



Université Larbi Ben M'hidi -Oum El Bouaghi
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département de Mathématiques et Informatique

Thèse de Doctorat

Pour l'obtention du diplôme de Docteur 3ème cycle LMD en Informatique

Option : Ingénierie des Systèmes Distribués

Par :

MESSACI Assia

Thème :

Contribution à l'interaction 3D collaborative en réalité virtuelle

Soutenue le 06 juin 2023

Devant le jury composé de :

M. DERDOUR Makhoulf	Professeur	UOEB	Président
M. TAOUCHE Chérif	Professeur	UOEB	Examineur
M. BENBELKACEM Samir	MRA	CDTA	Examineur
M. BELHOCINE Mahmoud	DR	CDTA	Directeur de thèse
M ^{me} ZENATI-HENDA Nadia	DR	CDTA	Co-directrice de thèse
M. ZAIDI Sofiane	MCB	UOEB	Invité

Remerciements

En tout premier lieu, je remercie Dieu, tout puissant et miséricordieux de m'avoir donné la santé, la force, la patience et la volonté de mener à bien ce modeste travail.

Je remercie mon directeur de thèse, monsieur Mahmoud BELHOCINE directeur de recherche au Centre de Développement des Technologies Avancées (CDTA) d'avoir accepté de prendre la direction de cette thèse. Je lui suis également reconnaissante de m'avoir accordé généreusement le temps nécessaire pour partager avec moi sa grande expérience et de m'avoir assuré un encadrement rigoureux, tout en me donnant toute fois la possibilité de trouver par moi-même mon cheminement personnel.

Mes sincères remerciements s'adressent également à madame Nadia ZENATI directrice de recherche au CDTA, qui m'a toujours soutenue et encouragée. Son écoute, ses critiques, sa disponibilité et ses conseils constructifs m'ont guidée tout au long de cette thèse. Je tiens à saluer son dynamisme, son énergie et sa générosité qui ont déteint sur moi durant ces quelques années de travail ensemble.

Je tiens à remercier et à exprimer mon profond respect à monsieur Makhouf DERDOUR professeur à l'université Larbi Ben M'hidi pour avoir accepté de présider mon jury de thèse.

Mes plus vifs remerciements vont également à monsieur Chérif TAOUCHE professeur à l'université Larbi Ben M'hidi et à monsieur Samir BENBELKACEM maître de recherche A au CDTA pour avoir accepté d'examiner ce travail. Je remercie monsieur Sofiane ZAIDI maître de conférence B à l'université Larbi Ben M'hidi d'avoir accepté d'être dans le jury.

J'adresse ma gratitude à tous les membres de l'équipe IRVA (Interaction en Réalité Virtuelle et Augmentée) du CDTA, merci d'être bienveillants et de m'avoir si bien intégré au sein de l'équipe. J'ai une pensée particulière pour Hayet et Abdelkader qui m'ont encouragée et soutenue durant mes années de thèse.

Un grand merci à tous mes amis et collègues du CDTA.

Je tenais à remercier monsieur Samir OTMANE professeur à l'université d'Évry-Val-d'Essonne, Université Paris Saclay de l'accueil qu'il m'a réservé et du temps qu'il m'a consacré lors de mes stages au sein du laboratoire IBISC. Je remercie également tous les membres de l'équipe IRA2 (Interaction, Réalité Augmentée et Robotique Ambiante) pour leur aide et conseils.

À tous ceux et celles qui ont participé de près ou de loin à l'élaboration du présent travail.

À mes très chers parents.

À ma sœur.

À mon frère et sa petite famille.

À Rassim

À Esra et Anfel

À toute ma famille.



Table des matières

Liste des figures	ix
Liste des tables	ix
Liste des abréviations	x
Résumé	xi
Introduction générale	1
1 Interaction 3D en Réalité Virtuelle	5
1.1 Introduction	5
1.2 Réalité virtuelle : historique et définitions	7
1.3 Applications en réalité virtuelle	9
1.4 Définitions	10
1.4.1 Réalité augmentée	10
1.4.2 Environnement virtuel	13
1.4.3 Interaction	13
1.4.4 Paradigme et métaphore d'interaction	14
1.4.5 Environnement virtuel collaboratif (EVC)	14
1.4.6 Interaction 3D collaborative	15
1.5 Interaction 3D en réalité virtuelle	16
1.5.1 Interaction 3D mono-utilisateur	17
1.5.1.1 Les tâches d'interaction 3D	17
1.5.1.2 Classification des techniques d'interaction 3D	20
1.5.2 Interaction 3D collaborative	38

1.5.2.1	Classification des techniques d'interaction 3D Collaborative	39
1.6	Bilan et discussion	46
1.7	Conclusion	49
2	Reconnaissance de gestes : technologies et techniques	50
2.1	Introduction	50
2.2	Les interfaces utilisateur naturelles	51
2.3	Définition et représentation du geste	52
2.4	Techniques d'acquisition et de reconnaissance de gestes	53
2.4.1	Approches basées capteurs	54
2.4.1.1	Capteurs mécaniques	54
2.4.1.2	Capteurs magnétiques	56
2.4.1.3	Capteurs acoustiques	56
2.4.1.4	Capteurs Inertiels	56
2.4.1.5	Dispositifs tactiles	57
2.4.1.6	Autres types de capteurs	58
2.4.2	Approches basées vision	60
2.4.2.1	Techniques basées apparence	61
2.4.2.2	Techniques basées modèle 3D	69
2.5	Bilan et discussion	80
2.6	Conclusion	83
3	Contribution à l'interaction 3D en RV : "Zoom-fwd"	85
3.1	Introduction	85
3.2	Développement d'une technique d'interaction 3D en utilisant la reconnaissance de gestes dans un environnement virtuel	86
3.2.1	Conception du setup	86
3.2.2	Conception du système	87
3.2.3	Les gestes utilisés	88
3.2.4	Résultats obtenus	89
3.3	"Zoom-fwd" technique d'interaction 3D efficace pour la sélection des objets distants et occultés	90

3.3.1	Conception du Setup	90
3.3.2	Conception du système	91
3.3.2.1	Le module acquisition des données	92
3.3.2.2	Le module reconnaissance de gestes	92
3.3.2.3	Le module interaction 3D	94
3.3.3	Approche proposée	94
3.3.4	Évaluation de la technique " <i>Zoom-fwd</i> "	104
3.3.4.1	Hypothèse de recherche	105
3.3.4.2	Le protocole expérimental	106
3.3.4.3	Mesures et analyse statistique des données	110
3.4	Discussion	119
3.5	Conclusion	121
4	Système 3D collaboratif	123
4.1	Introduction	123
4.2	Système 3D collaboratif	124
4.2.1	Modélisation de la collaboration	124
4.2.1.1	Modèle des "3C"	124
4.2.1.2	Modèle 3D collaboratif	126
4.3	Exemple informatique du système collaboratif 3D	130
4.3.1	Système d'apprentissage collaboratif pour la maintenance industrielle	130
4.3.1.1	Conception du Setup	130
4.3.2	Conception du système global	132
4.3.2.1	Interaction 3D	132
4.3.2.2	Interaction 3D Collaborative	133
4.3.2.3	Visualisation	137
4.4	Évaluation du système 3D collaboratif	137
4.4.1	Hypothèse de recherche	138
4.4.1.1	Questions de recherche	138
4.4.1.2	Hypothèses de recherche	138
4.4.2	Le protocole expérimental	139
4.4.2.1	Les participants	139

4.4.2.2	Conception du protocole expérimental	139
4.4.2.3	Tâches expérimentales	140
4.4.3	Procédure expérimentale	140
4.4.4	Mesures et analyses statistiques des données	141
4.5	Conclusion et perspectives	141
Conclusion générale		143
BIBLIOGRAPHIE		145
Annexe		166
4.6	Questions démographique	166
4.7	Questionnaires d'évaluation	167
4.7.1	Questionnaire d'évaluation USE	167
4.7.2	Questionnaire d'évaluation sur l'Utilisabilité du Système SUS	168

Liste des figures

1.1	Les trois composantes de RV [7]	8
1.2	Exemples d'applications en réalité augmentée : a) Superposition des vaisseaux sanguins sur la vidéo laparoscopique [19], b) Application MARTA (Mobile Augmented Reality Technical Assistance) [20]	12
1.3	Le continuum Réel-Virtuel [23]	12
1.4	Techniques de navigation : a) direction du regard, b) déplacement en utilisant une carte, c) déplacement en utilisant une accélération constante [40]	21
1.5	Technique de la main virtuelle simple [42]	24
1.6	La techniques Go-Go[45]	25
1.7	Technique PRISM [48]	26
1.8	Technique 3D bubble cursor [49]	27
1.9	Technique Ray-casting [50]	28
1.10	Technique Image plane [52]	29
1.11	Techniques Ray-flexible [54]	29
1.12	a) Technique Flashlight [55] b) Technique Aperture [56]	30
1.13	Technique Eye-casting [57]	31
1.14	Technique Depth-Ray [49]	31
1.15	Technique monde en miniature [59]	32
1.16	Technique Voodoo Dolls [61]	33
1.17	Technique Indirect-touch [63]	33
1.18	Technique Scaled-world grab [64]	34
1.19	Technique HOMER [29]	35
1.20	Technique SQUAD [69]	36
1.21	Quelques contrôleurs physiques	37
1.22	Menu graphique 3D [71]	37

1.23	a) Un EVC pour la visualisation de molécules [72] b) CEDAR : EVC pour construire et visualiser des pipelines [76]	40
1.24	La course se réalise avec deux participants. L'écran est partagé, les utilisateurs peuvent communiquer en direct mais agissent uniquement sur leur véhicule, ils ne peuvent pas agir sur le véhicule adverse.	40
1.25	Répartition des données partagées entre les utilisateurs [80])	42
1.26	A : Répartition de données partagées par métaphores homogène(a : aperçue général de la co-manipulation, b : scénario de la co-manipulation, c,d : vues des différents utilisateurs [78]. B : Répartition de données partagées par métaphores hétérogène(a : vue du premier utilisateur, b : vue du deuxième utilisateur [77].	42
1.27	a : Elastique(les rayons interagissent avec le cube). b : Coudé (les rayons interagissent avec le cube). c : Déformable (les rayons interagissent avec le cube) [82].	44
1.28	Accès concurrent aux données partagées ([79]	44
1.29	Manipulation collaborative de deux utilisateurs en utilisant la technique SkeweR [80]	45
2.1	Les gants « CyberGlove II » [95]	55
2.2	L'Animazoo"IGS-190" contient 18 capteurs pour la reconnaissance de geste en plus des marqueurs visuels [95]	55
2.3	Manipulation d'un objet 3D sur un écran en utilisant le capteur inertiel d'un smartphone [97]	57
2.4	Le principe de la technologie 3DTouch d'Apple	58
2.5	A) radar solide à courte portée, captant l'énergie réfléchi par les mains, B) le signal résout le mouvement dans la gamme des millimètres, C) la reconnaissance de gestes spécialement conçue pour reconnaître des gestes subtils et peu exigeants en se basant sur le signal Soli[98].	59
2.6	Interaction avec des appareils intelligents par le capteur Soli [99]	59
2.7	concept de Airtouch montrant l'espace de détection de la main [100]	60
2.8	Représentation des gestes de la main, adaptée de [102]	61
2.9	Technique basée sur les contours actifs [104]	62
2.10	Technique de suivi et de reconnaissance en utilisant les couleurs [105]	63
2.11	Technique de gants colorés proposée dans [106]	64
2.12	Technique d'interaction basée marqueurs couleurs [107]	65
2.13	Technique de reconnaissance de gestes par les propriétés de la main [108]	66

2.14	Diagramme du flux de la technique de reconnaissance de geste par filtre de Gabor[118]	67
2.15	La table interactive proposée dans [125]	68
2.16	Principe du projet "RemoteGestures" [126]	68
2.17	Quelques exemples de caméras de profondeurs. a) Creative Camera1. b) Azure Kinect DK. c) Xtion Pro2. d) Kinect. e) Kinect 2.f) Zed3 camera	69
2.18	Exemple d'un modèle 3D de la main [127]	70
2.19	Modèles 3D générés [130]	71
2.20	Modèles 2D à partir des contours d'après (Imai, 2004)	72
2.21	Le principe de la technique 3Gear [132]	73
2.22	Technique de reconnaissance et de suivi de geste de la main proposée dans [134]. a) Initialisation : modèle estimé de geste. b) Convergence vers le geste réel en appliquant une fonction d'optimisation de l'énergie. c) Exemple d'application en RV [133]	74
2.23	Déroulement de la technique d'interaction : 1) pointage vers l'objet désiré. 2) un zoom est appliqué sur l'image. 3) l'utilisateur utilise la main virtuelle simple pour saisir et manipuler l'objet. [137]	75
2.24	Le dispositif Leap-Motion [143]	76
2.25	Cameras RealSense de Intel	77
2.26	Les gestes prédéfinis par le système RealSense	78
2.27	Casques a) : Oculus Rift DK1, b) : Oculus Rift DK2, c) : oculus Rift-S, d) : HTC Vive, e) : Oculus Quest.	79
3.1	Dispositif matériel utilisé [1]	87
3.2	L'architecture logicielle du système développé[1]	88
3.3	Les différents gestes utilisés	88
3.4	Les résultats obtenus : a : navigation dans l'environnement virtuel, b : sélection des objets, c : manipulation des objets sélectionner [1]	89
3.5	Dispositif immersif utilisé [3]	91
3.6	Architecture logicielle du système développé[3]	92
3.7	Les différentes articulations de la main fournis par contrôleur de Leap-Motion [149]	93
3.8	Le principe de fonctionnement de la technique "Zoom-fwd" [3]	95
3.9	Déplacement dans l'environnement virtuel[3]	96
3.10	Illustration du fonctionnement de la partie Zoom-sélection [3]	98

3.11	Le principe du fonctionnement de la partie Zoom-sélection [3]	99
3.12	principe de fonctionnement de la sélection précise : a : la sélection de tous les objets appartenant au même entourage, b : sélection de l'objet souhaité.	101
3.13	Illustration de la partie sélection précise : a) sélection du premier crayon, b) sélection du deuxième crayon, c)sélection du troisième crayon, d) sélection du gobelet [3].	102
3.14	Principe de fonctionnement de la partie sélection occultée[3]	104
3.15	Premier Scénario de sélection : sélection de la tablette [3]	107
3.16	Deuxième scénario de sélection : sélection du livre [3]	108
3.17	Troisième scénario de sélection : sélection d'une statuette [3]	108
3.18	Le temps d'exécution des différentes tâches d'interaction [3]	114
3.19	Comparaison entre une tâche complexe et une tâche simple : premier cas	114
3.20	Comparaison entre une tâche complexe et une tâche simple : deuxième cas	115
3.21	Scores pour chaque question [3]	117
3.22	Résultats subjectifs groupés par critères [3]	117
3.23	Signification du score au SUS [152]	119
4.1	Le modèle 3C par Ellis [154]	125
4.2	Le modèle représentant un EVC avec une interaction 3D collaborative	126
4.3	Dispositif matériel utilisé.	131
4.4	Environnement virtuel développé.	131
4.5	L'architecture du système global	132
4.6	Manipulation collaborative asynchrone	135
4.7	Manipulation collaborative simultanée en séparant les degrés de liberté	136
4.8	Manipulation collaborative simultanée en utilisant une priorité	137
4.9	Le protocole expérimental	140



Liste des tables

1.1	Grille des échelles	47
1.2	Classification des techniques de sélection et de manipulation	48
2.1	Classification des techniques de reconnaissance de gestes	82
3.1	Test de normalité	112
3.2	Analyses descriptives et statistiques pour les données objectives [3]	113
3.3	Les questions post-expérimental et les moyennes des notes données pour chaque technique	116



Liste des abréviations

RV :Réalité Virtuelle

RA :Réalité Augmentée

3D :Trois Dimensions

EVC :Environnement Virtuel Collaboratif

TC :Travail Collaboratif

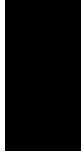
IHM :Interaction Homme Machine

VO :Vision par Ordinateur

EV :Environnement Virtuel

TCAO :Travail Collaboratif Assisté par Ordinateur

NUI :Natural User Interface



Résumé

La réalité virtuelle est une discipline qui se situe à la croisée de plusieurs domaines tels que l'Interaction Homme Machine (IHM), la Vision par Ordinateur et le Travail Collaboratif Assisté par Ordinateur. L'interaction 3D est la composante motrice de la réalité virtuelle. Elle permet à l'homme d'être un acteur capable de changer le cours des événements dans un environnement synthétique et ainsi d'interagir avec des entités virtuelles. Il existe différentes techniques d'interaction 3D qui sont dédiées aux différentes tâches d'interaction 3D à savoir la navigation, la sélection, la manipulation et le contrôle d'application. Dans la plupart des cas, les techniques d'interaction 3D traitent l'interaction d'une manière générale et ne répondent pas aux exigences de certaines applications telles que la précision en particulier avec des objets distants, occultés et se trouvant dans un milieu encombré.

Les environnements collaboratifs quant à eux, sont des mondes peuplés, dans lesquels les utilisateurs interagissent entre eux pour réaliser des objectifs ensemble. La complexité de ces environnements est liée à l'interaction d'un groupe d'utilisateurs avec des entités partagées.

Dans ce travail, nous avons développé dans un premier temps une technique d'interaction 3D "*Zoom-fwd*" basée sur la reconnaissance de gestes, cette dernière permet à l'utilisateur de sélectionner et/ou manipuler précisément des objets virtuels distants et occultés sans avoir à se déplacer. Dans un second temps, nous avons présenté un formalisme d'interaction

3D collaborative, ce dernier coordonne les actions d'un groupe d'utilisateurs sur un environnement partagé tout en gardant la cohérence de ce dernier. Notre technique d'interaction 3D peut être considérée comme un moyen de contourner les limitations de certaines techniques d'interaction 3D classiques afin de les rendre utilisables dans des contextes différents.

Mots clés : Réalité virtuelle, Interaction 3D, Techniques d'Interaction 3D, Interaction 3D Collaborative, Travail Collaboratif, Environnements Virtuels Collaboratifs.

Abstract

Virtual Reality is a research domain that is situated at the crossing of several domains as Human Computer Interaction (HCI), computer aided design and collaborative work, etc. The 3D interaction is a core component of virtual reality. It allows the user to be an actor able to interact with virtual entities. There exists different 3D interaction techniques that are dedicated to the different tasks of 3D interaction that are navigation, selection, manipulation and command control. In most cases, the 3D interaction techniques are only used to perform one of the 3D interaction tasks without taking into account the requirements of some applications such as precision especially with distant, occluded objects in a crowded environment.

The collaborative environments are populated worlds, in which users interact between them to accomplish goals together. The complexity of these environments is related to the interaction of a group of users with shared entities.

In this work, the "*Zoom-fwd*" which is an efficient 3D interaction technique is presented. The proposed technique uses gesture recognition for different 3D interaction tasks like

selection and manipulation. This new approach allows an efficient interaction with distant and occluded objects, while providing a precise selection, even when the environment is crowded. On the other hand, we presented a collaborative 3D interaction formalism, which coordinates the group of user's actions on a shared environment.

Our 3D interaction assistance system is considered as a mean to bypass the limitations of some 3D interaction classical techniques in order to make them usable in different contexts.

Key Words : Virtual Reality, 3D interaction, 3D Interaction Techniques, 3D collaborative Interaction, Collaborative work, Collaborative Virtual Environments.

ملخص

الواقع الافتراضي هو تخصص يسمح للإنسان بالتفاعل مع الحاسوب (IHM)، وكذا خلق رؤية حاسوبية من أجل العمل التعاوني بمساعدة الكمبيوتر.

التفاعل ثلاثي الأبعاد هو المكون الرئيسي للواقع الافتراضي، حيث يسمح للإنسان أن يكون عنصراً فاعلاً قادراً على تغيير مسار الأحداث في بيئة اصطناعية وبالتالي التفاعل مع الكيانات الافتراضية. للتفاعل ثلاثية الأبعاد عدة تقنيات، وهي مختلفة من حيث المهام على سبيل المثال (التنقل والاختيار والمعالجة والتحكم في التطبيق). في معظم الحالات، تتعامل تقنيات التفاعل ثلاثي الأبعاد مع التفاعل بطريقة عامة ولا تفي بمتطلبات تطبيقات معينة مثل الدقة، لا سيما مع الكائنات البعيدة والمخفية وكذا الموجودة في بيئة مزدحمة.

من ناحية أخرى، فإن البيئات التعاونية هي عوالم مأهولة يتفاعل فيها المستخدمون مع بعضهم البعض لتحقيق الأهداف معاً. يرتبط تعقيد هذه البيئات بتفاعل مجموعة من المستخدمين مع الكيانات المشتركة.

في هذا العمل، قمنا أولاً بتطوير تقنية تفاعل ثلاثية الأبعاد "Zoom-fwd" تعتمد على التعرف على الإيماءات كمرحلة أولى، والتي تسمح للمستخدم بتحديد و / أو التعامل بدقة مع الكائنات الافتراضية البعيدة والمخفية دون الحاجة إلى التحرك. في المرحلة الثانية، قدمنا شكليات تفاعلية ثلاثية الأبعاد، يقوم هذا الأخير بتنسيق إجراءات مجموعة من المستخدمين في بيئة مشتركة مع المحافظة على اتساق الأخيرة.

يمكن اعتبار أسلوب التفاعل ثلاثي الأبعاد الخاص بنا كوسيلة للتحايل على قيود المرتبطة ببعض تقنيات التفاعل ثلاثية الأبعاد الكلاسيكية، وهذا من أجل جعلها قابلة للاستخدام في سياقات مختلفة.

الكلمات الرئيسية: الواقع الافتراضي، التفاعل ثلاثي الأبعاد، تقنيات التفاعل ثلاثية الأبعاد، التفاعل ثلاثي الأبعاد التعاوني، العمل التعاوني، البيئات الافتراضية التعاونية.



Introduction générale

Contexte

Les travaux de cette thèse s'inscrivent dans le domaine de l'interaction homme-machine (IHM) et de la réalité virtuelle (RV) et porte sur "contribution à l'interaction 3D collaborative en réalité virtuelle". Elle consiste à étudier et à améliorer l'interaction 3D qu'elle soit pour un environnement virtuel mono-utilisateur ou un environnement virtuel collaboratif. Nous commençons tout d'abord par identifier les différents verrous et challenges de l'interaction 3D dans un environnement virtuel mono-utilisateur et collaboratif, et apporter des solutions aux techniques d'interactions 3D existantes ou proposer des nouvelles modalités d'interactions 3D. Nous validons ces solutions par des protocoles d'évaluation. Les travaux de cette thèse ont été menés au sein de l'équipe "Interaction homme-système Réalité Virtuelle/Augmentée" (IRV@) de la Division Productique et Robotique (DPR) du Centre de Développement des Technologies Avancées (CDTA). Ces travaux s'inscrivent dans le cadre du projet "Interaction 3D Multimodale et Collaborative dans un environnement de Réalité Virtuelle et Augmentée (IM@REV)".

Problématique de recherche

Ce travail concerne principalement le domaine de la réalité virtuelle dont les problématiques se situent au croisement de plusieurs disciplines notamment l'Interaction Homme Machine(IHM), la Vision par Ordinateur (VO)et le Travail Collaboratif Assisté par Ordinateur(TCAO). Ce domaine tente de plus en plus de nous fasciner et de changer notre manière de percevoir et d'analyser le monde qui nous entoure. Il permet à l'utilisateur d'être un acteur capable de changer le cours des évènements dans un environnement synthétique et ainsi d'interagir avec des entités virtuelles. Afin de visualiser et/ ou interagir avec le monde virtuel de nombreux périphériques matériels sont utilisés. Depuis le premier casque présenté par Sutherland en 1968 jusqu'à ce jour, nous avons pu voir une multitude de dispositifs de visualisation et/ou d'interaction tels-que casques, tablettes, smartphones, Kinect, Leap-Motion, etc. L'interaction 3D est la composante motrice de la réalité virtuelle. Elle permet à l'utilisateur d'interagir avec les composantes de l'environnement virtuel. Aujourd'hui, plusieurs techniques d'interaction 3D existent et chacune d'elles est dédiée à une des quatre tâches qui sont la navigation, la sélection, la manipulation et le contrôle d'application. Les techniques d'interaction 3D traitent l'interaction d'une manière générale et ne répondent pas aux exigences de certaines applications telles que la précision en particulier avec des objets distants, occultés et se trouvant dans un milieu encombré. Les environnements collaboratifs quant à eux, sont des mondes habités par plusieurs utilisateurs, dans lesquels ils interagissent entre eux pour réaliser des objectifs communs. La complexité de ces environnements réside dans leur interaction avec des entités partagées. En effet, lorsque plusieurs utilisateurs évoluent dans le même univers, il y a plusieurs contraintes à respecter notamment celles liées à la coordination des actions des différents utilisateurs. Ces avancées technologiques et scientifiques ont permis l'émergence des interfaces naturelles

(Natural User Interfaces), ces interfaces offrent un moyen d'interaction naturel et intuitif et cela en utilisant un des canaux de communication du corps humain (voix, toucher, geste, etc.). Actuellement, des limitations subsistent encore que ce soit sur le plan technologique dans le domaine du Travail Collaboratif ou encore au niveau des techniques et dispositifs d'interaction. Dans cette thèse, nous nous intéressons plus particulièrement aux questions suivantes :

1. Comment peut-on améliorer l'interaction 3D dans un environnement virtuel ?
2. Comment coordonner les différentes actions des utilisateurs pour garantir une cohérence d'un environnement virtuel partagé ?

Pour répondre à ces questions, il est nécessaire de s'intéresser d'une part aux différentes techniques d'interactions notamment aux techniques de sélection et de manipulations des objets virtuels et aux techniques de reconnaissances de gestes, et d'autre part, aux environnements virtuels collaboratifs et aux différentes techniques d'interaction virtuelles collaboratives. Dans cette optique, les objectifs qu'on s'est fixé sont au premier lieu de proposer une technique d'interaction 3D basé sur la reconnaissance de gestes, permettant une interaction naturelle, intuitive, rapide et précise notamment avec les objets distants, occultés et se trouvant dans un milieu encombré. Deuxièmement de formaliser les différentes actions collaboratives dans un environnement multi-utilisateurs.

