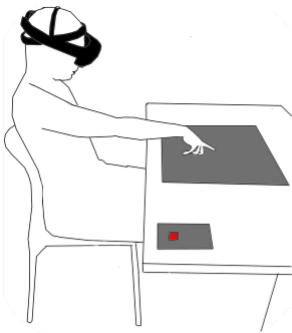
Q7. I found this technique frustrating

strongly disagree	disagree	neutral	agree	strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

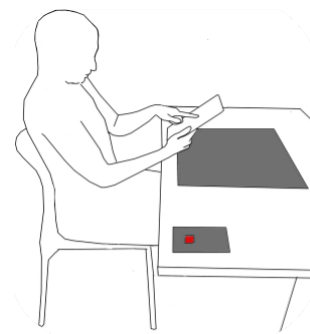
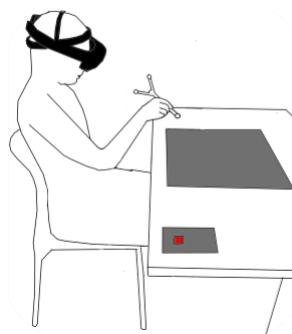
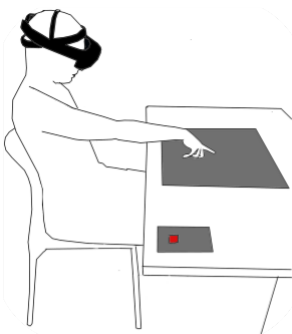
	No pain 1	Slight 2	Moderate 3	Severe 4
Eye	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Face	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Neck/ Shoulder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Arms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fingers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Back	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B. Chapitre I – Étude expérimentale 1 : Questionnaire final

1. Classification of the 3 techniques from the less tiring to the most tiring :
Please write 1 (most tiring), 2 or 3 (less tiring) under each technique



2. Classification of the 3 techniques in your preference order :
Please write 1 (less preferred), 2 or 3 (most preferred) under each technique



C. Chapitre I – Étude expérimentale 2 :

Questionnaire intermédiaire

Q1. It is easy to select the target with this technique:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2. I select the target quickly with this technique:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q3. This technique is physically demanding:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q4. This technique is mentally demanding:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q5. This technique is frustrating:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

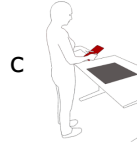
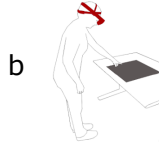
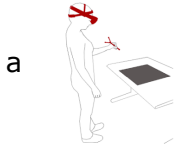
Q6. I felt I was accurate with this technique:

	strongly disagree	disagree	neutral	agree	strongly agree
When the target is visible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When the target is hidden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

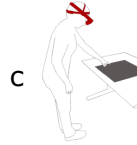
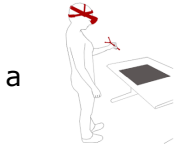
D. Chapitre I – Étude expérimentale 2 :

Questionnaire final

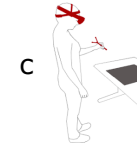
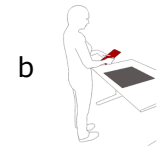
Q1. Choose the easiest technique to select a visible target



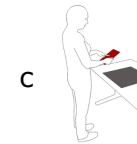
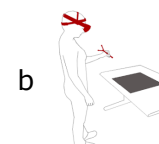
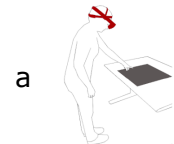
Q2. Choose the easiest technique to select a hidden target



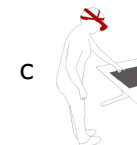
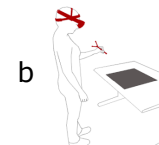
Q3. Choose the quickest technique to select a visible target



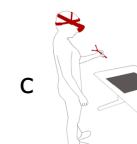
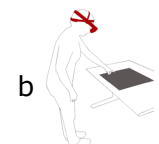
Q4. Choose the quickest technique to select a hidden target