Organisation de la thèse

Le manuscrit est organisé en deux grandes parties.

La première partie est consacrée à l'état de l'art et la seconde à la contribution. Chacune de ces deux parties est constituée de deux chapitres que nous résumons ci-dessous.

Le premier chapitre porte sur les concepts liés à l'interaction (techniques, paradigme et métaphore d'interaction) ainsi qu'aux différentes tâches d'interaction. Une classification des techniques d'interaction 3D en fonction des tâches d'interaction à savoir les techniques de sélection, les techniques de manipulation et les techniques de contrôle d'application a été élaborée. Nous présentons par la suite les techniques d'interaction virtuelles collaboratives ainsi que les différents modes de partages de données. Un bilan est dressé à la fin du chapitre et une discussion qui résume les différentes techniques d'interactions 3D dans un environnement habité par un ou plusieurs utilisateurs sont donnés. Dans le deuxième chapitre les différentes techniques et technologies existantes d'acquisition et de reconnaissance de gestes de la main pour l'interaction en RV sont présentés, ainsi que leur classification. A la fin de ce chapitre, nous dressons un bilan qui résume les différentes approches et techniques. Le troisième chapitre présente notre contribution dans le domaine de l'interaction 3D en RV. La première contribution concerne l'implémentation de la technique "Main virtuelle simple" en utilisant la reconnaissance de gestes [1]. Suite aux problèmes rencontrés durant l'élaboration de ce travail, nous avons proposé une nouvelle technique d'interaction "*Zoom-fwd*" [2] [3]. "*Zoom-fwd*" qui est une technique de sélection se basant sur le principe du zoom, pour répondre aux problèmes de sélection et de manipulation des objets virtuels distants et occultés dans un environnement encombré. Une évaluation est présentée à la fin de ce chapitre, afin d'évaluer cette technique d'interaction. Le quatrième chapitre présente un formalisme de l'interaction 3D Collaborative. Celui-ci va pouvoir nous aider à coordonner les actions de groupe des différents utilisateurs sur un environnement partagé. Afin de valider ce formalisme un exemple d'EVC à deux utilisateurs a été présenté ainsi que les différents résultats obtenus. Un protocole expérimental a été proposée. Une conclusion générale viendra clôturer cette thèse avec une mise au point sur ce qui a été présenté dans ce rapport et des perspectives à envisager comme suite de ce travail.

Chapitre

1

Interaction 3D en Réalité Virtuelle

1.1 Introduction

L'homme a toujours usé de son imagination pour s'évader de la réalité quotidienne pour différentes raisons (artistiques, culturelles ou professionnelles), et les moyens lui permettant d'y parvenir ont évolué au fur et à mesure avec les avancées technologiques. Ainsi, nous pouvons citer la peinture, les livres, la musique, la photographie et le cinéma. Cependant toutes ces représentations audiovisuelles sont figées, dans le sens où l'homme est seulement spectateur, contrairement à la Réalité Virtuelle (RV) qui lui procure la possibilité d'être acteur dans cette expérience. Que représente alors cette nouvelle technologie qui nous permet de réaliser ce que nous ne pouvons pas faire habituellement dans le monde réel ?

La RV est un oxymore, une expression qui désigne d'elle-même une contradiction : la réalité représente ce qui est réel, alors que le virtuel désigne ce qui ne l'est pas. Malgré son apparition au milieu du 19ème siècle avec l'invention d'un outil qui a reproduit une image en trois dimensions, le terme RV a été introduit pour la première fois en 1988 par [4]. Il décrit la RV comme étant une technologie qui utilise les outils pour créer un monde synthétisé, habité par plusieurs personnes, que nous pouvons appréhender par les cinq sens et avec lesquelles nous pourrions interagir, le tout par l'intermédiaire d'artefacts informatisés :

"We are speaking about a technology that uses computerized clothing to synthesize shared reality. It recreates our relationship with the physical world in a new plane, no more, no less. It does not affect the subjective world ; it does not have anything to do directly with what is going on inside your brain. It only has to do with what your sense organs perceive" [4].

La RV peut être définie par sa finalité, ses applications, ses fonctions ou encore les techniques d'interaction 3D sur lesquelles elle repose [5]. Plusieurs communautés scientifiques existent chacune définie la RV selon la catégorie au-quelle ils appartiennent.

Le Travail Collaboratif Assisté par Ordinateur (CSCW :Computer Supported Cooperative Work) a été introduit par Ellis [6]. Le travail coopératif regroupe toutes les activités humaines faisant intervenir un groupe de personnes, d'une manière simultanée ou pas. Ces activités sont diverses telles que la conception de produits industriels, l'enseignement, les relations commerciales, les jeux, etc.

La convergence de ces deux domaines a donné naissance aux Environnements Virtuels Collaboratifs (EVCs). Dans ce chapitre, nous commençons tout d'abord par un bref aperçu des concepts et définitions de base de la réalité virtuelle. Nous définissons ensuite les concepts liés à l'interaction (techniques, paradigme, métaphore d'interaction et les EVCs). Nous nous concentrons particulièrement sur les techniques d'interaction 3D (mono-utilisateurs et collaboratives). La première partie sera consacrée aux différentes techniques d'interaction mono-utilisateurs ainsi que leur classification selon les différentes tâches d'interaction. Nous abordons lors de la deuxième partie une classification des techniques d'interaction 3D collaboratives existantes dans la littérature. Un bilan et une discussion seront donnés pour conclure ce premier chapitre.

1.2 Réalité virtuelle : historique et définitions

La RV est un domaine de recherche qui se situe à la croisée de plusieurs disciplines telles que l'informatique graphique, la conception assistée par ordinateur, la simulation, la télé-opération, l'audiovisuel, le travail collaboratif. Ce classement des techniques d'interaction 3D collaboratives existantes dans la littérature. Un bilan et une discussion seront donnés pour conclure ce premier chapitre.

1.2 Réalité virtuelle : historique et domaine de recherche développe et utilise de nombreux périphériques matériels et des techniques logicielles pour chaque domaine applicatif [7]. Aujourd'hui, la définition de la RV n'est pas unique. Elle est désignée par plusieurs termes que nous trouvons dans la littérature scientifique tels que "Environnement Virtuel", "Espace Virtuel", "Monde Artificiel", "Réalité Synthétique" et bien d'autres encore. Les définitions de chacun de ces termes comportent, d'un domaine à l'autre, des nuances, parfois importantes. Les chercheurs d'un domaine ne donnent pas forcément la même définition que ceux d'un autre domaine pour le même terme. A titre d'exemple, la conception d'un philosophe de la RV, ne correspond pas à celle d'un informaticien. Cela dépend de l'usage du mot en question qui peut changer en fonction du contexte d'application.

Aukstakalnis et Blatner [8] considèrent la RV comme un moyen qui permet aux humains de visualiser, manipuler et interagir avec des ordinateurs et des données extrêmement complexes :

"A way for humans to visualise and interact with the artificial 3D environments created using computer graphics" [8].

Selon Burdea [7] les trois composantes de base pour la RV (voir Figure 1.1) sont : l'Immersion, l'Interaction et l'Imagination. L'utilisateur interagit dans un environnement

virtuel qui doit être représenté de manière réaliste pour donner une sensation d'immersion chez l'utilisateur. Cet environnement doit réagir en temps réel aux actions de l'utilisateur.

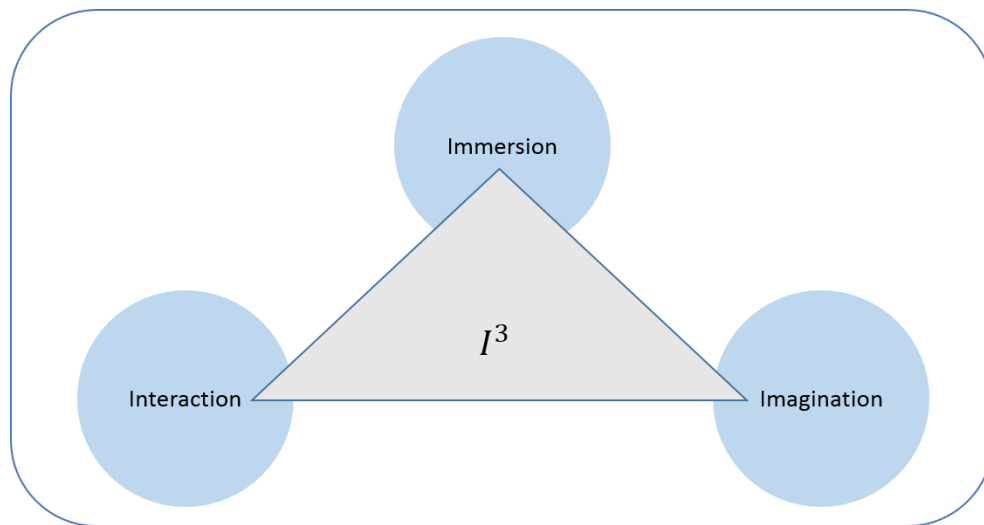


FIGURE 1.1 – Les trois composantes de RV [7]

Certains auteurs considèrent cette nouvelle technologie comme une extension des Interfaces Homme-Machine (IHM) classiques. Les interfaces résultantes dites "avancées" simulent un environnement réaliste et permettent à des participants d'interagir avec ceux-ci :

"Virtual Reality is an advanced human-computer interface that simulates a realistic environment and allows participants to interact with it" [6].

Fuchs [9] dans son livre intitulé "le traité de la réalité virtuelle " a donné deux définitions de la RV :

Une définition fonctionnelle :

"La réalité virtuelle va permettre de s'extraire de la réalité physique pour changer virtuellement de temps, de lieu et (ou) de type d'interaction : interaction avec un environnement simulant la réalité ou interaction avec un monde imaginaire ou symbolique" [9].

Une définition technique :

"La réalité virtuelle est un domaine scientifique et technique exploitant l'informatique et des interfaces comportementales en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en interaction en temps réel entre elles et avec un ou des utilisateurs en immersion pseudo-naturelle par l'intermédiaire de canaux sensori-moteurs". [9]

Quant à notre cas, nous considérons la RV comme une technologie qui immerge un (ou plusieurs) utilisateur(s) dans un monde synthétique représentant un environnement réel ou imaginaire et qui permet à ces utilisateurs d'être des acteurs capables de changer les propriétés de l'environnement et d'interagir avec les différentes entités composant l'univers simulé.

Pour mieux comprendre cette thématique, certaines définitions sont nécessaires à la compréhension de la suite.

1.3 Applications en réalité virtuelle

Les systèmes en RV ont beaucoup évolué durant cette dernière décennie avec l'avancement technologique. De par son originalité et sa pluridisciplinarité en matière de technologie, la

RV est impliquée dans plusieurs domaines. On pourra citer notamment les applications potentielles dans les domaines suivants : l'éducation [10] et [11], la médecine [12] et [13], l'architecture [14], l'industrie/maintenance, la robotique, ou encore les applications militaires [15] et [16].

1.4 Définitions

1.4.1 Réalité augmentée

Plusieurs définitions ont été données à la Réalité Augmentée (RA), dont les plus importantes sont citées ci-dessous. Azuma [17] a défini la Réalité Augmentée (RA) comme suit :

"Un système de réalité augmentée complète le monde réel avec des objets virtuels (générés par ordinateur) de telle sorte qu'ils semblent coexister dans le même espace que le monde réel" [17].

Cette définition est orientée vers des systèmes qui augmentent la vision. Mais toujours selon l'auteur, la RA peut potentiellement s'appliquer à tous les sens tels que le toucher ou l'ouïe. Les trois règles de base nécessaires pour le fonctionnement d'un système de RA :

- Combiner le réel et le virtuel ;
- Respecter les contraintes d'interactivité et de temps ;
- Respecter l'homogénéité et la cohérence entre deux mondes réel et virtuel.

Fuchs et ses coauteurs [9] ont défini la réalité augmentée par :

"la réalité augmentée regroupe l'ensemble des techniques permettant d'associer un monde réel avec un monde virtuel, spécialement en utilisant l'intégration d'Images Réelles (IR) avec des Entités Virtuelles (EV) : images de synthèse, objets virtuels, textes, symboles, schémas, graphiques, etc. D'autres types d'associations entre mondes réels et virtuels sont possibles par le son ou par le retour d'effort" [9].

Otmane [18] a défini la RA quant à lui comme suit :

"La finalité de la réalité augmentée est de permettre à une personne (ou plusieurs) des interactions multi sensorielles (audio, vidéo et haptique) avec un environnement qui fait coexister les deux mondes virtuel et réel" [18].

D'une manière générale, la RA consiste à augmenter la scène réelle avec des informations virtuelles supplémentaires. Ces augmentations peuvent prendre différentes formes telle que du texte explicatif, des formes géométriques (des flèches, des objets 3D etc.). La RA trouve son application dans des domaines aussi divers que la médecine [19] (Voir Figure 1.2.a), l'assistance automobile [20] (Voir Figure 1.2.b), le marketing (Ikea¹) ou la maintenance industrielle [21] [22].

1. IKEA, <http://info.ikea-usa.com/Catalog/>

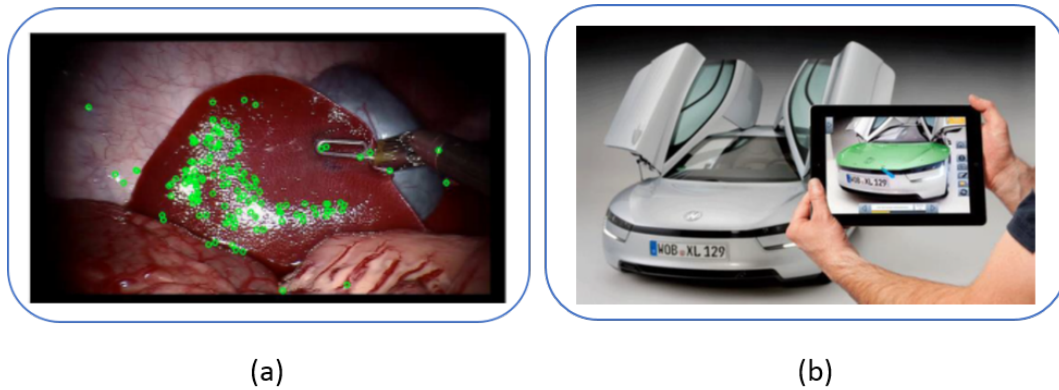


FIGURE 1.2 – Exemples d’applications en réalité augmentée : a) Superposition des vaisseaux sanguins sur la vidéo laparoscopique [19], b) Application MARTA (Mobile Augmented Reality Technical Assistance) [20]

Afin de résumer les différentes définitions données sur la RA et la RV, Milgram et ses associés [23] ont proposé un continuum linéaire qui va du réel au virtuel pour unifier les deux concepts de la RA et de la RV (voir Figure 1.3).

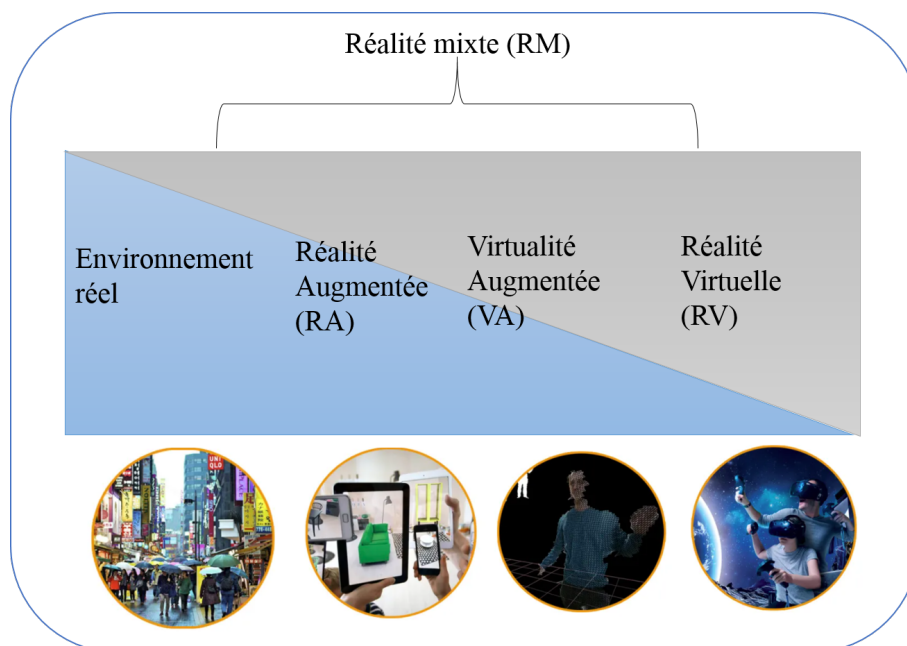


FIGURE 1.3 – Le continuum Réel-Virtuel [23]

1.4.2 Environnement virtuel

Le terme "environnement virtuel" (EV) a été introduit par des chercheurs du MIT (Massachusetts Institute of Technology) au début des années 90 comme synonyme de la réalité virtuelle. Heim et ses coauteurs[24] ont défini l'EV comme un espace interactif créé par des applications numériques dans le but d'accueillir un ou plusieurs utilisateurs et leur permettre d'accomplir certaines tâches avec l'impression d'être dans un cadre spécifique. L'EV est une représentation numérique tridimensionnelle d'un lieu réel ou imaginaire, dans lequel les utilisateurs peuvent interagir et explorer en temps réel.

Il existe différents types d'EVs selon le degré d'immersion qu'ils procurent à l'utilisateur :

- EV non-immersif (Non-Immersive Virtual Environment "NIVE");
- EV semi-immersif (Semi-Immersive Virtual Environment "SIVE");
- EV totalement immersif (Fully-Immersive Virtual Environment "FIVE") [25].

1.4.3 Interaction

Depuis l'apparition de la RV, les chercheurs se sont particulièrement intéressés à l'interaction qui peut être considérée comme la composante motrice de tout système interactif. L'interaction peut être définie comme étant un langage de communication entre l'homme et la machine. Ce langage correspond à l'ensemble des actions/réactions réciproques entre l'homme et l'ordinateur par l'intermédiaire d'interfaces sensorielles, d'interfaces motrices et de techniques d'interactions [26].

1.4.4 Paradigme et métaphore d'interaction

Le paradigme d'interaction 3D se réfère à l'approche et aux principes utilisés pour concevoir et faciliter l'interaction des utilisateurs avec des environnements virtuels. Il s'agit d'un ensemble de règles et de techniques qui visent à créer des expériences d'interaction fluides, intuitives et efficaces dans environnement mixte [27], [28] et [29]. La métaphore d'interaction en réalité virtuelle est un concept qui utilise des représentations ou des analogies de comportements et d'interactions du monde réel pour faciliter l'interaction des utilisateurs avec des environnements virtuels[26]. Selon Fuchs [9], une métaphore d'interaction est une image symbolique d'une action ou d'une perception utilisée pour réaliser une tâche précise dans un environnement virtuel. Otmane [18] a proposé de redéfinir l'interaction 3D par sa finalité. Ainsi, la finalité de l'interaction 3D est de permettre l'utilisation de dispositifs matériels et de techniques logicielles adaptées en vue d'une utilisation performante et crédible des tâches d'interaction 3D. Cela en proposant une démarche de recherche qui exploite les possibilités d'assister les tâches d'interaction 3D en considérant les éléments suivants :

- Le comportement des utilisateurs ;
- Les contraintes de l'environnement dans lequel évoluent les utilisateurs ;
- Les outils d'assistance (les guides virtuels et les retours multi sensoriels).

1.4.5 Environnement virtuel collaboratif (EVC)

Les EVCs sont vus comme une extension des technologies de la réalité virtuelle mono-utilisateur adaptées pour supporter plusieurs utilisateurs et plusieurs types d'applications. Une définition a été proposée par [30] :

"Un EVC est un système distribué qui permet à des individus de se partager des informations par l'interaction individuelle et collaborative dans un monde virtuel".

Un EVC est un environnement partagé par plusieurs utilisateurs, ces derniers agissent sur les mêmes objets pour réaliser un objectif commun. La complexité de ces environnements réside dans l'interaction du groupe d'utilisateurs avec les entités partagées. Plusieurs contraintes sont à respecter, notamment celles qui sont liées à la coordination des actions et à la communication des utilisateurs ainsi qu'à la gestion de l'interaction d'un utilisateur dans le monde virtuel vis à vis des objets et des autres utilisateurs [31].

1.4.6 Interaction 3D collaborative

Il existe des situations où la nature de la tâche est très complexe et difficile ou même impossible à réaliser par une seule personne, d'où la nécessité de faire appel à un groupe de personnes. Ces derniers doivent travailler en collaboration pour effectuer la tâche qui leur est assignée.

L'interaction de plusieurs utilisateurs avec l'environnement virtuel est limitée et de nombreuses recherches sont en cours.

L'objectif principal des recherches menées dans le domaine de l'interaction 3D multi-utilisateurs en RV est de permettre aux utilisateurs de collaborer et d'interagir de manière immersive dans un environnement virtuel partagé, et cela d'une manière facile et efficace.

Actuellement, on retrouve deux catégories de techniques d'interaction 3D multi-utilisateurs asynchrones et synchrones. Les techniques d'interactions collaboratives asynchrones permettent une interaction des différents utilisateurs à tour de rôle, pour celles synchrones deux catégories existent : une première catégorie sépare les degrés de liberté de l'objet à manipuler.

Dans ce cas, les utilisateurs utilisent des techniques d'interaction mono-utilisateur en agissant sur les degrés de liberté qui leurs sont affectés. Un mécanisme d'adaptation est mis en place. Pour la seconde catégorie, le mouvement final de l'objet dans l'EVC est calculé par une fonction à partir de la position et de l'orientation des utilisateurs.

Cela implique l'utilisation d'une nouvelle technique dans les cas multi-utilisateurs. Globalement toutes ces techniques manquent de considération pour les contraintes de coordination et de communication entre les utilisateurs et se concentrent exclusivement la manipulation[32].

1.5 Interaction 3D en réalité virtuelle

L'interaction est la composante principale dite motrice de la RV. Elle permet aux utilisateurs d'interagir dans l'EV [33]. Une technique d'interaction définit la méthode qui permet de réaliser une tâche dans l'environnement. C'est un protocole de communication avec le système qui représente les actions/réactions. Dans ce qui suit, nous détaillons les techniques utilisées pour interagir dans des environnements virtuels. Une première partie sera consacrée aux techniques d'interactions mono-utilisateurs (techniques pour un seul utilisateur), aux différentes tâches d'interaction 3D et à leur classification. La seconde partie quant à elle dédiée aux techniques d'interaction 3D collaboratives (techniques d'interactions 3D multi-utilisateurs) ainsi que leur classification.

1.5.1 Interaction 3D mono-utilisateur

1.5.1.1 Les tâches d'interaction 3D

Dans les environnements mixtes (virtuels ou augmentés), l'interaction est primordiale. Nous pouvons concevoir ou imaginer des multitudes de tâches à réaliser pour des applications spécifiques de RA ou de RV. Néanmoins considérer des tâches d'interaction de base est un moyen pratique qui permet de réaliser des tâches complexes dans toute application. Plusieurs classifications ont été proposées dans la littérature.

En 1995, Mine propose une première classification [27] basée sur quatre tâches fondamentales : la navigation, la sélection, la manipulation et la mise à l'échelle. Il a défini également une cinquième tâche qui dérive des quatre précédentes : les menus virtuels et l'interaction par widgets. En 1997, Hand et Chris [34] ont introduit les bases d'une nouvelle classification qui sera repris par Bowman en 1999 [29]. Cette dernière classe les différentes techniques d'interaction selon quatre tâches principales d'interaction 3D : navigation, sélection, manipulation et contrôle d'application. Nous nous baserons sur cette taxonomie au cours de notre étude.

a) Navigation

Comme dans le monde réel, l'utilisateur a besoin d'explorer et se déplacer dans la scène virtuelle pour réaliser certaines tâches. L'action "se déplacer" varie de l'action "bouger sa tête" pour observer un objet de plus près ou d'un autre angle, à l'action "se promener" dans un parc ou visiter un musée. La navigation définit tout ce qui concerne le déplacement de l'être humain dans un espace virtuel. Elle désigne l'ensemble des méthodes qui permettent de connaître la position d'un objet par rapport à un système de référence ou à un point fixe déterminé [35].

LaViola divise la tâche de navigation en deux sous tâches [35] :

Le déplacement : il représente la composante principale de la navigation ; il se rapporte au déplacement physique dans le monde virtuel (déplacement d'un endroit à un autre) ;

La recherche d'itinéraire (WayFinding) : elle représente la composante cognitive de la navigation. Cette tâche permet à un utilisateur de se repérer et de s'orienter dans le monde virtuel [35].

b) Sélection

L'action "sélectionner un objet" est une tâche très courante dans la vie quotidienne. En effet, pour pouvoir acquérir un objet, l'humain a toujours besoin de le désigner parmi d'autres objets. Ce processus s'appelle également désignation, il s'inspire souvent de la désignation dans le monde réel. La tâche de sélection, appelée aussi tâche d'acquisition d'une cible [36], représente la désignation d'un objet ou d'un ensemble d'objets afin d'accomplir un objectif donné au sein de l'environnement virtuel [37].

La Désignation : c'est le choix d'un ou plusieurs objets dans le monde virtuel ;

la Validation : c'est la confirmation de la sélection.

c) Manipulation

Les tâches précédentes permettent à l'homme d'avoir l'illusion d'habiter un monde virtuel, explorer ce monde, atteindre certains objets. Dans la plupart de ces cas, l'homme reste toujours un spectateur immergé dans son environnement virtuel. La tâche de manipulation quant à elle permet à l'utilisateur d'être un acteur capable de changer les propriétés de l'environnement virtuel ou des objets le constituant. Elle représente la composante active du système d'interaction 3D [31].

Avant de manipuler un objet on doit impérativement le sélectionner avant.

La tâche de manipulation peut être classée en trois sous-tâches :

Le positionnement, qui consiste à modifier la position d'un objet ;

La rotation, qui permet de changer l'orientation d'un objet ;

La mise à l'échelle, qui joue sur le facteur d'échelle d'un objet suivant les trois axes engendrant ainsi un rétrécissement ou un agrandissement.