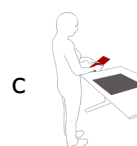
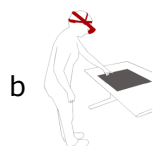
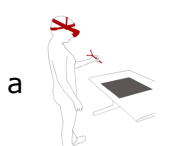
Q5. Choose the least physically demanding technique



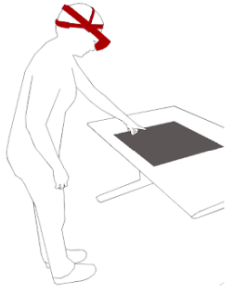
Q6. Choose the least mentally demanding technique



Q7. Choose the least frustrating technique

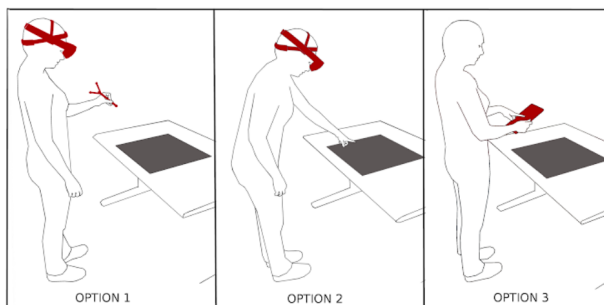


Q8. With this technique, you passed through virtual objects to select a hidden target

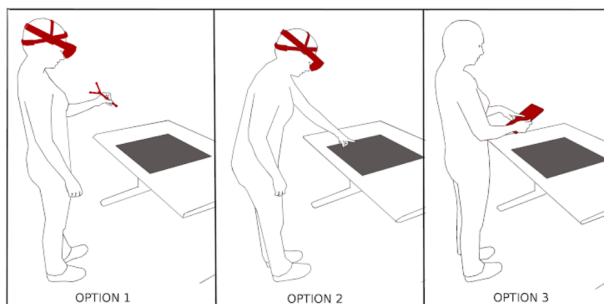


- ☐ Never ☐ Sometimes ☐ Often ☐ Always
☐ I dont remember

Q9. Rank the 3 techniques from the least accurate technique to the most accurate technique



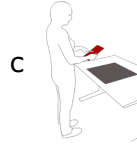
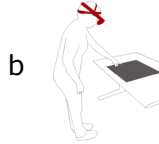
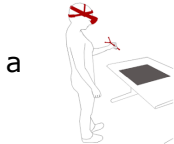
Q10. Rank the 3 techniques in your preference order



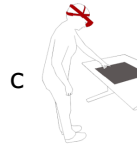
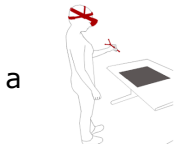
E. Chapitre I – Étude expérimentale 2 :

Questionnaire final

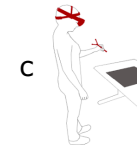
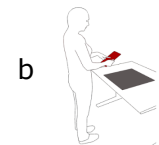
Q1. Choose the easiest technique to select a visible target



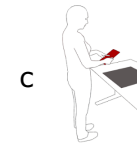
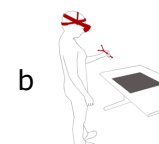
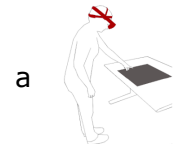
Q2. Choose the easiest technique to select a hidden target



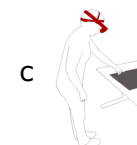
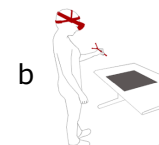
Q3. Choose the quickest technique to select a visible target



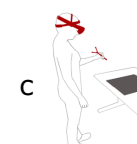
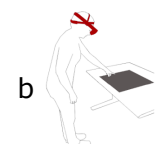
Q4. Choose the quickest technique to select a hidden target



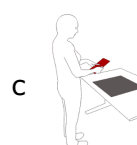
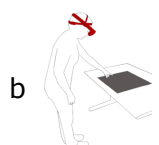
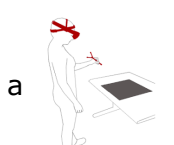
Q5. Choose the least physically demanding technique



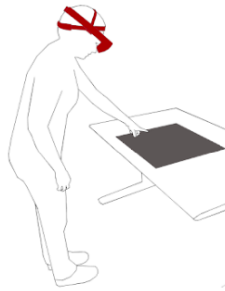
Q6. Choose the least mentally demanding technique



Q7. Choose the least frustrating technique

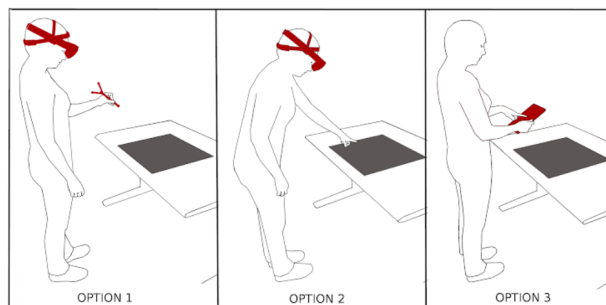


Q8. With this technique, you passed through virtual objects to select a hidden target

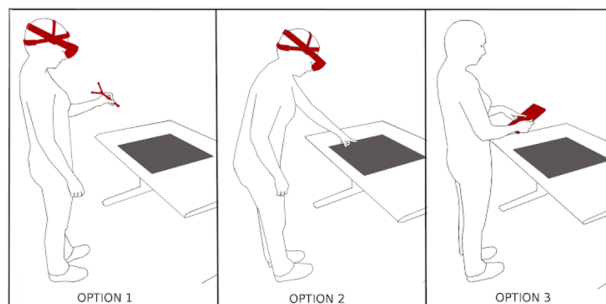


- ☐ Never ☐ Sometimes ☐ Often ☐ Always
☐ I dont remember

Q9. Rank the 3 techniques from the least accurate technique to the most accurate technique



Q10. Rank the 3 techniques in your preference order



F. Chapitre II – Questionnaire Intermédiaire

[Mental Demand] How mentally demanding was the task? // Quel degré d'activité mentale était exigé pour réaliser la tâche ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[Physical Demand] How physically demanding was the task? // Quel degré d'activité physique était exigé pour réaliser la tâche ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[Temporal Demand] How hurried or rushed was the pace of the task? // Quelle pression temporelle avez-vous ressentie pour réaliser la tâche ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[Performance] How successful were you in accomplishing what you were asked to do? // Avec quel succès pensez-vous avoir réalisé la tâche ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[Effort] How hard did you have to work to accomplish your level of performance? // Quel effort deviez-vous fournir pour accomplir la tâche ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[Frustration] How insecure, discouraged, irritated, stressed, and annoyed were you? // Avez-vous ressenti, durant votre tâche, de l'insécurité, du découragement, de l'irritation, du stress ou de l'agacement ?

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[TRANSPARENCY] The transparency filter applied on the objects placed in front of the cursor is useful // Le filtre de transparence appliqué sur les objets placés devant le curseur est utile

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[TRANSPARENCY] The transparency filter applied on the objects placed in front of the lens is useful // Le filtre de transparence appliqué sur les objets placés devant la lentille est utile

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

[FEEDBACK] The visual feedback representing the field of view of the lens is useful // Le feedback représentant le champ de vision de la lentille est utile

1 2 3 4 5 6 7

Very Low ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very High

Question pour
RaySlider

Question pour
RayLens

G. Chapitre II – Questionnaire final

[Mental Demand] The least mentally demanding technique to select a target // Choisissez la technique la moins mentalement demandante pour sélectionner une cible

- ☐ Ray + Lens // Rayon + Lentille
- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray only // Rayon seulement

[Physical Demand] Choose the least physically demanding technique to select a target // Choisissez la technique la moins physiquement demandante pour sélectionner une cible

- ☐ Ray only // Rayon seulement
- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray + Lens // Rayon + Lentille

[Temporal Demand] Choose the least temporally demanding technique to select a target // Choisissez la technique la moins temporellement demandante pour sélectionner une cible

- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray only // Rayon seulement
- ☐ Ray + Lens // Rayon + Lentille

[Performance] Choose the the most successful technique to select a target // Choisissez la technique la plus efficace pour sélectionner une cible

- ☐ Ray + Lens // Rayon + Lentille
- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray only // Rayon seulement

[Effort] Choose the technique demanding the least effort to select a target // Choisissez la technique demandant le moins d'effort pour sélectionner une cible

- ☐ Ray only // Rayon seulement
- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray + Lens // Rayon + Lentille

[Frustration] Choose the least frustrating technique to select a target // Choisissez la technique la moins frustrante pour sélectionner une cible

- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray only // Rayon seulement
- ☐ Ray + Lens // Rayon + Lentille

[Accuracy] Choose the most accurate technique to select a target // Choisissez la technique la plus précise pour sélectionner une cible

- ☐ Ray only // Rayon seulement
- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray + Lens // Rayon + Lentille

[Quick] Choose the quickest technique to select a target // Choisissez la technique la plus rapide pour sélectionner une cible

- ☐ Ray + Cursor // Rayon + Curseur
- ☐ Ray only // Rayon seulement
- ☐ Ray + Lens // Rayon + Lentille

Rank the 3 techniques in your preference order // Classez selon votre ordre de préférence

H. Chapitre III – Questionnaire Inter-médiaire

[Mental Demand] How mentally demanding was the task? // Quel degré d'activité mentale était exigé pour réaliser la tâche ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

[Physical Demand] How physically demanding was the task? // Quel degré d'activité physique était exigé pour réaliser la tâche ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

[Temporal Demand] How hurried or rushed was the pace of the task? // Quelle pression temporelle avez-vous ressentie pour réaliser la tâche ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

[Performance] How successful were you in accomplishing what you were asked to do? // Avec quel succès pensez-vous avoir réalisé la tâche ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

[Effort] How hard did you have to work to accomplish your level of performance? // Quel effort deviez-vous fournir pour accomplir la tâche ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

[Frustration] How insecure, discouraged, irritated, stressed, and annoyed were you? // Avez-vous ressenti, durant votre tâche, de l'insécurité, du découragement, de l'irritation, du stress ou de l'agacement ?

	1	2	3	4	5	6	7	
Very Low	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very High

I. Chapitre III – Questionnaire final

[Mental Demand] The least mentally demanding technique to select a target // Choisissez la technique la moins mentalement demandante pour sélectionner une cible

- ☐ Balloon
- ☐ Mixed
- ☐ DT

[Physical Demand] Choose the least physically demanding technique to select a target // Choisissez la technique la moins physiquement demandante pour sélectionner une cible

- ☐ DT
- ☐ Mixed
- ☐ Balloon

[Temporal Demand] Choose the least temporally demanding technique to select a target // Choisissez la technique la moins temporellement demandante pour sélectionner une cible

- ☐ Mixed
- ☐ DT
- ☐ Balloon

[Performance] Choose the the most successful technique to select a target // Choisissez la technique la plus efficace pour sélectionner une cible

- ☐ Balloon
- ☐ Mixed
- ☐ DT

[Effort] Choose the technique demanding the least effort to select a target // Choisissez la technique demandant le moins d'effort pour sélectionner une cible

- ☐ DT
- ☐ Mixed
- ☐ Balloon

[Frustration] Choose the least frustrating technique to select a target // Choisissez la technique la moins frustrante pour sélectionner une cible

- ☐ Mixed
- ☐ DT
- ☐ Balloon

[Accuracy] Choose the most accurate technique to select a target // Choisissez la technique la plus précise pour sélectionner une cible

- ☐ DT
- ☐ Mixed
- ☐ Balloon

[Quick] Choose the quickest technique to select a target // Choisissez la technique la plus rapide pour sélectionner une cible

- ☐ Mixed
- ☐ DT
- ☐ Balloon

Rank the 3 techniques in your preference order // Classez selon votre ordre de préférence