Cependant, d'autres paramètres peuvent être modifiés, tels que la couleur, la texture, le maillage de l'objet et bien d'autres. La manipulation en RV peut être directe ou indirecte selon l'action de l'utilisateur [33] :

La manipulation directe, la plus naturelle, puisque l'utilisateur agit directement sur l'objet afin d'appliquer les modifications souhaitées ;

La manipulation indirecte, dans laquelle l'utilisateur modifie les propriétés de l'objet au travers d'entités graphiques 2D (menus, boutons, etc) ou 3D (menus, outils virtuels, widgets, etc).

d) Contrôle d'application

Cette tâche combine toutes les techniques qui permettent à un utilisateur de communiquer avec un système.

Pour ces tâches, plusieurs techniques d'interactions ont été proposées. Ainsi, l'utilisateur peut interagir différemment à travers une même tâche avec son monde virtuel. L'utilisateur peut, par exemple, bouger sa tête pour observer l'objet avec qui il veut interagir, comme il peut se déplacer vers ce dernier en trouvant le chemin optimum [29]. De ce fait, nous présentons dans ce qui suit une classification des différentes techniques d'interactions.

1.5.1.2 Classification des techniques d'interaction 3D

Plusieurs techniques d'interaction ont été développées pour les environnements 3D. Ces dernières sont classifiées en général selon les tâches d'interaction. L'état de l'art étant riche, plusieurs travaux récapitulent les différentes techniques d'interaction pour les environnements virtuels ainsi que leurs classifications notamment [38] et [39].

a) Les techniques de navigation

Les techniques d'interaction définies pour la tâche de navigation concernent le déplacement, autrement dit le contrôle du mouvement du point de vue de l'utilisateur, qui est l'un des fondements de base des environnements virtuels. Le déplacement est défini comme étant le contrôle du mouvement du point de vue de l'utilisateur dans un environnement virtuel selon Bowman [29]. Ce dernier propose une classification qui répartit les techniques de navigation en trois catégories selon le type de déplacement (recherche, exploration, manœuvre, etc.) :

- Choix de la direction ou de la cible : spécification de la direction ou de la cible visée par le déplacement ;
- choix de la vitesse/accélération du mouvement : variation de la vitesse et/ou de l'accélération du déplacement ;
- choix des conditions d'entrée : désignation des entrées requises par le système pour démarrer, continuer ou arrêter le déplacement.

Plusieurs techniques de navigation appartiennent aux trois catégories mentionnées précédemment [35].

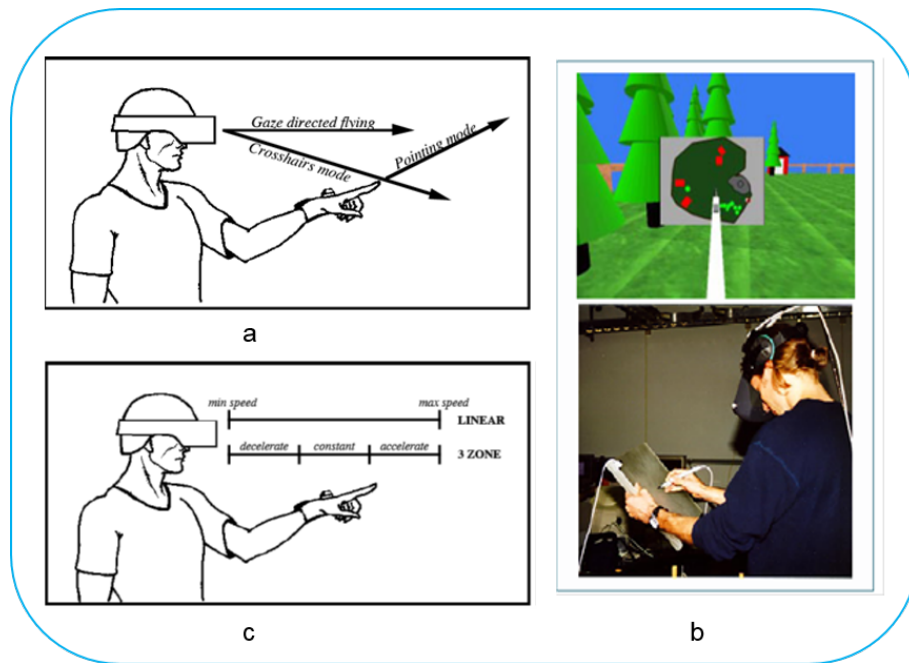


FIGURE 1.4 – Techniques de navigation : a) direction du regard, b) déplacement en utilisant une carte, c) déplacement en utilisant une accélération constante [40]

i. **Direction du regard** : la technique Direction du regard (Gaze directed steering) [27] est l'une des techniques de navigation qui utilise la direction du mouvement de l'utilisateur, et en particulier sa tête. Cette dernière permet de déplacer le point de vue dans la direction du regard à partir d'un suivi de la tête (tracking de la tête). L'inconvénient de cette technique c'est qu'elle ne permet pas à l'utilisateur de voir sur les côtés en se déplaçant (voir Figure. 1.4.a). De plus elle n'est efficace que dans des environnements virtuels totalement immersifs [35].

ii. **Direction de la main (pointing)** : l'utilisateur spécifie constamment la direction du mouvement en utilisant sa main. Cette technique assure le tracking de la main tout le temps (voir Figure 1.4.a). L'avantage de cette technique c'est qu'elle permet de se déplacer et de regarder dans différentes directions [35].

iii. Déplacement en utilisant une carte (Goal driven) : l'utilisateur déplace une icône ou un avatar, le représentant, sur une carte 2D par un stylet ou son doigt jusqu'à un nouvel endroit sur la carte (voir Figure 1.4.b), ce qui provoque son déplacement dans l'environnement virtuel [29].

iv. Déplacement en utilisant une accélération constante : l'utilisateur commence à se déplacer avec une vitesse lente ensuite la vitesse augmente continuellement jusqu'à ce que l'accélération devienne constante. Un ajustement de ce paramètre est nécessaire en fonction de la taille de l'environnement virtuel (voir Figure 1.4.c).

b) Les techniques de sélection et de manipulation

Étant donné que l'utilisateur est dans un EV, nous devons lui permettre d'interagir avec cet environnement. Dans ce sens, la sélection permet de désigner un ou plusieurs objets virtuels et de les impliquer dans le processus de manipulation. Il existe une variété de techniques qui permettent à l'utilisateur de sélectionner et de manipuler des objets virtuels. Bowman a classifié [29] les techniques de sélection et de manipulation par décomposition en tâches. Celles-ci se basent sur un ensemble de blocs qui constituent une tâche de sélection ou de manipulation. Chaque bloc exécute une action élémentaire par le biais d'un certain nombre de ses composants. Cette approche permet donc de structurer l'espace de conception et de modélisation des techniques d'interaction afin qu'on puisse mettre en place de nouvelles techniques d'interaction en utilisant les composants déjà existants.

Poupyrev et Ichikawa ont introduit une autre classification des techniques de sélection et de manipulation par métaphore [28]. En effet, la plupart des techniques d'interaction utilisent des métaphores de base ou une combinaison de ces métaphores. Chaque métaphore constitue le modèle fondamental d'une technique d'interaction. Une première classification

a divisé les techniques de sélection et de manipulation en deux grandes familles : les techniques exocentriques et les techniques égocentriques, selon la position de l'utilisateur et de la distance entre l'utilisateur et l'objet virtuel.

Laviola et ses coauteurs [35] ont repris dans leur livre les deux précédentes classifications et ils ont introduit une troisième classification basée sur l'isomorphisme. Les approches isomorphiques dites aussi naturelles, préservent une correspondance naturelle entre les actions de l'utilisateur et les effets résultants, alors que les techniques non-isomorphiques dites aussi magiques, permettent des interactions moins réalistes, potentiellement basées sur des outils virtuels ayant des effets "surnaturels". Généralement, ces techniques font référence à une métaphore de contact ou de pointage [41].

Nous présentons dans ce qui suit une classification basée sur les travaux de LaViola [35]. Cette classification est par métaphore d'interaction (saisie, pointage, métaphore indirecte et hybride).

1. Les techniques par métaphore de saisie (GraspingMetaphors) :

Dans cette catégorie de techniques, l'utilisateur sélectionne et/ou manipule les objets du monde virtuel en utilisant la métaphore de saisie.

- **Main virtuelle** : l'idée principale de cette technique est tiré de la métaphore de la main de l'utilisateur, elle permet à l'utilisateur d'utiliser sa propre main pour sélectionner et/ou manipuler des objets virtuels. En effet, il s'agit de la technique main virtuelle simple (Virtual hand) proposée par [42], reprise par plusieurs auteurs notamment [43] et [44]. L'objet désigné est touché par l'avatar de la main de l'utilisateur, la confirmation de la sélection est réaliser en fermant la poignet ou en restant en contact avec l'objet un certain temps. Les mouvements de la main de l'utilisateur sont transmis à l'avatar pour manipuler l'objet (voir Figure. 1.5).

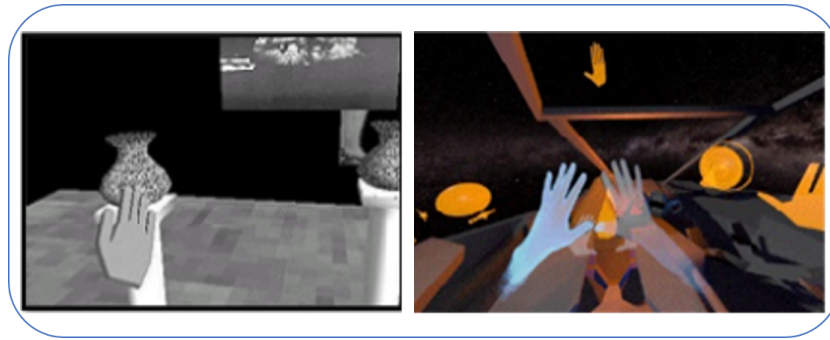


FIGURE 1.5 – Technique de la main virtuelle simple [42]

Bien que précise, naturelle et intuitive, cette technique pose plusieurs problèmes notamment l'accès aux objets distants ; l'utilisateur doit se déplacer vers ces objets jusqu'à ce qu'ils soient à sa portée.

- **La technique Go-Go** : cette technique vient palier aux problèmes de l'interaction avec les objets distants. Proposée par Poupyrev [45], elle est également appelée technique d'extension du bras (voir Figure. 1.6). Cette technique permet la sélection et la manipulation d'objets éloignés, en proposant une relation non-linéaire entre l'avatar de la main et la main réelle. En d'autres termes, le déplacement dans le monde réel se traduit par un grand déplacement dans le monde virtuel après un certain seuil. L'inconvénient avec la technique Go-Go est la sélection des objets petits et occultés. Les mouvements de faible amplitude de la main réelle sont traduits par des mouvements de grande amplitude par la main virtuelle [46]. Une extension de celle-ci a été proposée par Poupyrev [45] appelée Stretch Go-Go. Dans cette technique, le bras virtuel peut s'agrandir ou rétrécir selon sa position dans le monde virtuel. Ainsi, lorsque la main est dans une zone proche, celle-ci s'approche de l'objet avec une vitesse petite et constante. Dans le cas où nous sommes dans une zone intermédiaire, la main est au repos, et en zone distante, le bras s'allonge à une vitesse élevée et constante. Bowman et Hodges [47], ont proposé une autre variante appelée

FastGo-Go [47]. Comme pour la technique Go-Go, la position du bras est calculé par une fonction non-linéaire avec une valeur du seuil égale à zéro, ce qui fait que le bras s'allonge rapidement dès le début du mouvement.

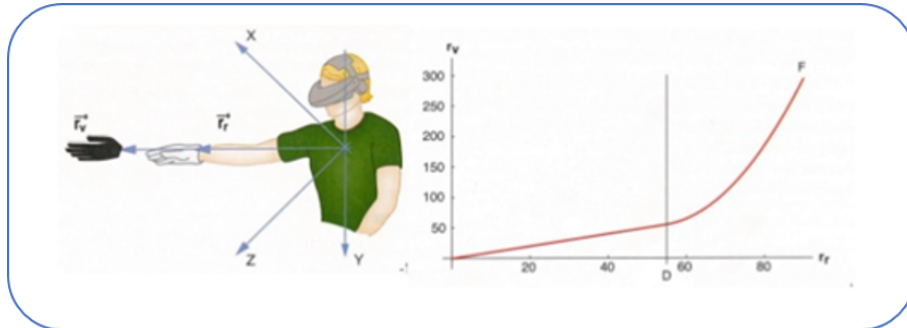


FIGURE 1.6 – La techniques Go-Go[45]

- **La technique PRISM** : proposée par Frees [48], cette technique est utilisée en association avec d'autres techniques telle-que Go-Go afin d'améliorer la précision des mouvements de la main de l'utilisateur (voir Figure. 1.7). En effet, les rotations et les translations se stabilisent lorsque les mouvements de la main ralentissent sous un certain seuil, diminuant automatiquement les effets des tremblements de la main dans le cas des manipulations précises.

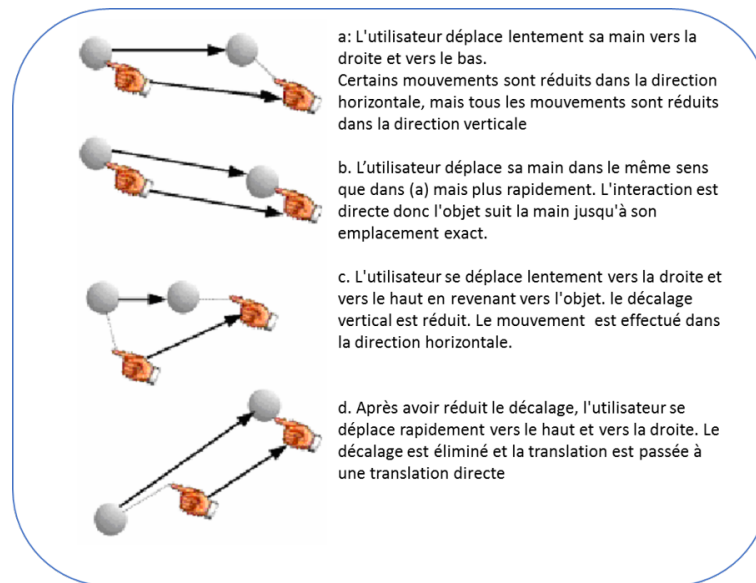


FIGURE 1.7 – Technique PRISM [48]

- **La technique 3D bubble cursor** : cette technique utilise la technique de main virtuelle pour contrôler un pointeur 3D (en guise de curseur 3D) et une sphère semi-transparente en 3D englobant la cible la plus proche [49]. Cette technique est efficace pour sélectionner de petites cibles dans des environnements denses, mais elle augmente considérablement le temps de sélection en raison des étapes supplémentaires qu'elle introduit. Les techniques de sélection utilisant la proximité des cibles semblent efficaces mais elles nécessitent l'utilisation d'un retour d'information approprié, notamment en 3D (Voir Figure. 1.8).

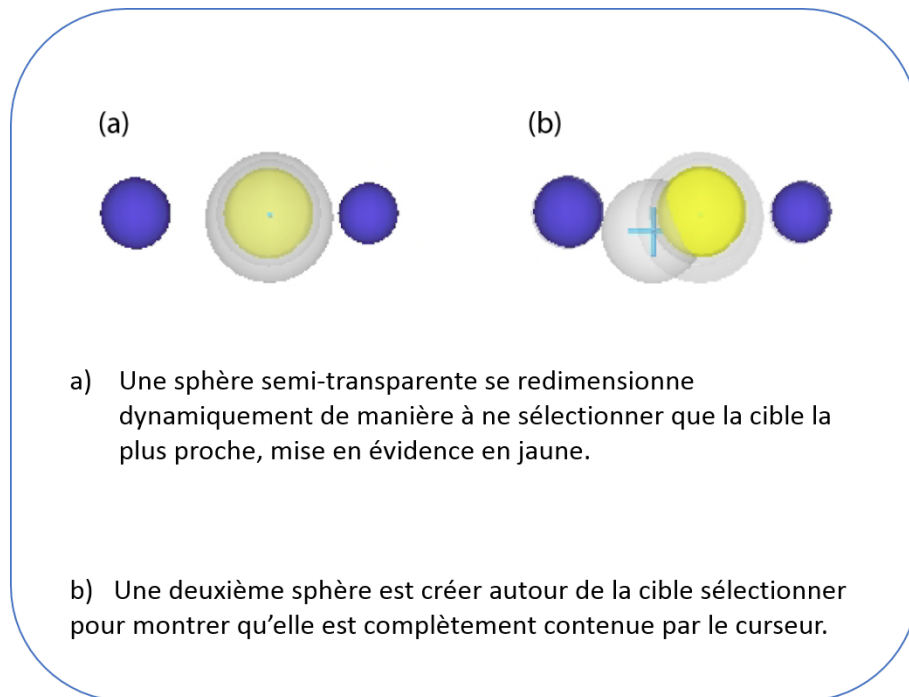


FIGURE 1.8 – Technique 3D bubble cursor [49]

2. Les techniques par métaphore de pointage (Pointing Metaphors)

Dans cette catégorie de techniques, l'utilisateur sélectionne et/ou manipule les objets du monde virtuel en utilisant la métaphore de pointage par la main.

- **La technique Ray-casting** : cette technique basée sur la métaphore du rayon virtuel, introduite par Bolt [50] puis reprise par plusieurs auteurs Zhai[36], De Amicis[51] et Argelaguet[43]. Un rayon laser infini est créé à partir de la main virtuelle de l'utilisateur. Le premier objet en intersection avec ce rayon laser est sélectionnable. Bien que très pratique, l'inconvénient avec la technique ray-casting c'est la sélection des objets petits ou lointains (voir Figure. 1.9).

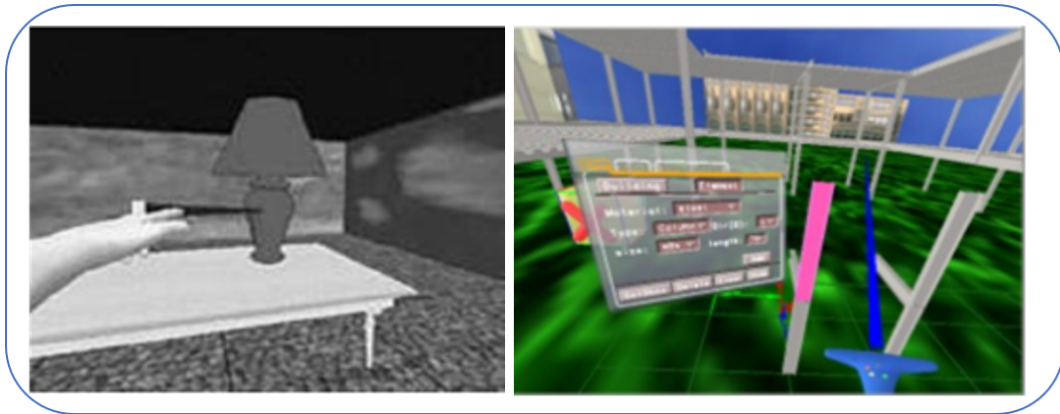


FIGURE 1.9 – Technique Ray-casting [50]

- **Les techniques Image plane pointing** : dans cette technique l'objet qu'on veut sélectionner ou manipuler se projettera sur un plan image 2D. Pierce et ses coauteurs [52] ont présenté un ensemble de variantes de la technique plan image, à savoir, la technique paume à plat (Lifting Palm), où l'objet souhaité doit être mis sous la paume de la main de l'utilisateur. La technique entre deux doigts (Head crusher), où l'utilisateur utilise ses deux doigts "pouce et indexe" et effectue un geste pour encadrer l'objet qu'il souhaite sélectionner (voir Figure. 1.10). La technique dirigée du doigt (StickyFinger) se base sur un rayon virtuel qui part de la tête de l'utilisateur passant par l'index de sa main pour sélectionner l'objet à manipuler. Pour finir, la technique de la main encadrante (Framing Hands), où la sélection s'effectue en encadrant par les deux mains réelles le ou les objets. Ces techniques ont été reprises par Lee afin de les évaluer [53].

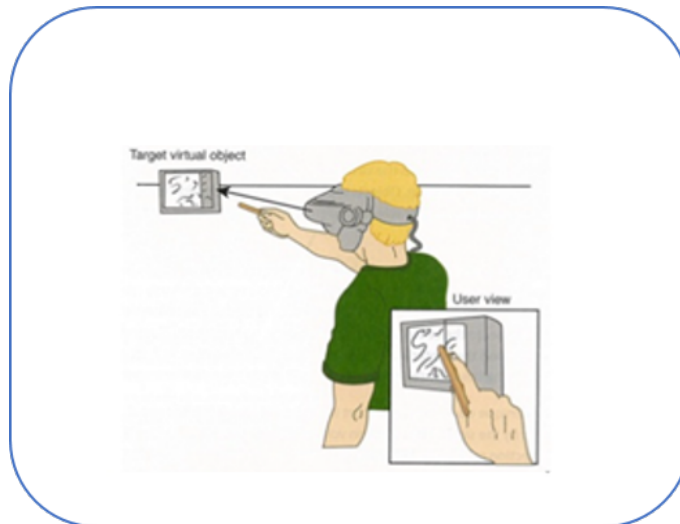


FIGURE 1.10 – Technique Image plane [52]

- **La technique du Ray-flexible** : cette technique propose un pointeur flexible pour résoudre le problème d'obstacles entre l'objet à sélectionner et la main virtuelle [54] (voir Figure. 1.11). l'avantage de cette technique c'est qu'elle permet de pointer plus facilement des objets occultés par d'autres objets dans la scène virtuelle. La courbure et la longueur de ce rayon est contrôlé par des deux mains de l'utilisateur.

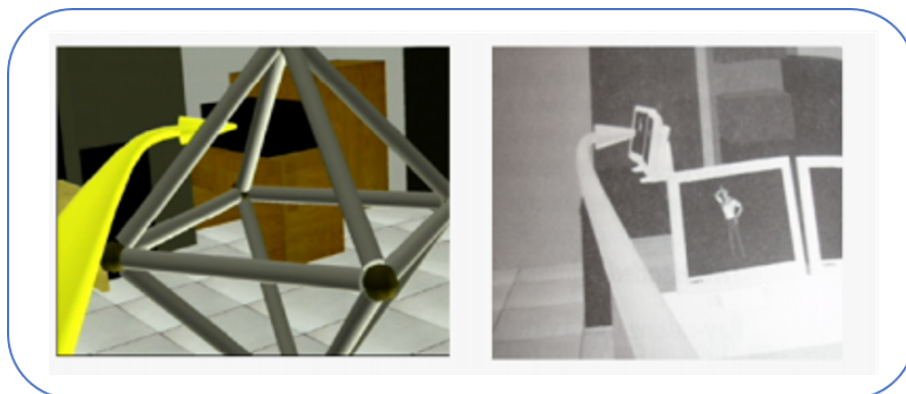


FIGURE 1.11 – Techniques Ray-flexible [54]

- **La technique Flashlight** : appelée aussi la technique spotlight, elle utilise un cône à la place du rayon pour sélectionner les objets distants ou petits où plusieurs

objets appartenant au cône peuvent être sélectionnables [55](Voir Figure 1.12.a). Une nouvelle variante de Flashlight appelée "Aperture" a été proposée à fin de varier l'angle d'ouverture du cône en fonction de la distance des objets à sélectionner [56] et [43](Voir Figure 1.12.b).

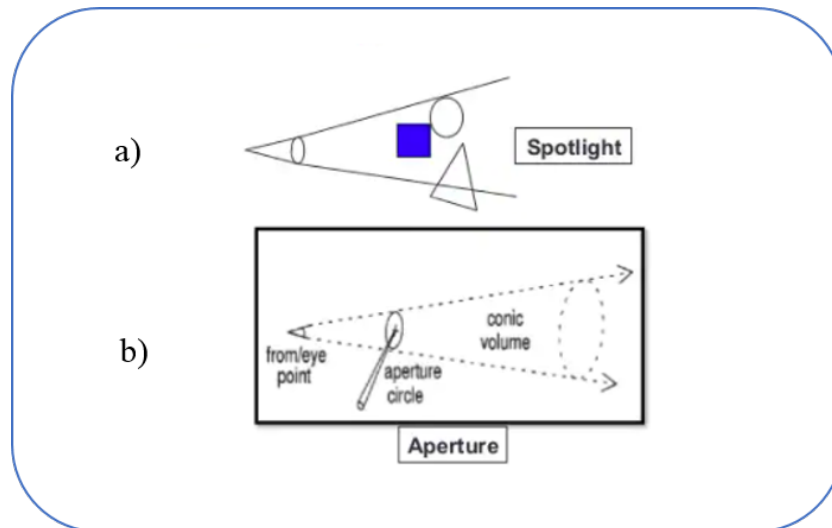


FIGURE 1.12 – a) Technique Flashlight [55] b) Technique Aperture [56]

- **La technique Eye-casting** : c'est une nouvelle approche de contrôle pour la sélection en utilisant la métaphore du rayon, appelée "Ray-casting from the Eye" [57]. Dans cette approche, l'origine du rayon est la position de l'œil mais son orientation est contrôlée par des rotations du poignet de la main (voir Figure 1.13). Puisque l'origine du rayon est l'œil donc l'ensemble des objets visibles et l'ensemble des objets sélectionnables est le même. Cette technique a montré une amélioration significative de la sélection dans un environnement encombré, cependant elle ne permet pas une sélection précise [43].

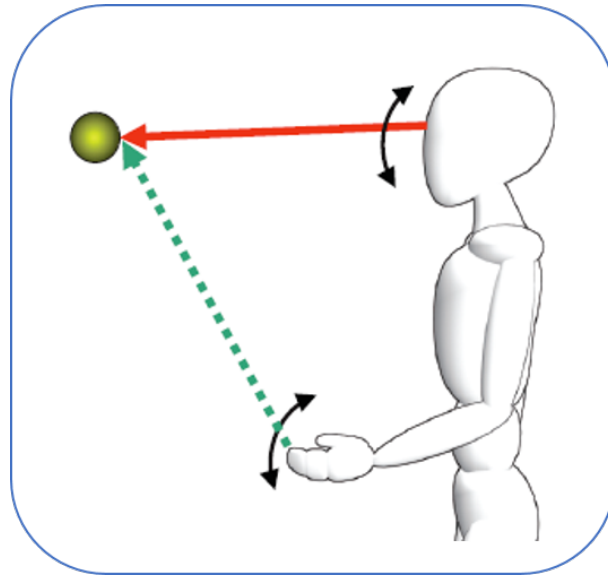


FIGURE 1.13 – Technique Eye-casting [57]

- **La technique Depth-Ray** : c'est une technique qui se base sur le ray-casting avec un marqueur fixe au milieu du rayon [58] et [49]. Parmi tous les objets intersectés par le rayon, celui le plus proche de ce marqueur est sélectionné (voir Figure.1.14). Cette technique permet de régler le problème d'occultation, puisque l'on peut ainsi sélectionner un objet qui se trouve derrière un autre. Cependant, les objets petits et éloignés restent difficiles à atteindre puisque le Depth-Ray ne permet pas de sélectionner les cibles qui ne sont pas traversées par le rayon.

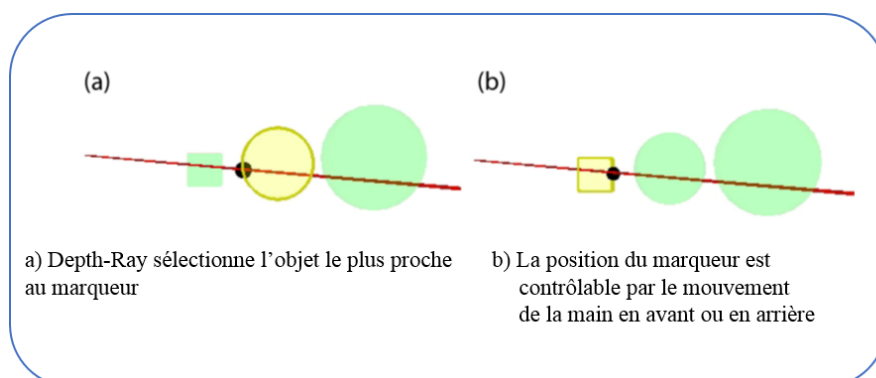


FIGURE 1.14 – Technique Depth-Ray [49]

3. Les techniques par métaphore indirecte On retrouve dans cette catégorie toutes les techniques de sélection et/ou de manipulation qui se font indirectement sur l'objet (la sélection et manipulation ne se font pas sur l'objet).

- **Le monde en miniature** : elle utilise une représentation en miniature de l'environnement virtuel et de permettre à l'utilisateur d'interagir indirectement sur les objets virtuels qui appartiennent à ce dernier [59](voir Figure.1.15). Les objets seront manipulés indirectement par l'utilisateur à travers la représentation en miniature du monde virtuel. Ainsi, ces actions sur les objets virtuels provoquent des actions réelles similaires sur un/des objets. Ce travail a été repris dans [60], ce dernier a été critiqué et évalué. L'inconvénient majeur de cette technique est la sélection et la manipulation des objets virtuels qui sont petits à l'origine.



FIGURE 1.15 – Technique monde en miniature [59]

- **Technique Voodoo Dolls** : cette technique appelée aussi "Poupées Vaudou" [61] des poupées sont créées pour représenter les objets virtuels existants en miniature. La sélection des objets à manipuler se fait par la technique "headcrusher". Une maquette miniature de l'objet sélectionné et de son environnement proche est créée

[62]. La manipulation peut alors se faire en utilisant la main virtuelle simple (voir Figure. 1.16).

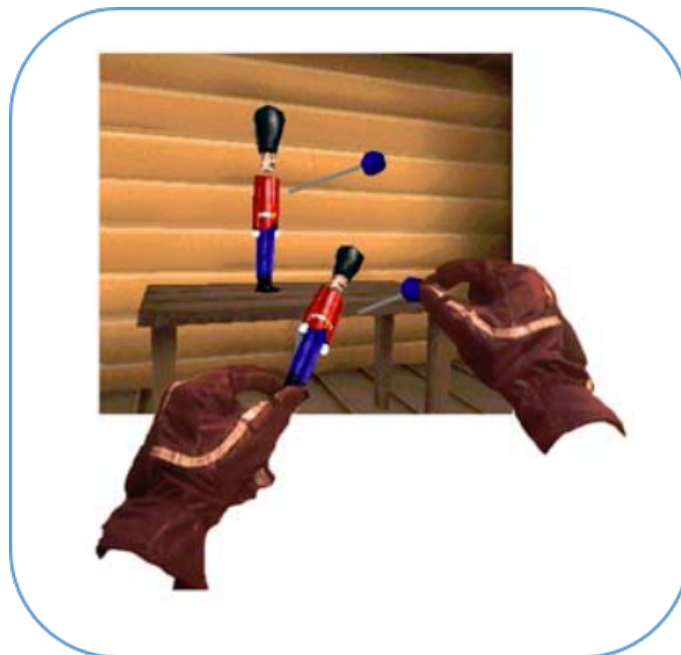


FIGURE 1.16 – Technique Voodoo Dolls [61]

- **La technique Indirect-touch** : proposée par Simeone [63], cette technique utilise des surfaces tactiles pour la sélection et/ou la manipulation, elle utilise des gestes tactiles pour manipuler du contenu 3D sur un dispositif tactile externe tel qu'une tablette ou un smartphone (Voir Figure. 1.17).

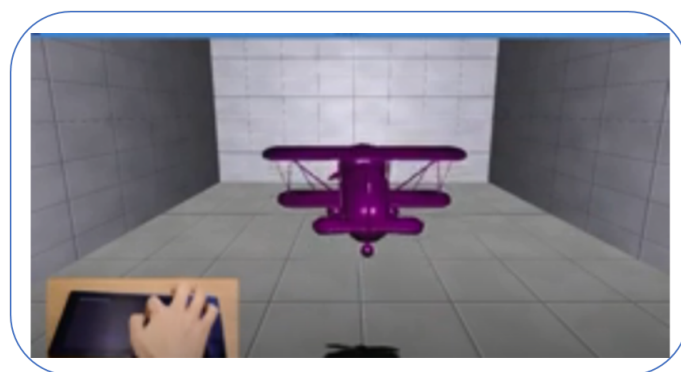


FIGURE 1.17 – Technique Indirect-touch [63]

4. Les techniques par métaphore hybride Ces techniques hybrides combinent au moins deux techniques de sélection et/ou de manipulation.

- **La technique Scaled-world grab** : lors de la sélection d'un objet, l'échelle de l'utilisateur et du monde 3D est augmentée de façon à ce que la main de l'utilisateur touche l'objet sélectionné [64]. Dans cette technique l'utilisateur ne perçoit pas le changement jusqu'au moment de son mouvement (Voir Figure. 1.18). Cette technique a été reprise et citée par plusieurs auteurs [65], [66], [35] et [67].

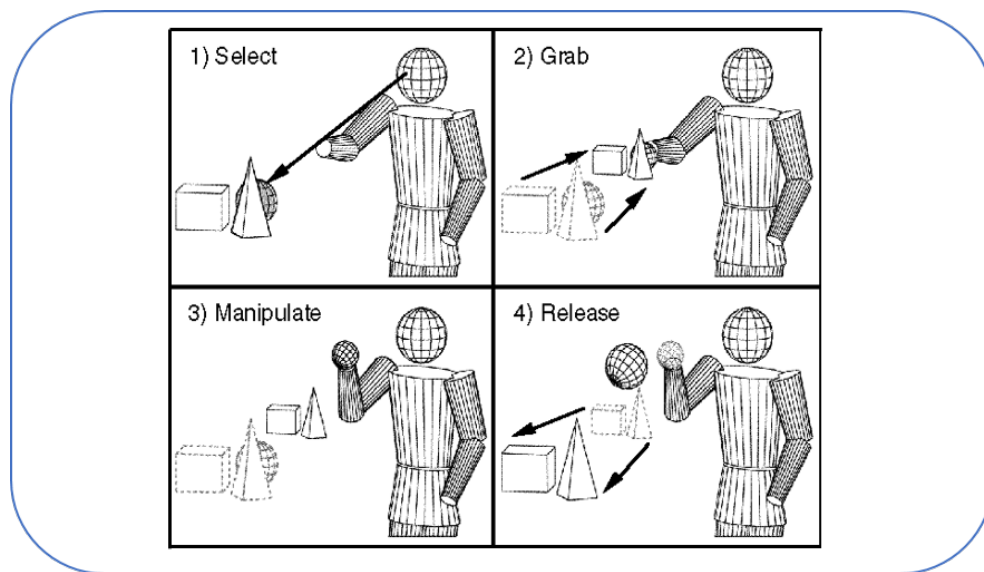


FIGURE 1.18 – Technique Scaled-world grab [64]

- **La technique HOMER (Hand-centered Object Manipulation Extending Ray-casting)** : cette technique combine les deux techniques précédentes le Ray-Casting pour la sélection d'un objet et la technique de la main virtuelle simple pour la manipulation [29].

Un objet sélectionné est attaché à la main et hérite de toutes les transformations que la main subie. La main virtuelle reprend sa position initiale dès que l'objet est relâché (Voir Figure 1.19). Plusieurs auteurs ont jugé que la technique HOMER était

une technique d'interaction très intéressante, elle a permis de résoudre plusieurs problèmes liés à la manipulation d'objets 3D distant. Cette technique a été reprise et citée par plusieurs auteurs [62], [48], [68], [66], [35] et [67].

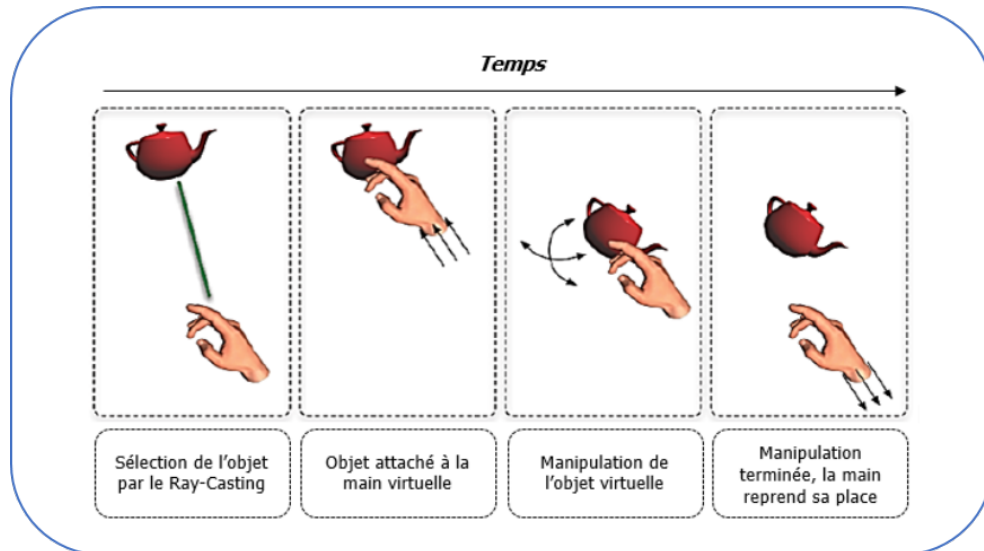


FIGURE 1.19 – Technique HOMER [29]

- **La technique SQUAD** : c'est une technique qui adapte le ray-casting pour projeter une sphère sur la surface en intersection la plus proche du rayon, afin de déterminer quels objets sont susceptibles d'être sélectionnés [69]. Les objets sélectionnables sont ensuite répartis sur quatre quadrants (QUADS) et l'utilisateur affine la sélection jusqu'à ce que l'objet souhaité puisse être sélectionné. Cette technique a montré des performances nettement meilleures pour les objets de petites tailles et les faibles densités d'objets, mais elle présente également une dégradation significative des performances pour les objets de grandes tailles et les densités importantes, en raison du nombre accru d'étapes pour sélectionner un objet (voir Figure 1.20). Cashion et ses associés [70] proposent une variante de SQUAD, appelée Expand, qui ajoute la possibilité de zoomer. Ils montrent qu'Expand est plus rapide que SQUAD pour des densités d'objets élevées.

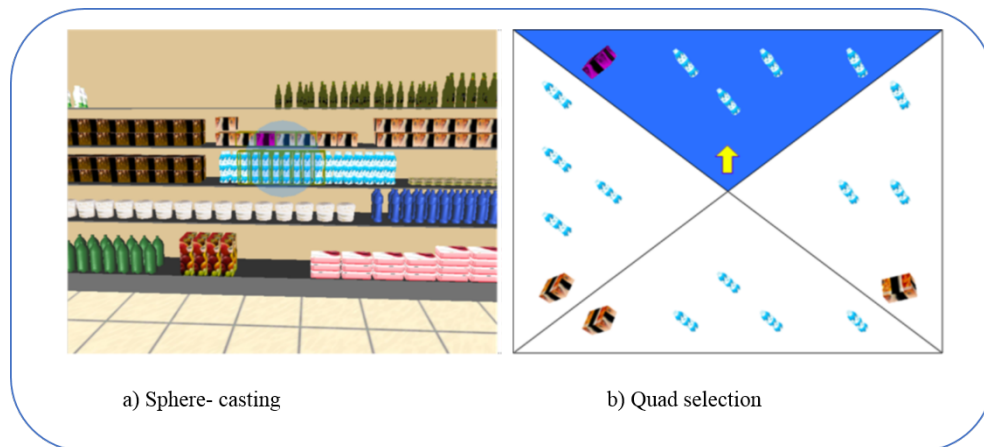


FIGURE 1.20 – Technique SQUAD [69]

c) Les techniques de contrôle d'application

Le contrôle d'application regroupe toutes les techniques de manipulation indirectes sur l'environnement 3D. Cette technique est différente des autres techniques, elle se situe à un niveau conceptuel différent, dans le sens où l'utilisateur agit sur l'application en utilisant les services offerts par l'application elle-même. Plusieurs techniques de contrôle d'application ont été développées. Elles sont regroupées :

1. Contrôleurs physiques : on retrouve dans cette catégorie les contrôleurs physiques tels que les boutons et les commutateurs offrant une solution légère pour le contrôle du système, analogiquement aux touches de fonction des ordinateurs de bureau (voir Figure 1.21).



FIGURE 1.21 – Quelques contrôleurs physiques

2. Menu graphique 3D : c'est l'équivalent 3D des menus 2D. Ils prennent une forme 3D des différents menus existant. Ils peuvent aussi prendre la forme 3D d'un menu 2D affiché dans le monde virtuel (afin de sélectionner et de manipuler on utilisera alors une technique de sélection/manipulation 3D pour manipuler les boutons du menu) (voir Figure 1.22)[71].



FIGURE 1.22 – Menu graphique 3D [71]

3. Commande vocale : une commande a partir d'un ensemble de mots, le lancement de ces commandes initialise la sélection. La reconnaissance vocale bien qu'elle soit une méthode naturelle d'interaction avec le système, elle n'est pas d'une efficacité garantie[35].

Les techniques citées ci-dessous sont des techniques mono-utilisateurs, pour des situations complexes où la tâche est difficile ou même impossible réalisable par une seule personne, l'interaction collaborative est nécessaire.

1.5.2 Interaction 3D collaborative

Dans un environnement virtuel collaboratif (EVC), l'utilisateur interagit seul ou en groupe. Dans le cas où l'utilisateur agit seul sur un objet, il utilise les techniques d'interaction mono-utilisateur citées dans la section 3.3.4.2 dans le but de se déplacer, sélectionner et manipuler dans l'EVC. Néanmoins ses actions peuvent être observées par les autres participants. Cependant toutes les techniques ne sont pas exploitables dans les EVC. En effet, certaines techniques sont difficiles à employer car elles ne permettent pas aux autres utilisateurs de comprendre ce que l'utilisateur est en train de réaliser et ne favorisent pas l'immersion de l'utilisateur. L'interaction multi-utilisateurs doit permettre à plusieurs utilisateurs d'un EVC d'interagir sur un même objet, c'est-à-dire sur les mêmes données. Le second point nécessaire est de permettre aux utilisateurs de visualiser les actions des autres participants. Il est nécessaire de proposer une sur-couche aux techniques mono-utilisateur afin que les utilisateurs puissent agir ensemble, sur les données partagées de l'objet (par exemple, la position et l'orientation).

1.5.2.1 Classification des techniques d'interaction 3D Collaborative

Les techniques d'interaction 3D multi-utilisateurs sont divisés en deux catégories : les techniques de sélection/manipulation asynchrones et les techniques de sélection/manipulation synchrones.

a) Techniques de sélection/manipulation asynchrones

Les méthodes asynchrone ont vu le jour dès les débuts de la recherche dans ce domaine. Le projet PIT (Protein Interactive Theater) [72] a proposé la première approche qui exploite les techniques mono-utilisateurs que plusieurs utilisateurs vont appliquer à un objet à tour de rôle.

Cette dernière est utilisé lorsque les utilisateurs sont physiquement proches ou co-localisés, ces utilisateurs partagerons et interagirons sur un environnement virtuel commun pour faire un meeting vidéo. Chacun des deux utilisateurs possède son propre point de vu de la scène, les positions de l'utilisateur sont données par le "tracking" de la tête par des capteurs positionnés sur les lunettes stéréoscopiques(Figure 1.23.a). Cependant, les deux utilisateurs se partagent un même dispositif d'interaction pour la manipulation des objets avec une main virtuelle simple. Les actions sur l'objet sont réalisées par les utilisateurs, chacun son tour, mais ils peuvent communiquer en direct pour se coordonner pour l'attribution des tâches [73].

Chow et ses coauteurs [74] ont listés plusieurs EVCs qui utilisent l'interaction asynchrone telle-que V-Mail (Imai, 1999), MASSIVE-3 [75], CEDAR [76] (Voir Figure 1.23.b).

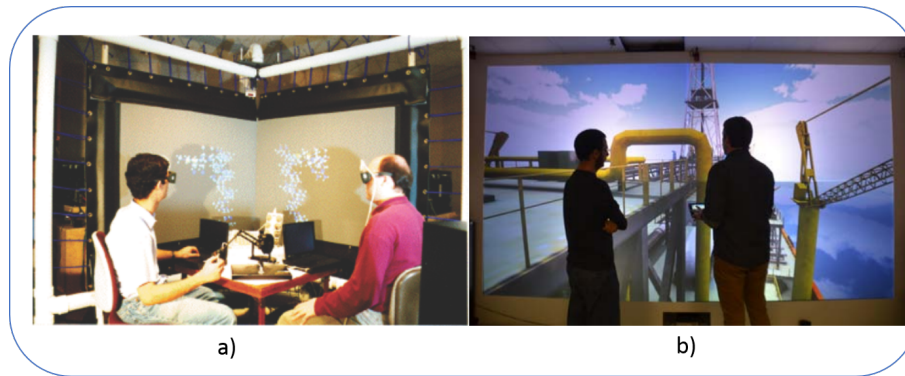


FIGURE 1.23 – a) Un EVC pour la visualisation de molécules [72] b) CEDAR : EVC pour construire et visualiser des pipelines [76]

Dans cette application chaque utilisateur dispose d'un dispositif d'interaction, avec le dispositif de sortie(écran). C'est le principe qu'on retrouve très souvent dans les jeux vidéo multi-joueurs où l'écran est divisé en deux ou plusieurs parties(voir Figure 1.24 : Extrait du jeu GranTurismo7²). Les utilisateurs agissent uniquement sur leurs données mais peuvent visualiser les données des autres participants [74].



FIGURE 1.24 – La course se réalise avec deux participants. L'écran est partagé, les utilisateurs peuvent communiquer en direct mais agissent uniquement sur leur véhicule, ils ne peuvent pas agir sur le véhicule adverse.

2. GranTurismo7 <https://fr.wikipedia.org/wiki/GranTurismo7>

Ce mode d'interaction ne permet pas une interaction simultanée (manipulation simultanée) ou un accès concurrent d'un objet par plusieurs utilisateurs [32]. Cependant, certaines techniques d'interaction peuvent être modifiées pour permettre l'interaction à plusieurs.

b) Techniques de sélection/manipulation synchrones

Pinho et ses coauteurs [77],[78] ont défini les techniques synchrones. Ces dernières permettent aux utilisateurs d'agir ensemble et en même temps sur des données partagées (un ou plusieurs objets). On distingue deux catégories : une première catégorie consiste à répartir des données partagées (degrés de liberté). On donnera l'exemple d'un utilisateur qui va agir sur les positions et l'autre sur les rotations de l'objet via des techniques d'interaction mono-utilisateurs. La seconde catégorie concerne "l'accès concurrent aux mêmes données". Cela permet de combiner les mouvements (position et orientation) des différents utilisateurs pour déterminer le mouvement final de l'objet. On retrouve dans cette catégorie des techniques qui exploitent des techniques mono-utilisateurs tels que : ray-casting ou main virtuelle simple [32].

Répartition des données partagées entre les utilisateurs (Degrés de liberté différents)

Ruddle et ses coauteurs [79] ont introduit cette première approche, qui permet de réaliser une interaction simultanée sur l'objet partagé en séparant les propriétés de ce dernier (Degrés de libertés) comme les rotations et positions afin de les assigner aux différents utilisateurs. On prendra l'exemple où un utilisateur réalisera les translations et l'autre effectuera les rotations. Dans cette catégorie, les utilisateurs peuvent agir en même temps sur le même l'objet (Voir Figure 1.25). Ce principe a été repris par plusieurs travaux notamment [80], [78], [32] et [73].

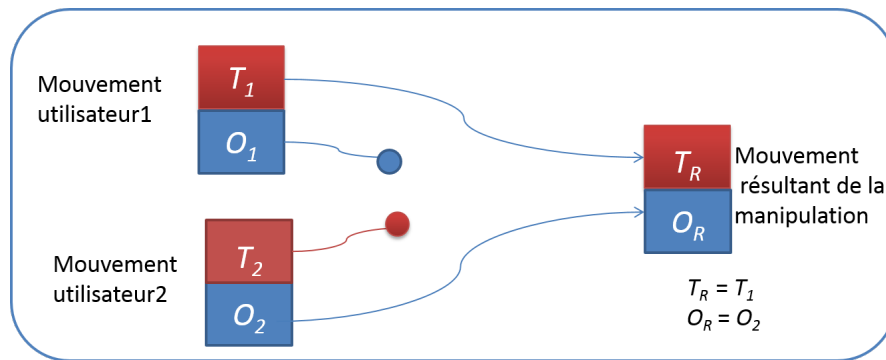


FIGURE 1.25 – Répartition des données partagées entre les utilisateurs [80])

Pinho et ses associés[78]ont classé cette approche en deux sous-familles, la première étant par métaphore coopérative homogène et la seconde par métaphores hétérogènes.

Pour le premier cas, les utilisateurs utilisent une même métaphore mono-utilisateur, par exemple une main virtuelle simple, Ray-casting ou autre [78] et [32]. Par ailleurs, la séparation des degrés de liberté implique que si l’un des utilisateurs peut manipuler l’orientation d’un objet, l’autre ne peut changer que sa translation, couleur ou autre (voir Figure 1.26.a).

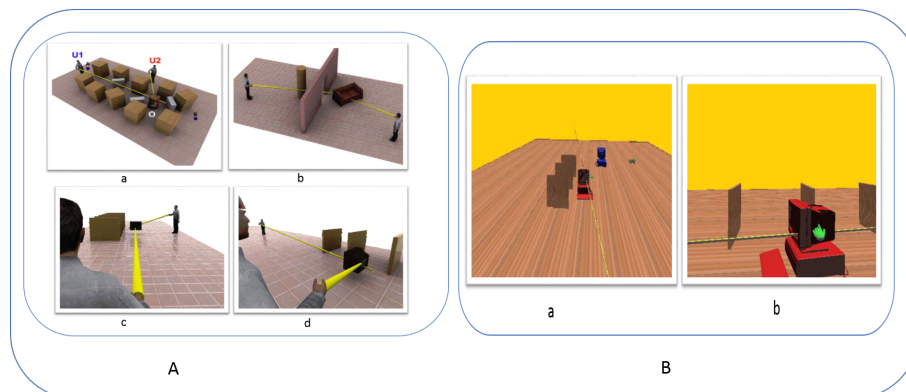


FIGURE 1.26 – A : Répartition de données partagées par métaphores homogène(a : aperçue générale de la co-manipulation, b : scénario de la co-manipulation, c,d : vues des différents utilisateurs [78]. B : Répartition de données partagées par métaphores hétérogène(a : vue du premier utilisateur, b : vue du deuxième utilisateur [77].

Pour les métaphores coopératives hétérogènes, les utilisateurs utilisent des métaphores mono-utilisateur différentes. Par exemple, un premier utilisateur applique des translations en utilisant une technique tel que Ray-casting et le deuxième utilisera une autre telle que main virtuelle simple ou autres (Voir Figure 1.26.b) [32] et [73]. Ces techniques s'avèrent très adaptées pour faire des ajustements et des manipulations difficiles (par exemple des manœuvres dans des espaces étroits) où elles conduisent à une réalisation plus rapide de tâches.

Accès concurrent aux données partagées

Une technique qui dérive de la technique Ray-casting pour la manipulation d'un objet partagé à deux utilisateurs a été proposée par Duval et Le Tenier [81]. Cette dernière se base sur le principe suivant : si un utilisateur manipule son rayon afin de déplacer un objet partagé, le rayon de l'autre utilisateur "agrippé" à ce même objet va aussi se déplacer. Cependant, le rayon du second utilisateur n'a pas bougé ce qui engendre une incohérence entre le déplacement réel et le déplacement virtuel. Duval et ses associés proposent de courber les rayons en fonction de la force qui est appliquée [73].

Le rayon peut prendre trois formes :

- *Elastique* : si une contrainte apparaît lors de la manipulation de l'objet par l'utilisateur, un élastique va s'étirer entre l'objet et le rayon (voir Figure 1.27.a).
- *Le rayon coudé* : le rayon va former un angle de manière à former une sorte de coude afin d'atteindre l'objet en question. Cette technique est utilisée par les deux utilisateurs (Voir Figure 1.27.b).
- *Déformable* : le rayon va se dédoubler, une partie correspondant au rayon initial ayant permis la sélection va rester stable [57]. La seconde partie va se plier pour

rester en contact avec l'objet en fonction des mouvements de l'objet suite à la collaboration(voir Figure 1.27.c).

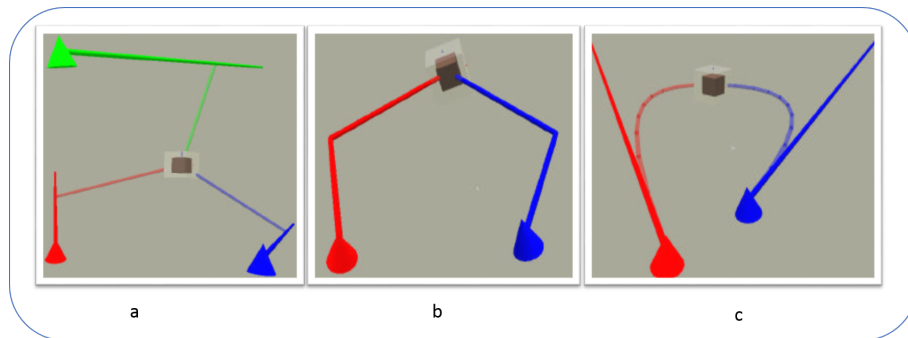


FIGURE 1.27 – a : Élastique(les rayons interagissent avec le cube). b : Coudé (les rayons interagissent avec le cube). c : Déformable (les rayons interagissent avec le cube) [82].

Ruddle et ses coauteurs[79] ont introduit une approche qui permet un accès simultané et concurrent aux mêmes données (degrés de liberté), Ruddle a proposé d'utiliser la moyenne des degrés de liberté de chaque utilisateur pour avoir la position finale de l'objet partagé (moyenne des translations et des orientations Figure 1.28).

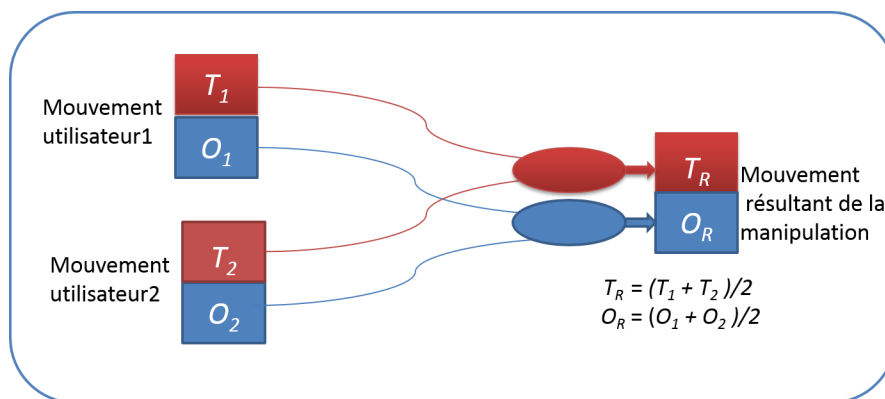


FIGURE 1.28 – Accès concurrent aux données partagées ([79])

En se basant sur ce principe Duval et ses coauteurs [80] proposent la technique SkeweR dont le déplacement s'effectue par des curseurs 3D. Un utilisateur approche son curseur

près d'un point de contrôle pour manipuler l'objet. S'il n'y a qu'un utilisateur, seuls des mouvements de rotations autour du centre de gravité de l'objet peuvent lui être communiqués (Voir Figure 1.29). S'il y a plusieurs utilisateurs, l'objet se manipule par une fonction de translation et de l'orientation.

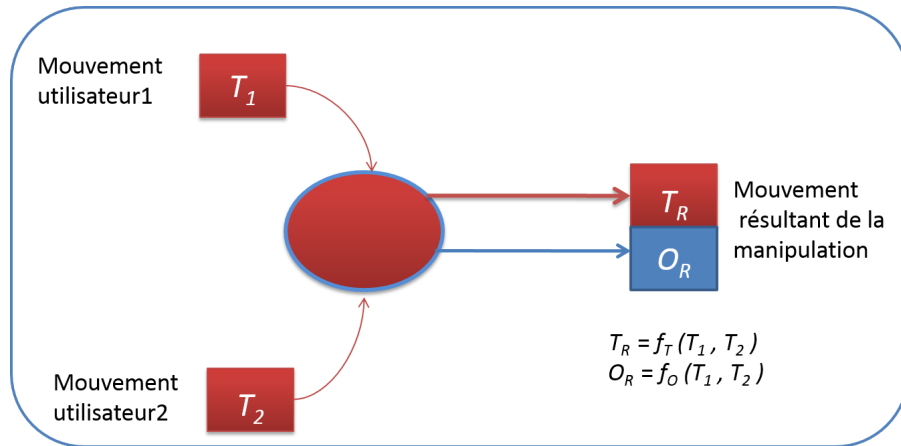


FIGURE 1.29 – Manipulation collaborative de deux utilisateurs en utilisant la technique SkeweR [80]

Tous ces mécanismes et approches ont été utilisés par plusieurs auteurs afin de mettre en place leur techniques d'interactions mono-utilisateurs dans des environnements virtuels collaboratifs de manière synchrone et/ou asynchrone.

Nous citerons Soares [83], qui a proposé et évalué la technique EGO-EXO, technique qui assure une co-manipulation synchrone (manipulation collaborative) par deux utilisateurs avec des points de vues différents.

Grandi dans [84] a proposé et évalué une interface 3D basée sur un smartphone pour la co-manipulation. Wieland dans [85] a présenté une comparaison de plusieurs techniques d'interaction 3D collaboratives : synchrones, asynchrones et même hybrides en réalité augmentée.

1.6 Bilan et discussion

Nous avons présenté dans la section 1.5 les techniques d'interaction 3D mono-utilisateurs et multi-utilisateurs ainsi que leurs classifications. Dans ce qui suit, nous allons nous concentrer sur les techniques de sélection et de manipulation. Ces techniques présentent chacune des avantages et des inconvénients qui ont été recensés lors de l'étude de l'état de l'art.

La plupart des techniques d'interaction 3D mono-utilisateurs existantes traitent l'interaction d'une manière générale et ne répondent pas aux exigences de certaines applications telles que la précision dans l'interaction que ça soit dans un environnement simple ou complexe. La complexité de l'environnement varie selon les objets le constituant (dense ou non, petits ou grands, loin ou proches, occultés ou visibles, etc).

Pour répondre à toutes ces contraintes, il est important d'avoir une interaction 3D avec un certain nombre de contraintes :

- Précision dans la sélection et/ou manipulation des objets ;
- Sélectionner et/ou manipuler les objets petits, lointains et occultés ;
- Avoir une interaction 3D rapide.

Nous avons effectué une étude comparative des techniques de sélection et de manipulation mono-utilisateurs.

La table ci-dessous présente une classification des techniques d'interaction dédiées à la sélection et à la manipulation (Voir Table 1.2).

Nous avons classifié ces techniques selon leurs précision et leur capacité de sélectionner des objets difficiles à atteindre (loin et/ou occultés).

TABLE 1.1 – Grille des échelles

Précision	Objets difficiles à atteindre
+ : Moins précise	+ : Sélection des objets difficiles à atteindre moins possible
+++++ : Très précise	+++++ : Sélection des objets difficiles à atteindre très possible

Nous avons attribué une note pour chaque critère en utilisant une échelle de 1 à 5 (de + à +++++) expliqué dans la table 1.1.

A l'issue de cette étude, nous avons relevé plusieurs challenges liés à l'interaction 3D mono-utilisateur dans le monde virtuel. Le premier challenge étant la précision. Le niveau de précision de chaque technique d'interaction 3D diffère d'une à l'autre. Nous avons remarqué que les techniques les plus précises sont les techniques qui utilisent directement ou indirectement la métaphore de la main virtuelle simple.

Le deuxième challenge consiste à sélectionner et/ou manipuler des objets difficiles à atteindre tels-que les objets petits, lointains et occultés.

Certains travaux présentés dans la section 1.5.2 s'intéressent aussi à l'interaction 3D collaborative, ce qui permet à plusieurs utilisateurs d'agir ensemble. La complexité de ces environnements est liée à l'interaction du groupe d'utilisateurs avec les éléments partagés.

Dans un monde partagé, plusieurs contraintes apparaissent, notamment celles liées à la coordination, la communication et l'interaction collaborative. L'interaction 3D collaborative est basée essentiellement sur des techniques d'interactions 3D mono-utilisateurs coordonnées entre elles via des protocoles de collaborations. Le challenge majeur dans les Environnements virtuels collaboratifs est de fournir des modèles et des formalismes de collaboration permettant de gérer et de contrôler les actions des utilisateurs dans l'environnement virtuel.

Nous avons constaté durant ce chapitre que l'interaction 3D ne manque pas d'intérêt. Les recherches dans ce domaine sont en pleine effervescence, certains problèmes et challenges

TABLE 1.2 – Classification des techniques de sélection et de manipulation

Techniques d'interaction 3D	Outils de sélection	Références	Précision	Objets difficiles à atteindre	Le temps d'exécution
Main virtuelle simple	Main	[42]	++++	+	++++
Go-Go	Main	[28]	+++	++	++
Fast Go-Go	Main	[29]	+++	++	+++
Stretch Go-Go	Main	[45]	+++	++	++
PRISM	Main	[48]	+	+++	++
3D Bubble Cursor	Main	[49]	+	+++	+
Ray-casting	Rayon	[50]	++	++	++++
Image plane	Rayon	[52]	+	++	+
Ray-flexible	Rayon	[54]	++	+++	+++
Flashlight	Rayon	[55]	+++	++	+++
Aperture	Rayon	[56]	+++	++	++
Eye-casting	Rayon	[57]	+	+++	++++
Depth-ray	Rayon	[49]	++	+++	+++
Le monde en miniature	Indirecte	[59]	+	+	+
Voodoo Dools	Indirecte	[61]	+++	+	+
Indirect touch	Indirecte	[63]	++	+	+
Scaled world grap	Hybride	[64]	+	+++	++
HOMER	Hybride	[29]	++++	++	+++
SQUAD	Hybride	[69]	+	+	++
Expand	Hybride	[70]	++	+	+

restent toujours d'actualité tels-que la précision de l'interaction notamment avec les objets distants, occultés.

Dans cette optique, nous proposons une technique d'interaction 3D notamment une technique de sélection qui va permettre de reprendre les précédents challenges et d'améliorer l'interaction 3D. Dans le monde partagé nous envisageons formaliser l'interaction 3D collaborative pour mieux coordonner les actions de groupe d'utilisateurs.

1.7 Conclusion

Tout au long de ce chapitre, nous avons abordé l'interaction 3D en réalité virtuelle, en présentant dans un premier temps les concepts fondamentaux et les définitions de base qui régissent cette technologie. Nous avons également présenté les environnements virtuels collaboratifs(EVCs). Nous avons abordé dans une première partie les différentes tâches de l'interaction mono-utilisateur ainsi que leur classification. Dans une seconde partie les techniques d'interaction 3D multi-utilisateurs ainsi que leur classification ont été présentés.

A partir de cet état de l'art, nous avons constaté que la majorité des techniques d'interaction 3D mono-utilisateurs existantes ne sont pas précises et ne permettent pas à l'utilisateur d'interagir avec des objets occultés. Ces techniques mono-utilisateurs ont besoin d'être coordonnées entre elles pour offrir des techniques d'interaction 3D multi-utilisateurs.

Nous présentons dans ce qui suit les différentes techniques et technologies de reconnaissance de gestes de la main. Ceci nous permettra de voir les possibilités en matière d'équipements et de techniques de reconnaissance de gestes pouvant nous être utiles pour la suite de notre travail.

Chapitre

2

Reconnaissance de gestes : technologies et techniques

2.1 Introduction

La notion d'interaction en RV se réfère aux méthodes, règles et techniques qui permettent à un utilisateur ou à plusieurs d'effectuer des tâches d'interaction au sein d'un environnement virtuel ou augmenté [18]. L'interaction 3D est donc, dite le langage de communication entre l'utilisateur et le monde virtuel.

Le développement technologique ouvre de nouvelles portes à la RV. Avec le concept d'immersion mobile en réalité virtuelle, la démarche d'aujourd'hui consiste à permettre à l'utilisateur en situation de mobilité d'interagir directement avec les objets virtuels. L'objectif étant d'offrir un moyen d'interaction à la fois naturel et intuitif, et cela en utilisant un des canaux de communication du corps humain (voix, touché, geste, etc.).

Les travaux dans ce domaine sont nombreux, notamment ceux qui s'intéressent à l'interaction naturelle et cela en présence d'un ensemble de techniques de reconnaissance de gestes de la main. Dans ce chapitre nous allons présenter les interfaces naturelles et les différentes techniques de reconnaissances de gestes de la main.

Nous abordons également dans ce chapitre, les différents outils et technologies existantes permettant la reconnaissance des gestes de la (les) main(s). L'étude menée dans ce chapitre,

sur les différentes techniques et technologies de reconnaissance de gestes de la main, nous a permis d'énumérer les différentes possibilités en matières d'équipements et de techniques de reconnaissance de gestes pouvant nous être utiles pour la suite de notre travail. Un bilan détaillé de cette première partie de la thèse sera dressé afin de nous permettre de dégager nos contributions relatives à la problématique définie dans le cadre de cette dernière.

2.2 Les interfaces utilisateur naturelles

De nombreuses interfaces dites "naturelles" (Natural User Interfaces ou NUI) on vu le jour depuis la dernière décennie, que ce soit dans des dispositifs de réalité virtuelle, de jeu vidéo ou en Interaction Homme Machine (IHM).

On peut effectuer des tâches primitives à priori différentes, par exemple la sélection ou l'édition pour une IHM, et la manipulation en réalité virtuelle [86]; [87]. Comme le note Lee [88] ces deux domaines se concentrent sur l'activité; le principe d'une interface étant de permettre à l'utilisateur de concrétiser une idée, tout en minimisant l'effort pour y parvenir. Il existe ainsi des points de convergence entre les modalités d'interaction de ces deux domaines qui tendent à exploiter des interfaces plus naturelles. Elles utilisent le corps humain comme interface et font appel aux compétences que l'humain a développé depuis sa naissance, issues de son interaction permanente avec son environnement [89]. Contrairement à une interface utilisateur non-naturelle, elle demandera l'usage d'un outil spécifique, requérant un apprentissage long et non-intuitif. Typiquement, on trouvera dans cette catégorie des interfaces complexes et spécialisées pour un métier et une tâche donnés. Le fait d'expérimenter la réalité façonne les connaissances de l'homme et lui permet par la suite d'interagir avec son environnement naturellement en utilisant ces mêmes connaissances [90].

Les interfaces utilisateur naturelles imitent les interactions que les humains ont avec leur environnement réel. Elles produisent dès lors des interactions pseudo-naturelles. Les interfaces utilisateur s'appuient sur le système sensori-moteur de l'homme, et concernent la vue, l'haptique, la parole, le mouvement, la proprioception, mais également le goût ou encore le toucher.

Plusieurs types d'interfaces naturelles existent, chacune utilise un canal sensori-moteur distinct. L'interaction avec des systèmes combinant plusieurs interfaces utilisateur naturelles est très complexes à réaliser, semble dès lors être un moyen efficace d'améliorer la naturalité d'une interaction.

2.3 Définition et représentation du geste

Avant de comprendre la nature ou l'objectif du geste, rappelons qu'il est basé sur un mouvement du corps. Baudel et ses coauteurs [91] montrent que la gestuelle a l'avantage d'être une forme naturelle de communication et d'être rapide à comprendre. Dans le traité de la réalité virtuelle Fuchs [87] a remarqué une particularité du geste qui le différencie des autres modalités sensorimotrices :

sa fonction motrice et sa fonction sensorielle sont indissociables. De plus, pour ajouter un peu de complexité, il existe une influence réciproque geste et action. Par exemple, Goldin[92] a noté que dans la communication orale, le geste reflète les pensées, y compris le non-verbal, mais il peut réciproquement influencer la pensée de celui qui parle. Dans ce cas, le geste se révèle être à la fois un moyen de communication et un "outil pour la pensée" du locuteur.

Cadoz et Claude [93] ont défini trois fonctions principales des gestes de la main :

- la fonction ergotique : la main joue le rôle d'organe moteur et agit sur le monde physique pour le transformer. Elle applique aux objets des forces pour les déplacer ou les déformer.
- la fonction épistémique : la fonction épistémique : la main joue le rôle d'organe de perception. Le sens du toucher (sens tactilo-proprio kinesthésique) donne des informations sur la forme, l'orientation, la distance, la grandeur, le poids, la température, les mouvements des objets, etc.
- la fonction sémiotique : la main joue le rôle d'organe d'expression pour l'émission d'informations visuelles. Cela comprend la langue des signes, le geste Co-verbal, qui accompagne la parole, ou les gestes permettant une communication basique lorsqu'on ne peut pas utiliser la parole, comme dans un environnement bruité ou en plongée sous-marine.

Dans cette thèse, nous nous intéressons plus particulièrement à la fonction sémiotique. Cette fonction est la plus riche et la plus complexe.

Afin d'étudier les systèmes de reconnaissance de gestes, nous commençons par définir le geste et le classifier. Nous poursuivons avec les différentes techniques de reconnaissance de gestes, ainsi que les technologies existantes permettant l'acquisition de geste.

2.4 Techniques d'acquisition et de reconnaissance de gestes

Plusieurs études ont été menées sur la reconnaissance des gestes et plusieurs classifications y ont été proposées selon le type de la technologie d'acquisition utilisée. Deux approches

ont été présenté dans la classification de Laviola [94] : les approches basées capteurs et les approches basées vision.

2.4.1 Approches basées capteurs

Ces approches se basent sur l'utilisation de certains outils tels-que les accéléromètres, l'écran tactile, les gants...etc. Quelques dispositifs, comme l'iPhone (Apple), incluent plusieurs détecteurs : par exemple, un écran tactile et un accéléromètre. D'autres dispositifs utilisent un seul détecteur : par exemple, les accéléromètres de Nintendo (Wii-mote). Par conséquent, nous pouvons classer ces dispositifs en plusieurs sous-catégories.

2.4.1.1 Capteurs mécaniques

Les capteurs mécaniques munissent la main de l'utilisateur des gants instrumentés comme les gants cybernétiques tel que "CyberGlove II" (voir Figure 2.1) développé par Immersion. Ces gants sont dotés de traqueurs qui sont en général des capteurs supplémentaires (de type magnétique ou acoustique) attachés au dos de la main ou au-dessus du poignet et qui fournissent des données sur la position et l'orientation de la main ainsi que les angles de jointure [95]. Cependant les gants instrumentés nécessitent des câblages et restent très encombrants et cher pour l'utilisateur [95].



FIGURE 2.1 – Les gants « CyberGlove II » [95]

Un costume de corps appelé "IGS-190" a été proposé par Animazoo pour capturer des gestes du corps (voir Figure 2.2), ce dispositif peut être utilisé en collaboration avec d'autres dispositifs. Kevin et ses coauteurs [95] proposent une méthode pour la modélisation des trajectoires dans la reconnaissance de gestes qui utilise des CyberGloves avec des marqueurs magnétiques. De même, le costume de corps "IGS-190" est combiné avec dix-huit dispositifs inertiels (gyroscopes) qui permettent la détection de mouvement.



FIGURE 2.2 – L'Animazoo "IGS-190" contient 18 capteurs pour la reconnaissance de geste en plus des marqueurs visuels [95]

2.4.1.2 Capteurs magnétiques

Les capteurs magnétiques sont certainement les capteurs les plus utilisés en réalité virtuelle, ils utilisent le champ magnétique à basse fréquence émis par un émetteur pour un récepteur afin de déterminer sa position et son orientation par rapport à la source magnétique. Ils ont pour avantages d'être petits, légers et portables de n pas être sensibles aux interférences acoustiques, et d'avoir un taux de rafraichissement élevé et une faible latence. Le principaux inconvénients se situent d'abord au niveau des interférences magnétiques : des objets métalliques peuvent perturber les mesures des récepteurs. Un autres problèmes est lié a la technologie puisque le champs magnétique décroît rapidement avec la distance. Le volume de travail des capteurs magnétiques reste assez limité [96].

2.4.1.3 Capteurs acoustiques

Les capteurs acoustiques convertissent les ondes sonores (de fréquence supérieur à 20 kHz) en signal électrique. Ces capteurs sont peu coûteux, légers et petits. Ils ne sont pas sujets aux interférences magnétiques, en outre ils sont assez précis. Par contre ils souffrent d'un taux de rafraichissement assez faible et d'interférences aux ondes acoustiques (notamment aux échos et aux bruits parasites). L'inconvénient le plus important vient du faite qu'ils sont sensibles à la visibilité : il est préférable qu'il n'y ait pas d'obstacles entre récepteurs et émetteurs [26].

2.4.1.4 Capteurs Inertiels

Les capteurs inertiels reposent sur l'utilisation d'accéléromètres et de gyroscopes. L'orientation des objets est calculée en intégrant conjointement les valeurs du gyroscope, ces dernières étant proportionnelles à la vitesse angulaire autour de chaque axe. De la même manière,

en intégrant doublement les valeurs des accéléromètres, le système retourne les variations de position. Récemment, ce type de capteurs est intégré dans la plupart des smartphones, tablettes et autres consoles et dispositifs. Katzakis et ses coauteurs [97] ont utilisé un smartphone doté d'un capteur inertiel pour manipuler des objets 3D dans une plateforme de RV (voir Figure 2.3).

La portée de ces capteurs est infinie. Ils sont rapides, ils ne souffrent d'aucun problèmes d'occlusion par un autre objets ni d'interférences magnétiques, restent abordables et occupent un espace très réduit. Cependant, ils ne sont pas précis lorsque les changements de positions sont lents, et restent sensibles à des phénomènes de dérives, puisque les erreurs s'accumulent en permanence.

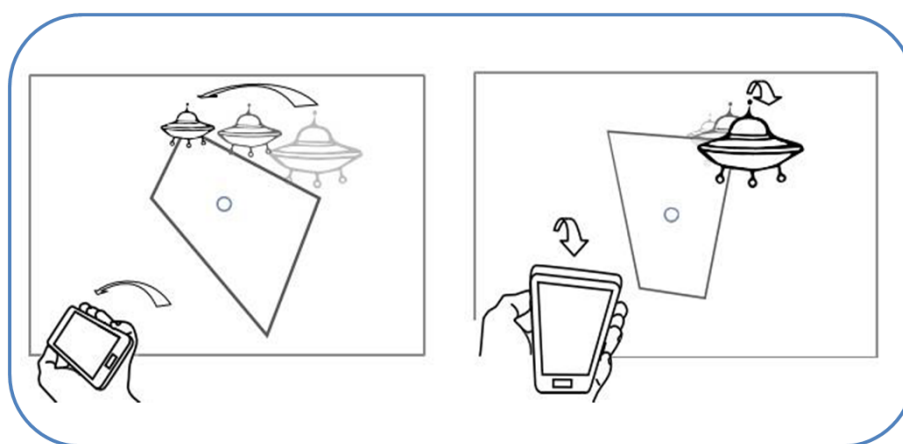


FIGURE 2.3 – Manipulation d'un objet 3D sur un écran en utilisant le capteur inertiel d'un smartphone [97]

2.4.1.5 Dispositifs tactiles

Cette technologie permet d'interagir physiquement avec un objet virtuel. Ce système restitue à l'utilisateur la perception du toucher, dans le cas d'un écran tactile, et la sensation de déplacement 3D dans l'espace en utilisant un dispositif à retour de force. "3DTouch" est une technologie présente au niveau "iPhone 6S" de Apple, cette dernière

permet d'effectuer une interaction 3D sur un écran 2D (voir Figure 2.4) et cela en exerçant un degré de pression par les doigts a travers un écran tactile.

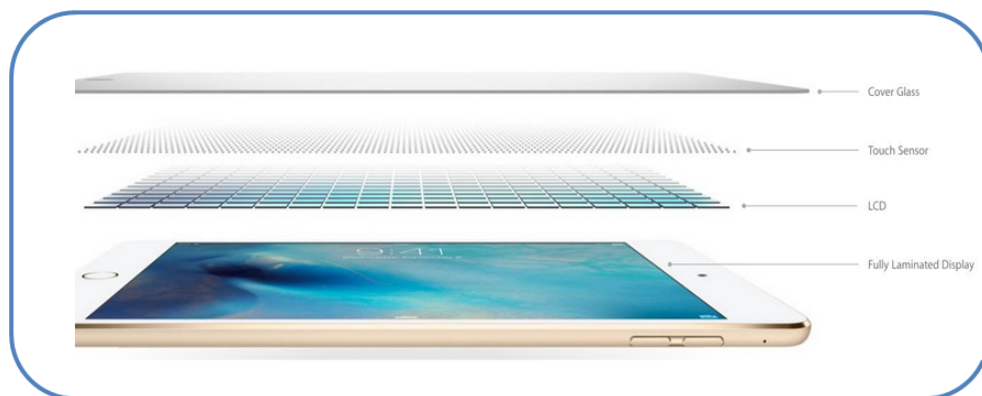


FIGURE 2.4 – Le principe de la technologie 3DTouch d'Apple

2.4.1.6 Autres types de capteurs

Récemment, un nouveau capteur appelé "Soli" [98] pour la détection des gestes de la main à été présenté par l'équipe de recherche ATAP de Google. Il s'agit d'une puce radar miniature de 9 mm de large, qui fonctionne dans le spectre radio à 60 GHz et qui envoie des ondes radio. Lorsque celles-ci rebondissent sur la main et les doigts, le radar peut en déduire leur position dans l'espace à l'échelle du millimètre, avec une fréquence équivalente à 10 000 images par seconde(voir Figure 2.5).

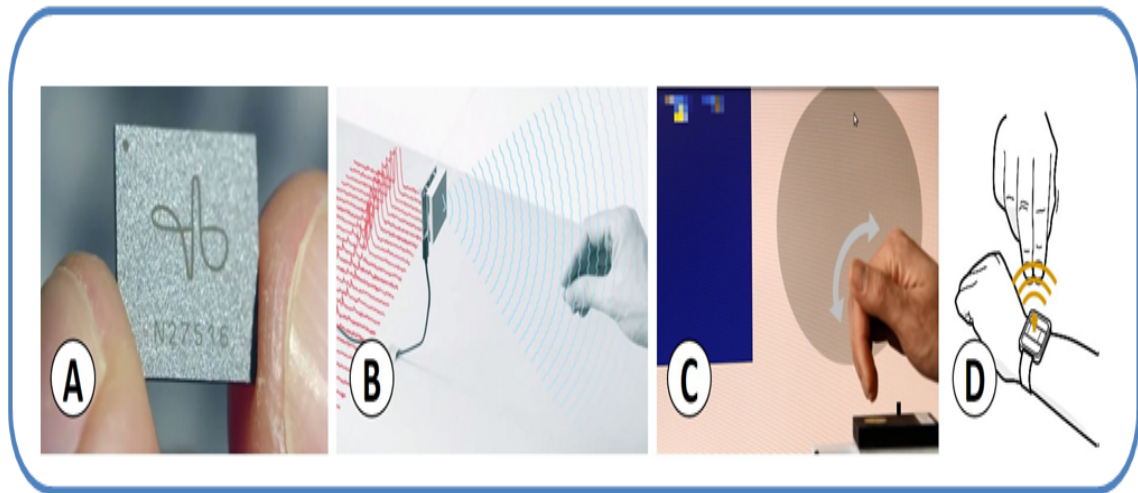


FIGURE 2.5 – A) radar solide à courte portée, captant l'énergie réfléchiée par les mains, B) le signal résout le mouvement dans la gamme des millimètres, C) la reconnaissance de gestes spécialement conçue pour reconnaître des gestes subtils et peu exigeants en se basant sur le signal Soli[98].

L'objectif du projet Soli est de permettre une interaction naturelle et intuitive avec différents objets connectés comme les smartwatches et les smartphones (voir Figure 2.6) en utilisant des mouvements de doigts ou de mains, et sans avoir à toucher l'appareil [99].

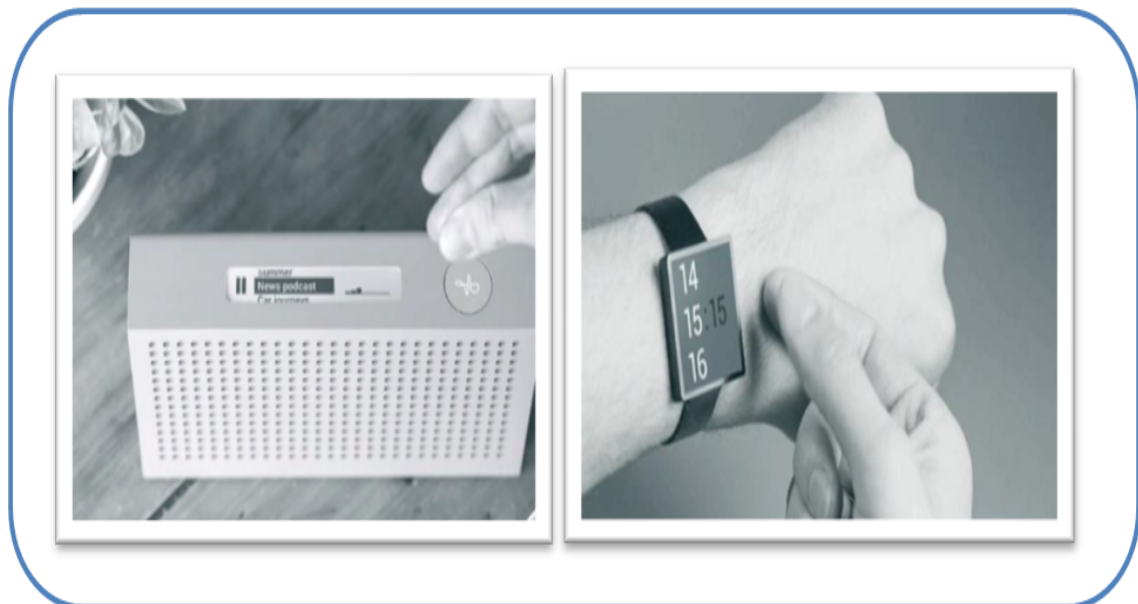


FIGURE 2.6 – Interaction avec des appareils intelligents par le capteur Soli [99]

Dans le même contexte, un prototype d'écran sans contact appelé "Airtouch" a été proposé par Du et associés dans [100]. Ce dernier se base sur des capteurs capacitifs (voir Figure 2.7) afin de déterminer la position du doigt en voisinage de l'écran. Ce qui permet à l'utilisateur d'interagir en 3D avec les objets virtuels. Cependant, le mouvement des doigts en profondeur est très limité.

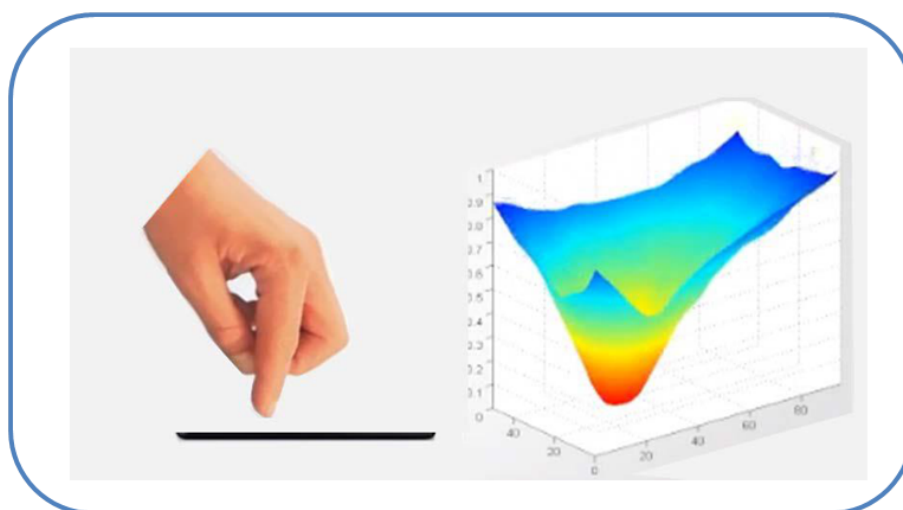


FIGURE 2.7 – concept de Airtouch montrant l'espace de détection de la main [100]

2.4.2 Approches basées vision

Les approches basées vision utilisent le flux vidéo capturé par des caméras pour reconnaître les gestes. Selon les différents points de vue du geste, ces approches constituent les différentes apparences du geste [101]. Plusieurs représentations ont été proposées afin de reconnaître et modéliser un geste de la main.

Kaaniche [102] divise les méthodes de reconnaissance de gestes en deux grandes catégories : les méthodes basées sur des modèles 3D et les méthodes basées sur l'apparence, comme illustré dans la Figure 2.8.

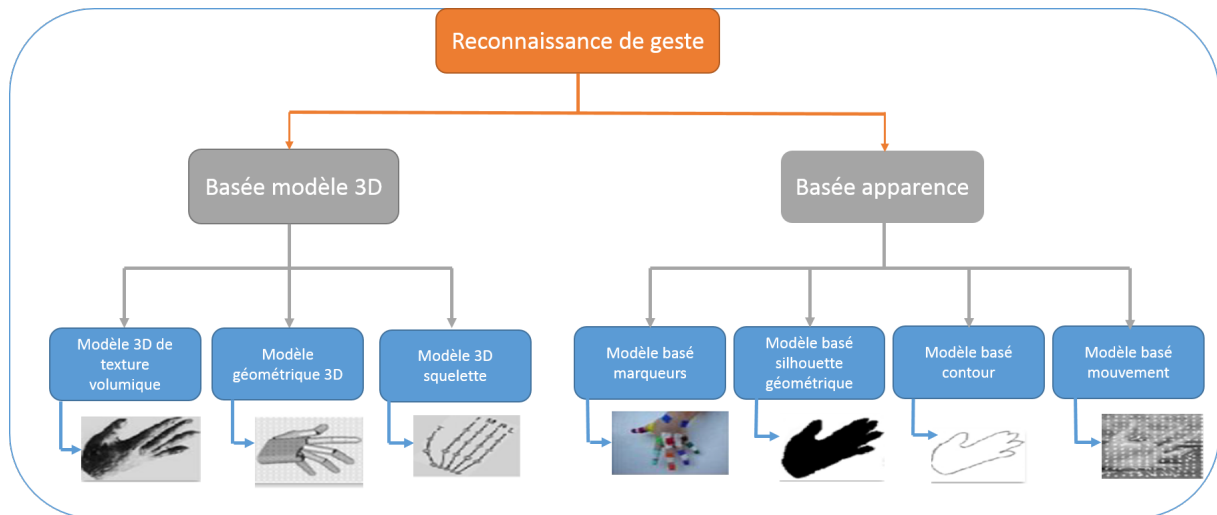


FIGURE 2.8 – Représentation des gestes de la main, adaptée de [102]

2.4.2.1 Techniques basées apparence

Dans cette catégorie on retrouve deux sous familles, les méthodes basées modèles 2D et des méthodes basées mouvements.

Plusieurs approches ont été proposées, la première étant basée sur des gabarits déformables qui sont généralement basés sur des contours actifs déformables ("Snakes" paramétré et ses variants) [103]. Ju et ses coauteurs [104] ont utilisé la théorie des contours actifs paramétriques "Snakes" pour l'analyse de gestes et d'actions de la main dans une séquence vidéo (voir Figure 2.9).

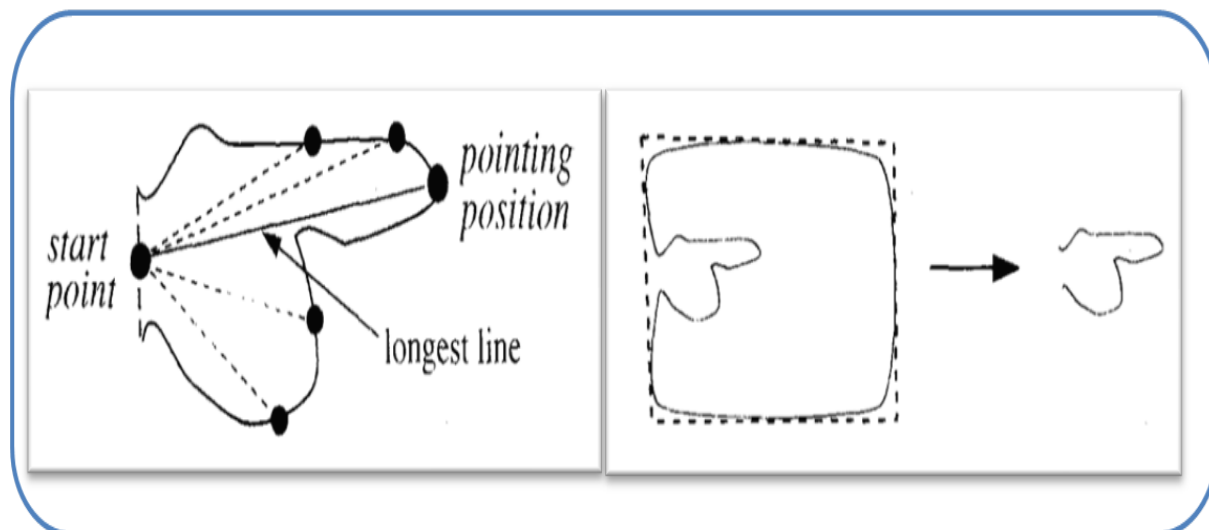


FIGURE 2.9 – Technique basée sur les contours actifs [104]

les travaux suivants seront basés sur des approches basées couleurs, ces approches utilisent des marqueurs (sur le corps, mains ou les doigts) faciles à reconnaître (gants, couleurs, cibles codées..etc) afin d'interagir avec le système.

La couleur de la peau est une caractéristique importante qu'on peut extraire d'une image afin de détecter et suivre les mains d'une personne. Bretzner et associé [105] ont proposé un algorithme de suivi et de reconnaissance des gestes de la main en utilisant des caractéristiques de couleur multi-échelles et des modèles hiérarchiques (voir Figure 2.10). Les postures de la main sont détectées et représentées en utilisant les caractéristiques d'images en couleur multi-échelle. Cependant, l'inconvénient majeur de ces méthodes c'est qu'elles ont du mal à différencier la main des autres objets dont la couleur est similaire à celle de la main, comme le visage ou bien le bras. Une solution a été proposée pour résoudre ce problème, faire porter aux utilisateurs des chemises blanches à manches longues ce qui devient intrusif.

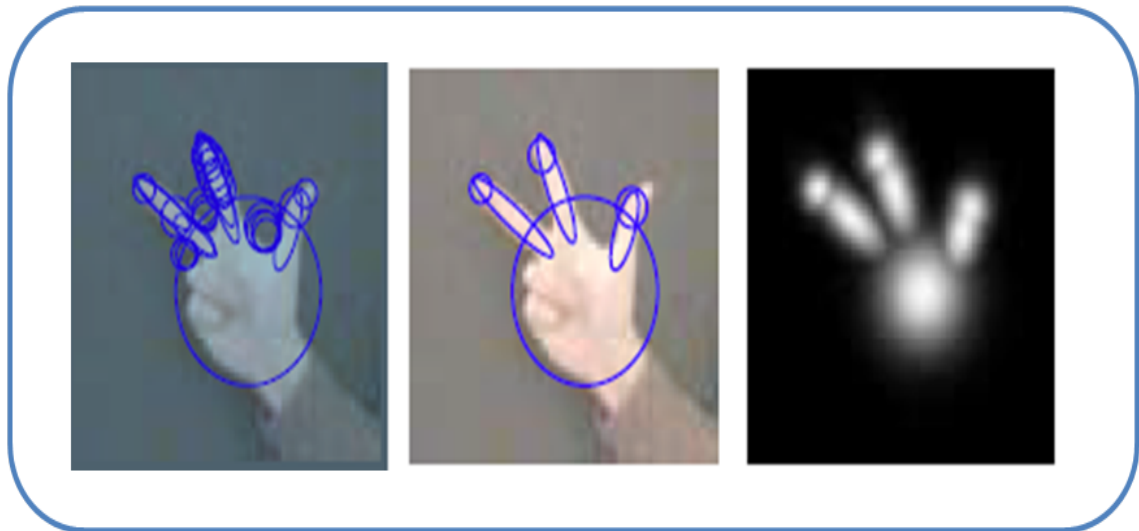


FIGURE 2.10 – Technique de suivi et de reconnaissance en utilisant les couleurs [105]

Wang et Popović[106] ont proposé une technique de reconnaissance du geste de la main en 3D en portant des gants colorés. Ces derniers sont modélisés et construits de manière à reconnaître quelques gestes de la main dynamiquement par une caméra (RGB). Cette technique offre à l'utilisateur une interaction naturelle avec l'environnement virtuel (voir Figure 2.11). Les résultats avec cette technique sont satisfaisants et convaincants, cependant, l'inconvénient avec ce système est qu'il ne reconnaît que cette forme de gants avec la configuration de couleurs montrée sur la Figure 2.11.



FIGURE 2.11 – Technique de gants colorés proposée dans [106]

Dans la même optique, Bellarbi [107] a proposé de porter des petits marqueurs à couleurs sur les doigts des mains de l'utilisateur (voir Figure 2.12). Cette technique permettra la manipulation des documents numériques projetés sur une table de réalité augmentée. Les marqueurs sont fabriqués avec des morceaux de papiers colorés pour ne pas encombrer et gêner l'utilisateur. Cependant, cette dernière permet de reconnaître un nombre limité de gestes statiques et assure une interaction en 2D. Certaines anomalies et présente certaines confusions sont observé : des couleurs portées avec celles de la scène.

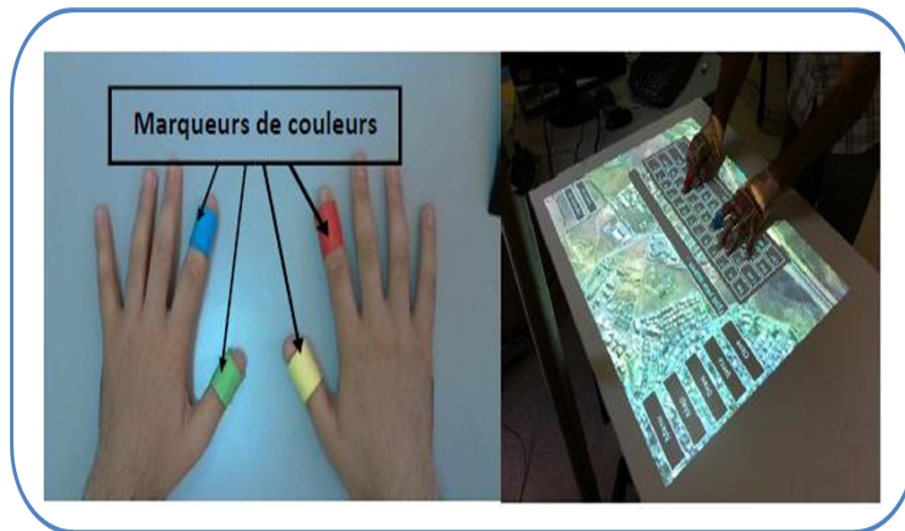


FIGURE 2.12 – Technique d’interaction basée marqueurs couleurs [107]

La forme géométrique de la silhouette de la main (surface, périmètre, convexité, l’orientation, etc) est une caractéristique avec laquelle on peut reconnaître la main de l’utilisateur.

Bridal [108] a présenté une méthode qui reconnaît les gestes de la main par la propriété géométrique de sa silhouette (Voir Figure 2.13).

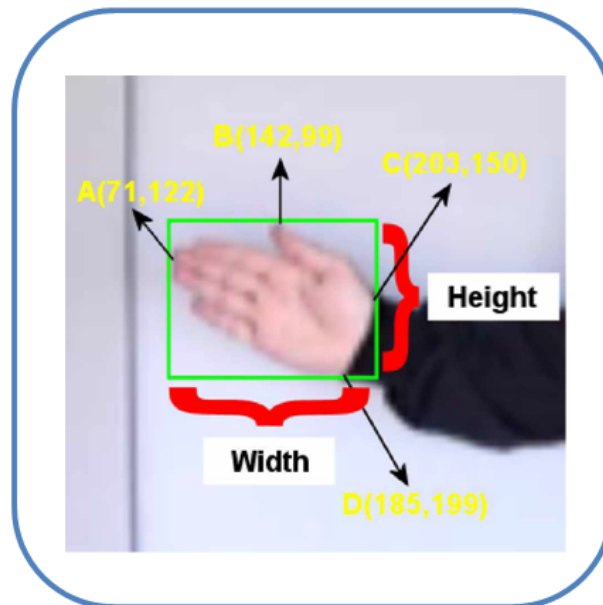


FIGURE 2.13 – Technique de reconnaissance de gestes par les propriétés de la main [108]

Pour pallier au différents problèmes rencontrés lors du port des marqueurs ou l'utilisation des différents capteurs, de nombreuses approches ont été proposées et ceci en traitant des images capturées par la camera. Certaines approches ont considéré le problème de la reconnaissance de gestes comme un problème de classification en faisant appel à des techniques de classification. Les réseaux de neurones ont été largement utilisés dans les systèmes de reconnaissance de gestes tel que [109], [110] et [111].

Les chaînes de Markov cachées (HMM) également sont largement utilisées notamment en reconnaissance de formes, en intelligence artificielle ou encore en traitement automatique du langage naturel. Nous citerons [112], [113] et [114].

Cependant, l'inconvénient avec ces approches c'est qu'elles ne garantissent pas des applications en temps réel. Plusieurs autres approches ont été proposées pour la reconnaissance de gestes, utilisant les moments invariants [115], les moments de Krawtchouk [116], Support

Vector Machine SVM [117], Gabor Filter [118](voir Figure 2.14), Dynamic Time Warpping DTW [119], Template Matching ou encore Randomized Decision Forests [120], [121].

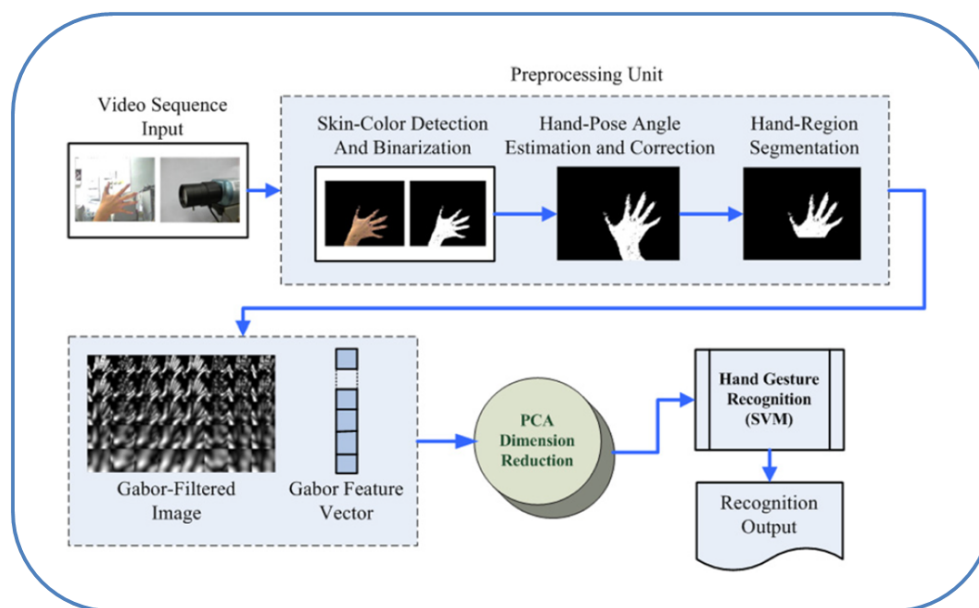


FIGURE 2.14 – Diagramme du flux de la technique de reconnaissance de geste par filtre de Gabor[118]

D'autres méthodes appliquent les descripteurs des points d'intérêt sur les images capturées, pour reconnaître les gestes de la main, on citera SIFT [122], SURF [123] ou encore LBP [124].

une technique de détection et de reconnaissance de gestes 2D de la main a été proposé par Bellarbi et ses associés [125], cette dernière utilise une simple webcam. Pour la détection et l'extraction de la main la technique proposée se base d'une part sur la binarisation adaptative et le calcul de l'histogramme, et pour la reconnaissance de gestes l'algorithme de Chain Code avec une version modifiée de la technique ASM (Approximate String Matching) est utilisé. Cette technique a été utilisée afin d'interagir en 2D avec une table interactive (voir Figure 2.15).

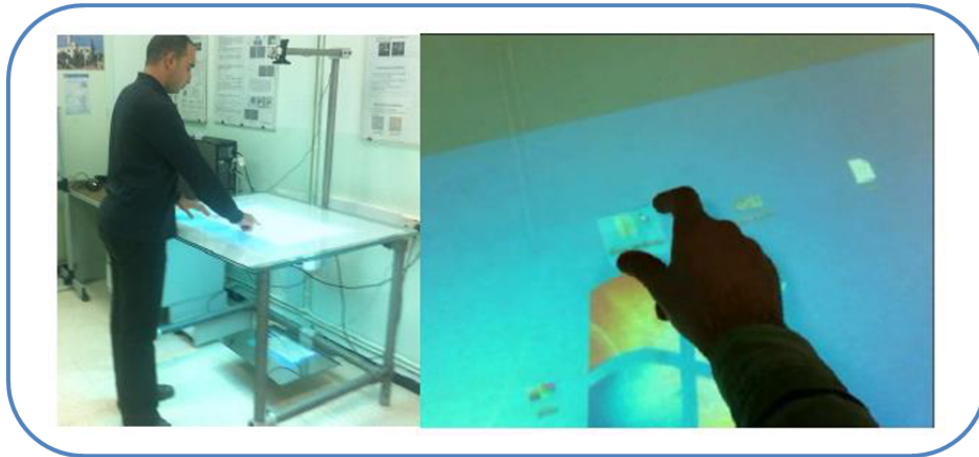


FIGURE 2.15 – La table interactive proposée dans [125]

Ce système a été utilisé par la suite comme une interface expert lors du projet de réalité augmentée [126] "RemoteGesture" qui a pour objectif de guider un technicien distant par les gestes d'un expert local (voir Figure 2.16).

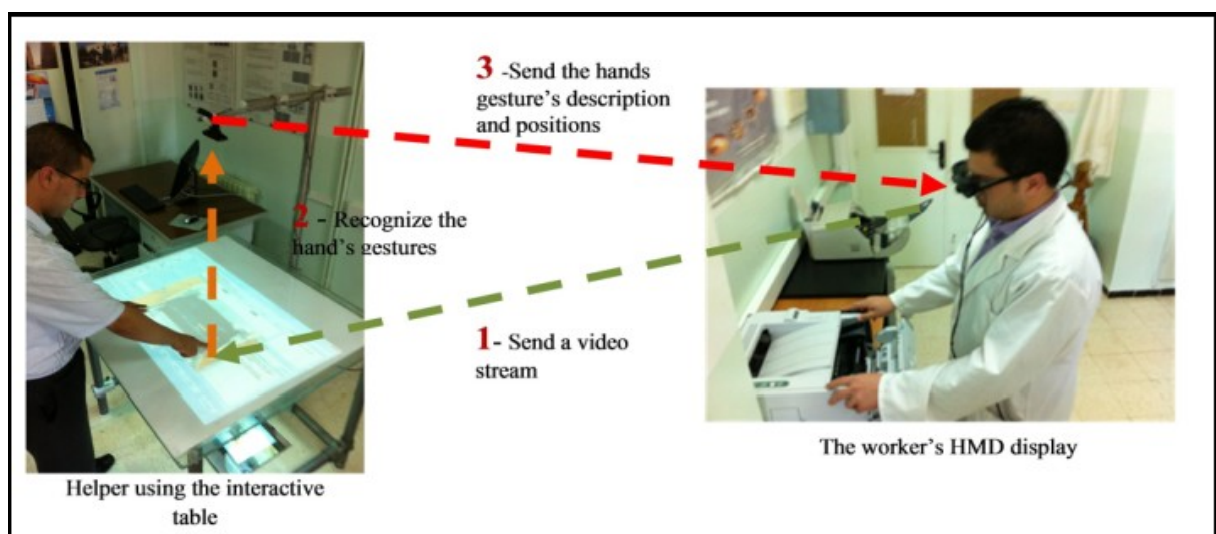


FIGURE 2.16 – Principe du projet "RemoteGestures" [126]

Ces techniques ont été utilisées dans de nombreux domaines et applications on citera le langage des signes, l'interaction avec des machines (robots, tables interactives, etc). Cependant, elles ne peuvent pas être utilisées en réalité virtuelle.

Avec l'émergence des caméras de profondeur, de nouvelles approches plus robustes ont vu le jour dans le but d'offrir une interaction 3D.

2.4.2.2 Techniques basées modèle 3D

Plusieurs approches de reconnaissance de gestes ont été proposées, et cela dans le but d'exploiter les avantages qu'offrent les caméras de profondeur, telles que Kinect de Microsoft, et Xtion de Asus (voir Figure 2.17).



FIGURE 2.17 – Quelques exemples de caméras de profondeurs. a) Creative Camera1. b) Azure Kinect DK. c) Xtion Pro2. d) Kinect. e) Kinect 2.f) Zed3 camera

Ces systèmes utilisent un modèle 3D de la main et tentent de faire correspondre ce modèle avec l'image observée. Deux manières sont possibles : estimation du modèle et mise en correspondance avec l'image ou extraction des caractéristiques puis l'estimation du modèle. Ces systèmes permettent d'évaluer les angles des articulations de la main. Ils fournissent donc une information très complète sur le geste effectué. L'évaluation des

paramètres du modèle est généralement difficile et coûteuse en terme de temps de calcul. Heap et Hogg[127] ont présentés une construction du modèle 3D de la main par les contours actifs (voir Figure 2.18).

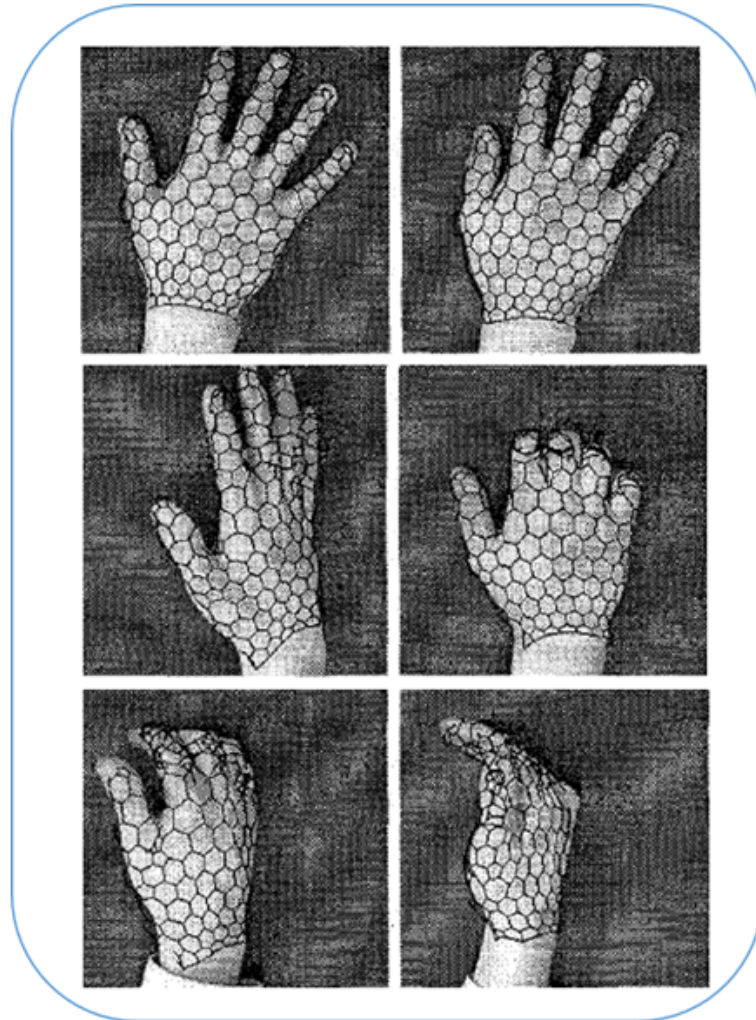


FIGURE 2.18 – Exemple d’un modèle 3D de la main [127]

Tomasi et ses associé [128] ont développé un modèle 3D de la main qui peut représenter la forme et les articulations de la main. Le modèle 3D peut être animé à l’aide de paramètres sur la pose de la main et les angles d’articulations. Dans leurs expérimentations ils ont implorés 24 signes de mais, chaque signe est représenté par 15 vues différentes.

Wu et ses coauteurs [129] ont quant à eux proposé le découplage de la pose de la main ainsi que les articulations des doigts et de les intégrer dans un processus itératif. Dans leur implémentation, les gestes sont effectués sur un fond clair et face à la caméra. Le système nécessite un calibrage manuel du modèle.

Athistos et Sclaroff [130] ont proposé une méthode qui permet de générer un ensemble de modèles qui pourrait correspondre à une image donnée, dans ce travail il s'agissait d'une main (voir Figure 2.19).

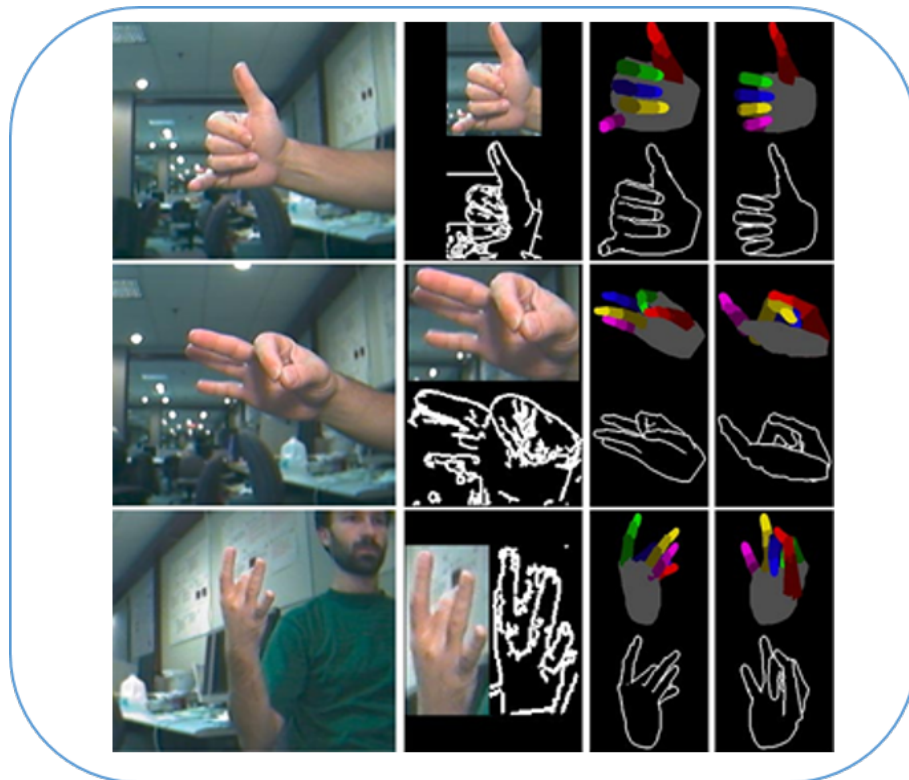


FIGURE 2.19 – Modèles 3D générés [130]

Pour surmonter le problème de complexité dans ce type de méthode qui est dû principalement au degré de liberté élevé et des différents points de vues, [131] ont proposé une approche par apparence 2D en utilisant les contours et qui permet d'estimer la posture 3D de la main (voir Figure 2.20).

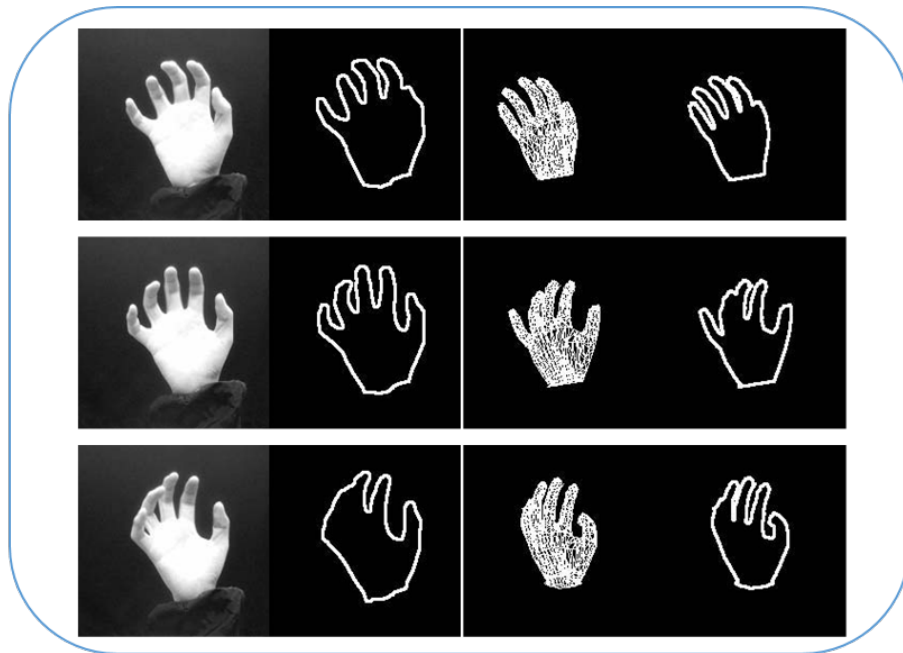


FIGURE 2.20 – Modèles 2D à partir des contours d’après (Imai, 2004)

Wang et ses associés [132] ont proposé une technique appelée 3Gear. L’idée de cette technique est de comparer les geste capturé par la Kinect à une panoplies de gestes appartenant a une base de données(voir Figure 2.21) afin de retrouvé celui qui lui suit le plus proche. La base données de gestes générée hors-ligne contient des images issues de la projection de différents gestes possibles à partir d’un modèle 3D de la main. Cette technique présente une reconnaissance robuste des gestes 3D de la main.

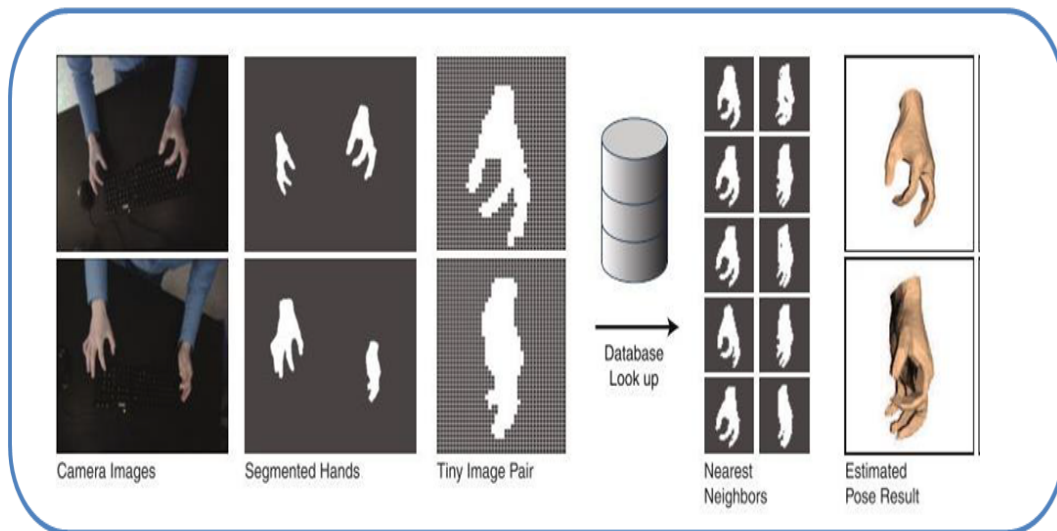


FIGURE 2.21 – Le principe de la technique 3Gear [132]

Avec un principe similaire, Khamis et ses associés [133] ont introduit un algorithme d'apprentissage automatique pour une recherche efficace dans une base de données des images de profondeur des gestes. Cette technique a présenté de meilleurs résultats. Cependant, elle est plus coûteuse en temps de calcul. Certains chercheurs ont modélisé la main selon des points d'articulation. Ils ont appliqué par la suite une fonction de coût, pour minimiser l'erreur entre le modèle 3D et la forme de la main réelle, afin d'estimer la position et le geste 3D de la main (voir Figure 2.22). Nous pouvons citer les travaux [134], [135] et [136].

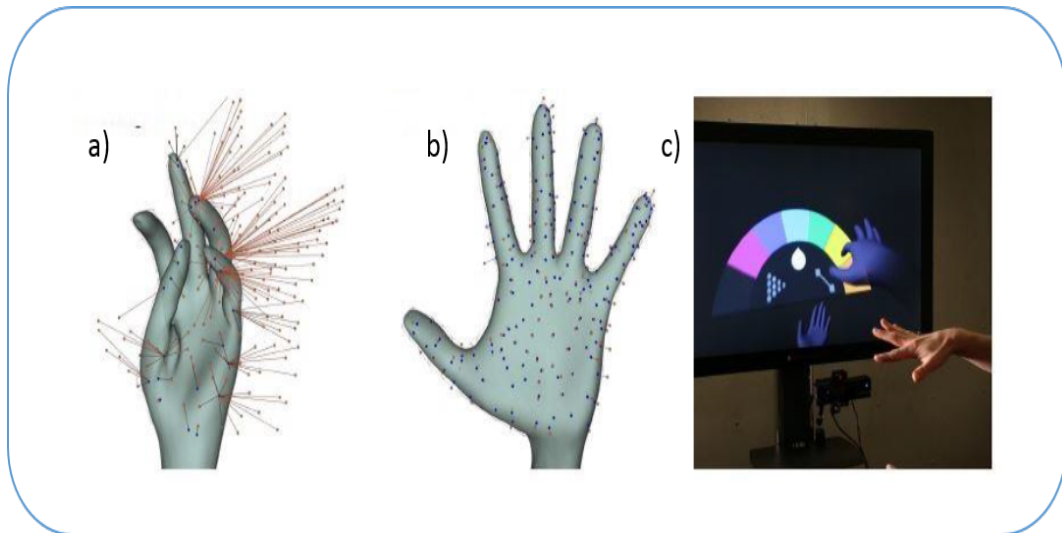


FIGURE 2.22 – Technique de reconnaissance et de suivi de geste de la main proposée dans [134]. a) Initialisation : modèle estimé de geste. b) Convergence vers le geste réel en appliquant une fonction d’optimisation de l’énergie. c) Exemple d’application en RV [133]

Bellarbi [137] a développé une technique d’interaction 3D basée sur la reconnaissance de gestes en RA, cette technique assure un suivi stable et précis de la scène réelle et assure également une interaction naturelle et intuitive avec les objets virtuels insérés (voir Figure 2.23).

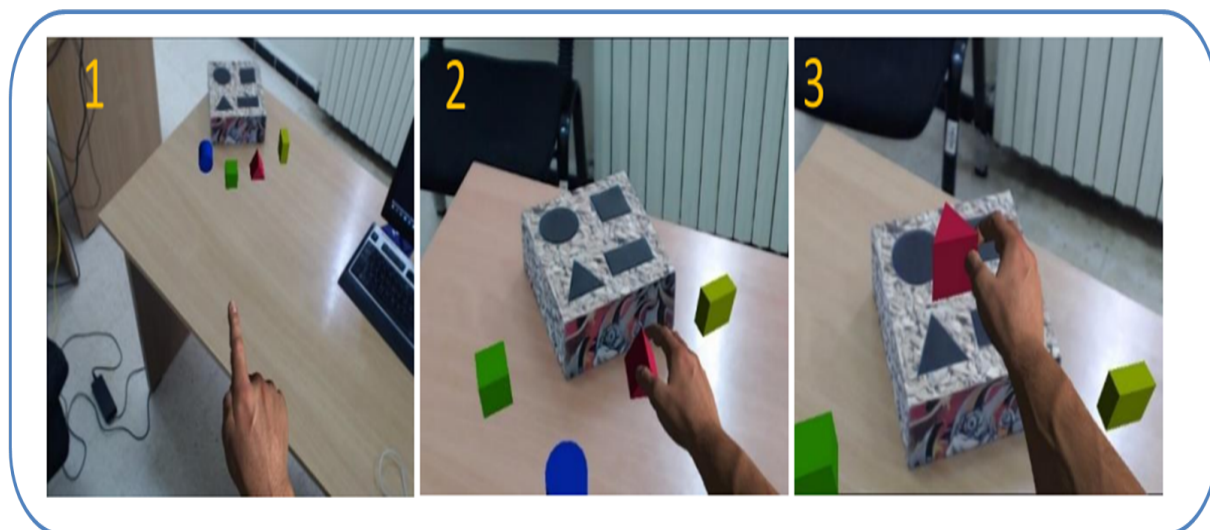


FIGURE 2.23 – Déroulement de la technique d'interaction : 1) pointage vers l'objet désiré. 2) un zoom est appliqué sur l'image. 3) l'utilisateur utilise la main virtuelle simple pour saisir et manipuler l'objet. [137]

Récemment, des techniques de reconnaissance de gestes 3D ont pris une autre dimension en termes de robustesse et de précision avec l'évolution des techniques d'apprentissage et l'apparition de l'apprentissage profond (Deep Learning). Nous pouvons citer plusieurs travaux tels-que [138] [139], [140], [141] et [142].

Ces techniques montrent une robustesse et une précision pour la reconnaissance de gestes de la main. On a constaté deux inconvénients majeurs : le premier étant matériel et les dispositifs de capture utilisés nécessitent une position fixe pour l'utilisateur (Kinect, Caméra, etc.), ce qui ne permettait pas une mobilité totale de l'utilisateur, le deuxième inconvénient concerne le champ de capture ou champs de vision qui est très limités pour ces capteurs.

Pour répondre a ces inconvénients Holz et ses associés [143] ont présenté le "Leap-Motion", un nouveau dispositif de reconnaissance et de suivi de gestes des mains, composé d'une partie soft et d'un contrôleur (matériel). Ce dernier est petit, léger de dimension 13 mm x 30 mm x 76 mm, il est doté de deux cameras infrarouge (voir Figure 2.24) et peut détecter

et reconnaître les mains dans un espace qui peut aller jusqu'à 50 cm. Ce dispositif est portable, il peut être placé sur n'importe quel casque de réalité virtuelle tel que Oculus-Rift, cette configuration permettra à l'utilisateur d'être mobile et d'avoir un champ d'action beaucoup plus important.



FIGURE 2.24 – Le dispositif Leap-Motion [143]

Le Leap-Motion, est très utilisé dans la RA et/ou RV. Son efficacité, sa mobilité, et son prix abordable le rendent accessible au grand public. Ce dernier a été utilisé dans plusieurs travaux nous citons [144], [145], [146] et [147].

RealSense³ une nouvelle technologie dévoilé par Intel en juillet 2016, elle consiste en un dispositif compact composé de trois types de capteurs (une caméra RGB de résolution 1080p, une caméra infrarouge et un projecteur laser infrarouge) (voir Figure 2.25).

3. <https://www.intel.fr/content/www/fr/fr/architecture-and-technology/realsense-overview.html>

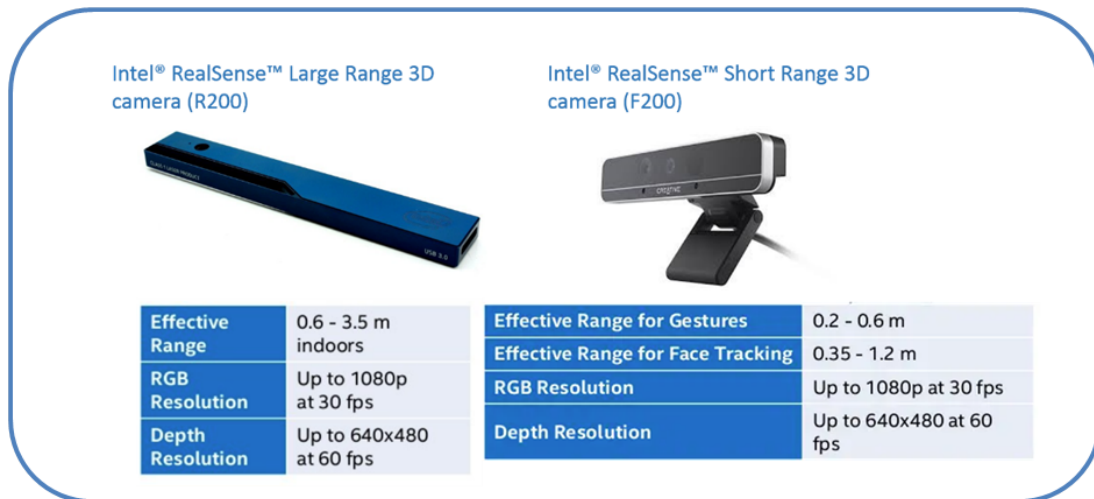


FIGURE 2.25 – Cameras RealSense de Intel

Cette combinaison de capteurs permet de reconnaître et suivre les mouvements et les gestes de l'utilisateur d'une façon robuste (Selon le constructeur Intel). Ainsi, le SDK fourni avec le système permet de reconnaître un ensemble de gestes de la main (voir Figure 2.26).










Grab		Start with an open hand (Big-5) and close together thumb and index finger, leave the rest of the fingers open naturally.
Release		Release pinchig gesture into an open hand (Big-5) pose.
Move		While holding thumb and index finger together, move in x-y-z axes.
Swipe		Swipes can be performed in both horizontal and vertical directions parallel to the screen.
Zoom In /Grow		You can use Big 5 or 2 finger-pinch as the activation and completion pose. Zooming is coupled to the distance between the 2 hands (similar to pinch-zooming on touchscreens).
Zoom Out / Shrink		Same as zoom in, in the opposite direction.
Push to Select		Both are good designs to consider. Developers should choose a style that fits their app.
Hover Select		Hold open palm (Big-5) gesture still.
Escape/Reset		Wave an open hand from side to side naturally to reset or escape from an application mode.

FIGURE 2.26 – Les gestes prédéfinis par le système RealSense

Pour afficher et visionner un environnement virtuel, plusieurs casques ont vu le jour depuis 2012, certains sont seulement utilisés pour l’affichage et d’autres contiennent des caméras intégrées pour la reconnaissance de gestes.



FIGURE 2.27 – Casques a) : Oculus Rift DK1, b) : Oculus Rift DK2, c) : oculus Rift-S, d) : HTC Vive, e) : Oculus Quest.

Oculus Rift DK1 est le premier casque d’affichage lancé en 2013(voir Figure 2.27.a), il dispose d’un écran 7 pouces 1280 x 800 pixels. La reconnaissance des mouvements est cadencée à 1000 Hz et se fait sur 3 axes grâce à un gyroscope, un accéléromètre et un magnétomètre. La version DK2 de l’Oculus Rift a été lancée en 2014 (voir Figure 2.27.b), Cette version intègre les dernières avancées réalisées sur réductions de vibration et du flou cinétique, sur les détections de mouvement et la nouvelle résolution de l’écran, à

savoir une dalle full HD de 1080p. Par la suite en 2016 il y a eu le lancement de HTC Vive (voir Figure 2.27.e).

Oculus Rift-S a fait son apparition en 2019 (voir Figure 2.27.c). Il se différencie par l'usage d'un seul écran LCD de 2560 x 1440 pixel, avec un taux de rafraîchissement maximum de 80 Hz. Il adopte également un nouveau système de tracking, utilisant cinq caméras intégrées au lieu des capteurs externes, pour se repérer dans l'espace et suivre le mouvement des contrôleurs.

Tous ces casques sont câblés ce qui rend l'utilisateur peut confortable, l'apparition de l'Oculus-Quest (Voir Figure 2.27.d), ce dernier est un appareil entièrement autonome et sans fil, il utilise deux écrans OLED, chacun avec une résolution individuelle de 1440 x 1600 et une fréquence de rafraîchissement de 72 Hz. Son système s'appuie sur quatre caméras infrarouges situées à chaque coin du casque pour suivre le casque ainsi que les manettes dans l'espace délimité. L'Oculus-Quest assure le suivi et la reconnaissance de quelques gestes des mains de l'utilisateur même dans sa vision périphérique.

Cependant, l'Oculus-Quest est incapable de détecter les mains qui se chevauchent dans le champ de vision ainsi que les gestes combinés. De plus, l'algorithme de suivi des mains utilisées sur plusieurs caméras a créé une sorte de coupure qui fait que parfois l'appareil perd ses capacités de suivi lorsque les deux mains se trouvent à un certain niveau de proximité. L'algorithme de reconnaissance de gestes utilisé dans ce dernier a besoins de raffinement pour arriver à la stabilité voulue.

2.5 Bilan et discussion

Dans la précédente partie du chapitre, nous avons abordé les différentes approches et techniques permettant la reconnaissance des gestes de la main de l'utilisateur ainsi

que les différents types de capteurs et de technologies utilisés. Ainsi, nous présentons une classification de quelques techniques de reconnaissance de gestes proposées dans la littérature. Nous nous focalisons sur les techniques basées vision, nous avons classifié ces dernières selon le type de capteur utilisé et la forme du geste reconnu (voir Table 2.1). Selon Table 2.1, nous constatons qu'avec l'émergence des nouveaux capteurs de profondeur tel que la Kinect, les caméras de RealSense ou encore le Leap-Motion, la reconnaissance de gestes en 3D est devenue possible (sans pour autant encombrer l'utilisateur avec des câbles ou du matériel physique gênant). Pendant cette étude on a remarqué que l'utilisation de ce type de capteurs améliore de plus en plus les résultats en termes de stabilité et de précision. Ce qui reste insuffisant pour développer des techniques d'interaction 3D en se basant sur la reconnaissance de geste, la stabilité et la précision reste un élément très important lors du choix du matériel à utiliser.

La Kinect est un capteur très utilisé pour la détection de personnes, l'identification de leurs postures et leurs mouvements utilisé dans différentes applications allant des interactions gestuelles aux jeux vidéo. Le premier inconvénient de ce dernier étant le champ d'utilisation qui est réduit à 120 degrés ce qui rend l'interaction 3D très limitée, le deuxième problème c'est l'immobilité de la Kinect ce qui permet à l'utilisateur de rester figer dans une position précise.

Le casque Oculus-Quest est un casque non câblé ce qui permet à l'utilisateur d'être libre lors de son déplacement, ce dernier permet une vision périphérique ce qui est un avantage considérable. Ce casque assure le suivi et la reconnaissance de quelques gestes des mains de l'utilisateur, mais malheureusement les gestes de la main reconnus par l'Oculus-Quest ne sont stable ni précis ce qui rend l'utilisation du Quest pas très apprécié lors du développement des techniques d'interaction 3D. Le Leap-Motion reste le capteur le

TABLE 2.1 – Classification des techniques de reconnaissance de gestes

Techniques	Classes	2D / 3D	Capteur (s)	Points forts	Points faibles
[109]	Basées apparence	2D	RGB webcam	Interaction naturelle	Temps réel
[114]		2D	RGB webcam	Reconnaissance robuste	Temps réel
[110]		2D	RGB webcam	Interaction naturelle	Temps réponse
[106]		3D	RGB webcam	Interaction naturelle	Forme de gants fixe
[107]		2D	RGB webcam	Reconnaissance facile	Confusion de couleurs
[118]		2D	RGB webcam	Interaction précise	Stabilité
[119]		2D	RGB webcam	Temps réel	Stabilité
[115]		2D	RGB webcam	Interaction précise	Stabilité
[125]		2D	RGB webcam	Stabilité	Nombre de gestes limités
[127]		3D	RGB webcam	Suivi robuste	Reconnaissance dynamique
[130]		3D	RGB webcam	Reconnaissance robuste	Mise en oeuvre complexe
[131]		3D	RGB webcam	Mise en oeuvre simple	Temps de calcul
[132]		3D	Kinect	Interaction précise	Positions des mains
[143]		3D	Leap-Motion	Reconnaissance robuste	Stabilité
[136]		3D	Kinect	Interaction naturelle	Temps de réponse
[144]	Basées modèle 3D	3D	Leap-Motion	Interaction précise	Apprentissage des gestes
[135]		3D	Kinect	Interaction naturelle	Temps de réponse
[139]		3D	Kinect	Interaction robuste	Position fixe de l'outil
[133]		3D	Kinect - V2	Interaction naturelle	Position de l'outil
[145]		3D	Leap-Motion	Interaction précise	Apprentissage des gestes
[146]		3D	Leap-Motion	Interaction rapide	Apprentissage des gestes
[142]		3D	Kinect	Interaction robuste	Position fixe de l'outil
[134]		3D	Kinect - V2	Interaction naturelle	Position de l'outil
[137]		3D	Leap-Motion	Interaction précise	Temps de réponse

plus avantageé pour la reconnaissance de gestes (poids, taille, champs d'action, portable sur un casque, etc).

Nous avons constaté durant ce chapitre que l'évolution de la technologie des interfaces naturelles utilisateur (Natural User Interface, NUI) a donné un nouvel élan au domaine de l'interaction 3D avec plus de possibilités et de moyens, d'interagir de manière naturelle. L'utilisation des NUI pour développer des techniques d'interaction 3D peut mener à des applications réelles et concrètes.

L'autre défi qui s'ajoute à ces derniers, est celui de fournir une interaction naturelle qui permet à l'utilisateur de surpasser les problèmes liés aux différentes interfaces physiques. L'approche basée sur les interfaces naturelles NUI(Natural User Interfaces) facilite grandement l'interaction 3D pour l'utilisateur, en lui offrant des interfaces naturelle , intuitive et flexible.

Dans ce qui suit, nous souhaitons utiliser la reconnaissance de gestes pour développer une nouvelle technique d'interaction 3D, cette dernière permettra une interaction naturelle, intuitive, rapide et précise notamment avec les objets distants, occultés et se trouvant dans un milieu encombré.

2.6 Conclusion

Tout au long de ce chapitre, nous avons abordé les modalités d'interaction naturelles notamment le geste de la main. Nous nous sommes penchés par la suite sur les techniques et technologies existantes permettant la reconnaissance des gestes de la (les) main (s). L'étude menée dans ce chapitre, sur les différentes techniques et technologies de reconnaissance de gestes de la main, nous a permis de voir les possibilités en matières d'équipements et

de techniques de reconnaissance de gestes pouvant nous être utiles pour la suite de notre travail.

Nous présentons dans la seconde partie de ce mémoire nos contributions scientifiques dans ce domaine, qui concerne les deux parties à savoir l'interaction 3D et le travail collaboratif dans un environnement virtuel.

Chapitre

3

Contribution à l'interaction 3D en RV : "*Zoom-fwd*"

3.1 Introduction

L'interaction 3D est une composante principale de la RV, elle permet aux utilisateurs de communiquer avec l'EV.

A l'issu de l'état de l'art ainsi que le bilan présenté dans la partie précédente de ce manuscrit, nous avons constaté que les techniques d'interactions dans le domaine de la RV sont nombreuses mais elles ne répondent pas aux exigences de certaines applications telles que la précision lors de l'interaction aussi bien dans un environnement simple que complexe. La complexité de l'environnement varie selon les objets le constituant (petits ou grands, loin ou proches, occultés ou visibles, etc).

Afin d'interagir avec un environnement virtuel l'utilisateur a besoin de matériels ou d'outils physiques qui sont une contrainte supplémentaire (poids de ces outils, limites logicielles, etc). Dans cette optique, nous avons proposé une technique d'interaction 3D qui pallie à différents problèmes liés aux interfaces physiques.

Cette dernière appelée "*Zoom-fwd*" : "*Zoom-forward*" basée essentiellement sur la reconnaissance de gestes [1], [2] et [3]. C'est une technique de sélection et de manipulation. Dans ce chapitre, nous présentons cette technique, ses principes ainsi que sa conception

et son implémentation. Une étude expérimentale est présentée à la fin du chapitre, afin d'évaluer ses performances.

3.2 Développement d'une technique d'interaction 3D en utilisant la reconnaissance de gestes dans un environnement virtuel

Dans un travail précédent [1], nous avons développé une application utilisant la technique 3Gear. L'objectif étant de permettre à un utilisateur d'interagir avec les objets 3D dans un environnement virtuel en utilisant les gestes de la main, c'est à dire d'implémenter une technique de sélection et de manipulation permettant de réaliser l'interaction souhaitée.

3.2.1 Conception du setup

Le système utilisé se présente comme suit (Voir Figure 3.1). Il comporte une Kinect qui est positionnée au-dessus de l'utilisateur, sert de capteur qui reconnaît les gestes des mains de l'utilisateur. L'utilisateur porte un HMD (Head Mounted Display) ou un casque vidéo qui lui permettra une certaine immersion.

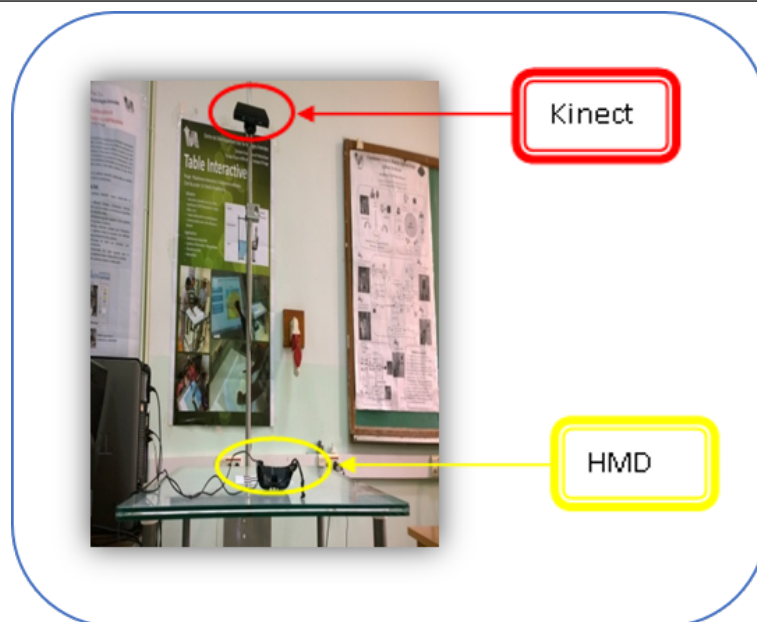


FIGURE 3.1 – Dispositif matériel utilisé [1]

3.2.2 Conception du système

Notre système est basé sur le processus suivant (Voir Figure 3.2), le module acquisition des données permet de collecter des données depuis la Kinect (images) et le HMD (orientation de l'utilisateur). Ces données sont transmises aux modules détection. Le module de détection extrait les mains de l'utilisateur dans les images fournies par la Kinect. Une fois les mains de l'utilisateur détectées, le module de suivi détermine la position et la représentation des mains 3D. Le module de reconnaissance donne une interprétation aux gestes de l'utilisateur, cette interprétation peut être vue dans l'interface virtuelle comme une interaction 3D avec des objets dans l'environnement virtuel (sélection, manipulation et navigation).

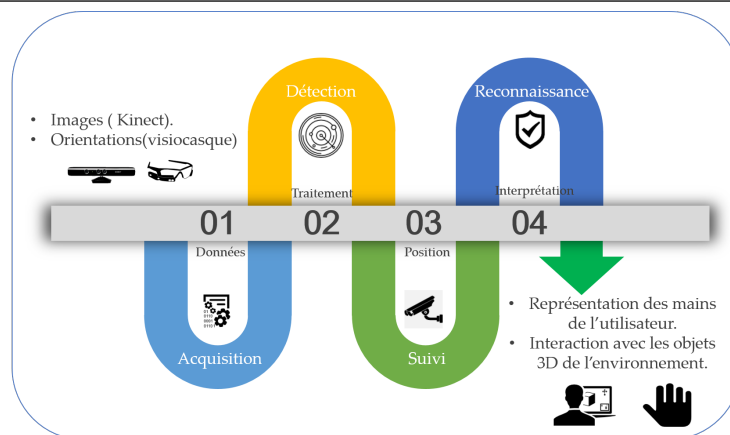


FIGURE 3.2 – L'architecture logicielle du système développé[1]

3.2.3 Les gestes utilisés

Nous avons défini trois gestes pour interagir avec l'environnement (Voir Figure 3.3). Par conséquent, l'utilisateur pointe son doigt afin de naviguer dans l'environnement virtuel, cette action lui permet de se déplacer vers l'avant. Alors que le geste L lui permet de reculer. De plus, l'utilisateur peut pointer ou faire des gestes L à deux mains pour se déplacer plus rapidement. En plus de cela, l'utilisateur peut tourner sa tête vers la gauche / droite pour naviguer dans le sens du regard. La manipulation des objets déjà sélectionné se fait par le geste de pincement.

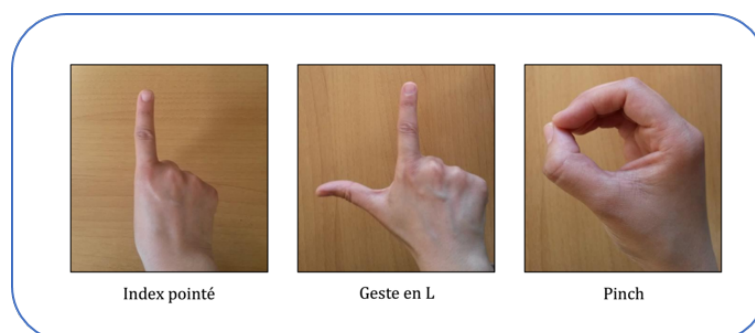


FIGURE 3.3 – Les différents gestes utilisés

3.2.4 Résultats obtenus

Cette application intuitive et naturelle utilise la "Main Virtuelle Simple" comme technique d'interaction 3D (sélection Figure 3.4.b / manipulation Figure 3.4.c). Ce système permet à l'utilisateur d'être plus flexible lorsqu'il effectue un ensemble de tâches (mains libres). De plus, les techniques choisies réduisent les erreurs dues à une mauvaise détection et à la variation de luminosité et offrent un moyen d'interaction économique et pratique.

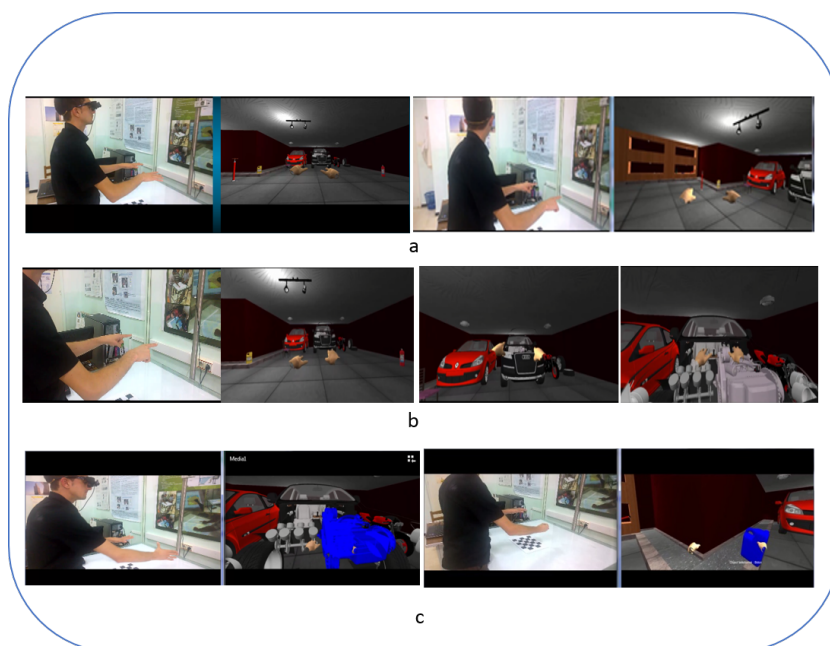


FIGURE 3.4 – Les résultats obtenus : a : navigation dans l'environnement virtuel, b : sélection des objets, c : manipulation des objets sélectionnés [1]

Durant cette première expérience, nous avons constaté une fidélité lors de la reconstruction des différents gestes de la main en temps-réel ce qui facilite la tâche pour l'utilisateur. Nous avons également remarqué une grande précision lors de l'interaction 3D surtout lors de la manipulation des différents objets qu'ils soient grand ou petits.

Cependant nous avons relevé un certain nombre de problèmes.

Le premier est lié principalement au matériel utilisé (Kinect), où l'utilisateur devrait tout le temps avoir les mains dans la même position qu'au moment du calibrage, ce qui est difficile car cela exige une concentration constante de l'utilisateur dans l'EV.

L'autre problème rencontré est celui du champ de vision de la Kinect qui est très étroit ce qui limite son utilisation.

Le troisième problème est lié à la technique d'interaction implémentée : "Main virtuelle simple". Nous avons constaté durant les tests qu'elle ne permettait pas l'interaction avec les objets éloignés et occultés dans un environnement encombré. Pour pallier aux problèmes cités précédemment, nous avons donc remplacé la Kinect par le Leap-Motion et le casque Vusix par l'Oculus Rift (pour une meilleure immersion de l'utilisateur).

Nous avons implémenté "*Zoom-fwd*" une nouvelle technique d'interaction 3D, qui va permettre une interaction précise et rapide avec tous types d'objets : lointains, petits, grands, proches, accessibles, partiellement et /ou complètement occultés.

3.3 "*Zoom-fwd*" technique d'interaction 3D efficace pour la sélection des objets distants et occultés

3.3.1 Conception du Setup

L'outil principal utilisé est le Leap-Motion [143]. C'est un outil léger, portable et offre une grande fidélité lors de la reconstruction des mouvements de la main de l'utilisateur dans l'EV. La configuration de notre système est composée d'un Leap-Motion qui est placé sur le casque de réalité virtuelle "Oculus Rift DK2" [148]. Cet emplacement permet à l'utilisateur une immersion totale et un déplacement facile et sans contrainte. Le retour

visuel de la reproduction des gestes et le résultat des actions sont affichés sur le casque (Voir Figure 3.5).



FIGURE 3.5 – Dispositif immersif utilisé [3]

3.3.2 Conception du système

La première étape consiste à identifier le déroulement des opérations de notre système par l'identification des différents modules, de leurs rôles ainsi que des données et des informations qui circulent entre eux (Voir Figure 3.6).

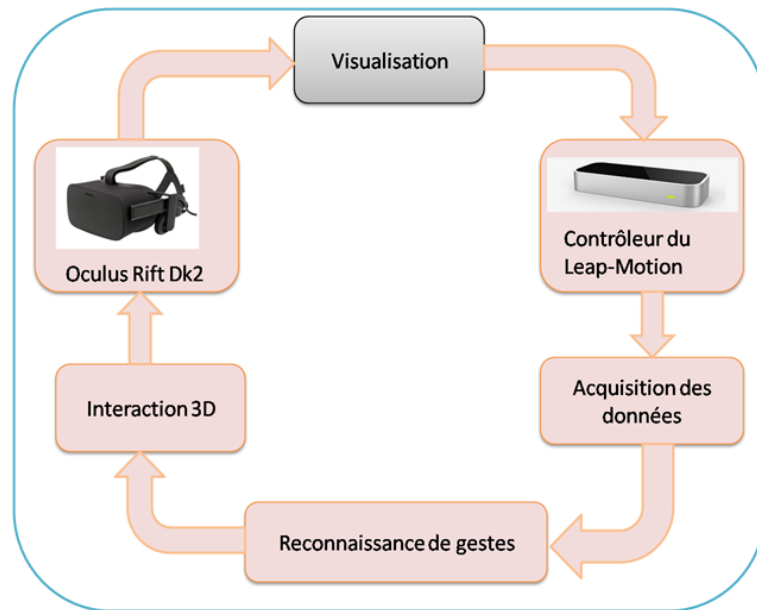


FIGURE 3.6 – Architecture logicielle du système développé[3]

3.3.2.1 Le module acquisition des données

Ce module collecte les différentes données. Le Leap-Motion fournit les informations spatiales 3D des articulations de mains de l'utilisateur. L'orientation de la tête de l'utilisateur est récupéré à partir de l'oculus rift. Ces informations sont transmises au module reconnaissance de gestes.

3.3.2.2 Le module reconnaissance de gestes

La méthode basée-3D-squelette du modèle basée-modèle-3D pour la reconnaissance des gestes est utilisée cela en se basant sur les différentes positions des jonctions des deux mains et de la forme du squelette de la main.

Ce module utilise les informations fournies par le module précédent(positions et orientations des différentes articulations des deux main ainsi que les positions et orientations des deux poignets), ce module détermine la position et la représentation des mains en

3D. Pour ce faire, un modèle squelette 3D des mains (voir Figure 3.7) est considéré pour chaque main (gauche/droite).

Les articulations de la main sont considérées comme des points de repère locaux. La position et l'orientation 3D de chaque articulation est donné par rapport au repère du poignet. L'articulation du poignet est à son tour donné par rapport au repère du Leap-Motion.

Pour concevoir et définir les différents gestes de la main avec lesquels l'utilisateur va interagir, on utilisera la distance entre les différentes articulations de la main ainsi que leur orientations. Ces informations sont transmises au module d'interaction 3D.



FIGURE 3.7 – Les différentes articulations de la main fournis par contrôleur de Leap-Motion [149]

3.3.2.3 Le module interaction 3D

Ce module gère le rendu graphique de l'interaction 3D et transmet à l'EV (Oculus Rift).

Pour concevoir et/ou implémenter cette partie, nous devons savoir quelles métaphores les utilisateurs doivent utiliser et quelles techniques d'interaction 3D doivent être utilisées pour interagir avec l'EV.

3.3.3 Approche proposée

- **Les étapes de l'approche proposée**

Nous avons mentionné dans le premier chapitre section 1.6 que le problème que nous voulons traiter est la précision lors de l'interaction 3D, en particulier avec un objet distant, occulté et se trouvant dans un milieu encombré. Notre contribution se situe au niveau de la sélection des objets, qui est une étape cruciale dans l'interaction. Pour obtenir une technique d'interaction 3D efficace, nous avons introduit des techniques de navigation, de manipulation et de contrôle d'application. Le principe de fonctionnement de la technique "*Zoom-fwd*" est montré dans le schéma suivant (Voir figure 3.8)

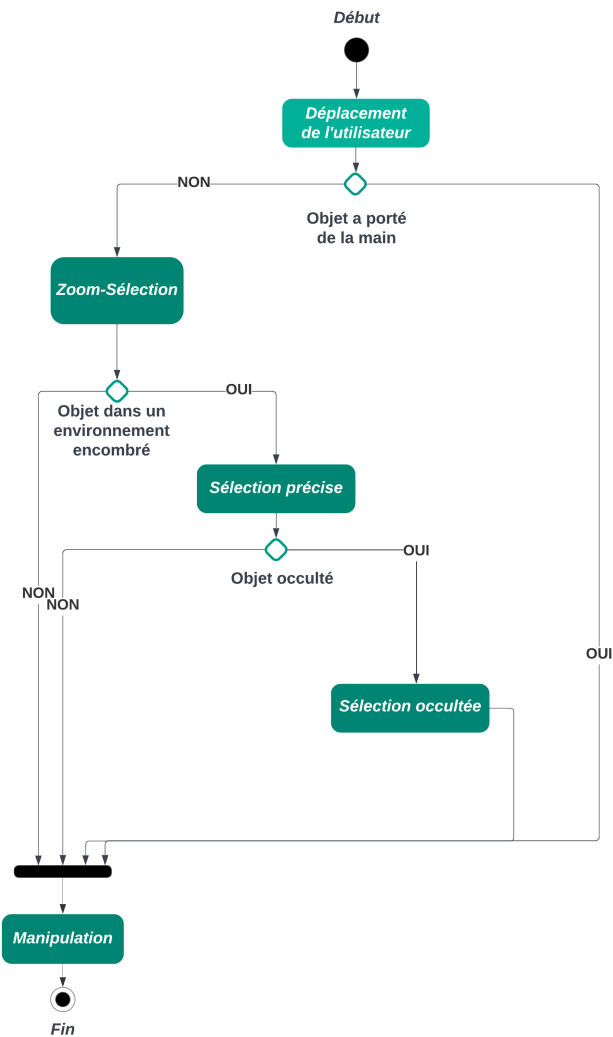


FIGURE 3.8 – Le principe de fonctionnement de la technique "Zoom-fwd" [3]

Dans ce qui suit nous expliquerons les différentes parties de l'approche proposée.

a) Déplacement de l'utilisateur dans l'environnement virtuel

L'utilisateur se déplace dans l'environnement réel et ses mouvements seront efficacement reproduits dans l'EV tant qu'il se trouve dans la zone de couverture du Leap-Motion (Voir figure 3.9). Cette zone étant limitée, nous avons ajouté la possibilité de se

déplacer sans se déplacer réellement, en utilisant les gestes de la main. L'utilisateur peut se déplacer en avant, en arrière, à gauche ou à droite dans le monde virtuel.

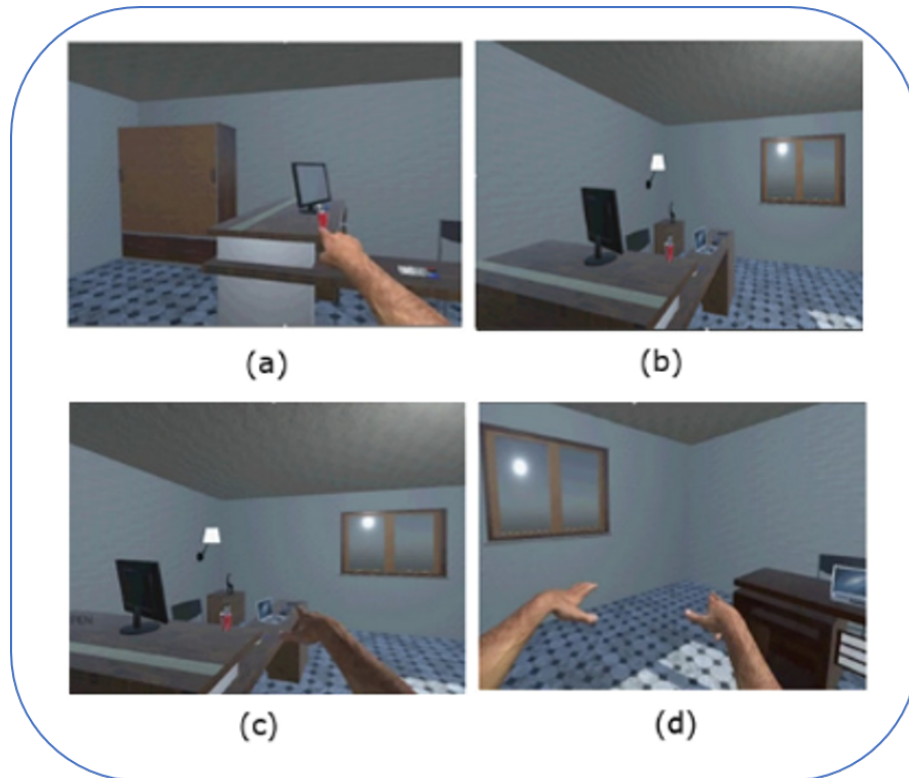


FIGURE 3.9 – Déplacement dans l'environnement virtuel[3]

b) Sélection 3D dans l'environnement virtuel

La technique "*Zoom-fwd*" comporte plusieurs étapes pour sélectionner l'objet souhaité (Voir figure 3.8) :

1. La première étape est de vérifier si l'objet que l'on veut sélectionner est proche de la main de l'utilisateur et non occulté par d'autres objets alors il sera directement sélectionné et manipulé par la suite ;
2. Si la première étape est non vérifiée, dans ce cas on peut avoir deux scénarios soit que l'objet est occulté ou lointain (l'objet n'est pas atteint par la main de l'utilisateur). Les deux étapes seront activées respectivement : Sélection occultée et Sélection précise.

c) Manipulation 3D

Après avoir utilisé la technique "*Zoom-fwd*" pour sélectionner l'objet souhaité, la tâche suivante est la manipulation. Nous choisissons la métaphore de la main virtuelle qui est utilisée pour déplacer l'objet de la position de départ à la position souhaitée.

• Principe de fonctionnement de la technique "*Zoom-fwd*"

La technique de sélection "*Zoom-fwd*" est utilisée lorsqu'on veut sélectionner un objet qui n'est pas à la portée de la main de l'utilisateur soit par le fait qu'il est loin, occulté par un autre objet ou qu'il se trouve dans un milieu encombré et que sa sélection nécessite beaucoup de précision.

d) Zoom-sélection

Cette partie est activée quand l'objet que l'on veut sélectionner est loin, et que la main de l'utilisateur ne peut pas l'atteindre directement (Voir figure 3.10). Dans ce cas, l'utilisateur effectue un geste de la main (pointe son doigt). Un zoom dynamique est ensuite appliqué jusqu'à ce que l'objet désiré soit suffisamment proche pour être à la portée de la main de l'utilisateur. Ce zoom dépend de la distance entre la main de l'utilisateur et l'objet désiré. Le principe est expliqué ci-dessous (Voir figure 3.11).

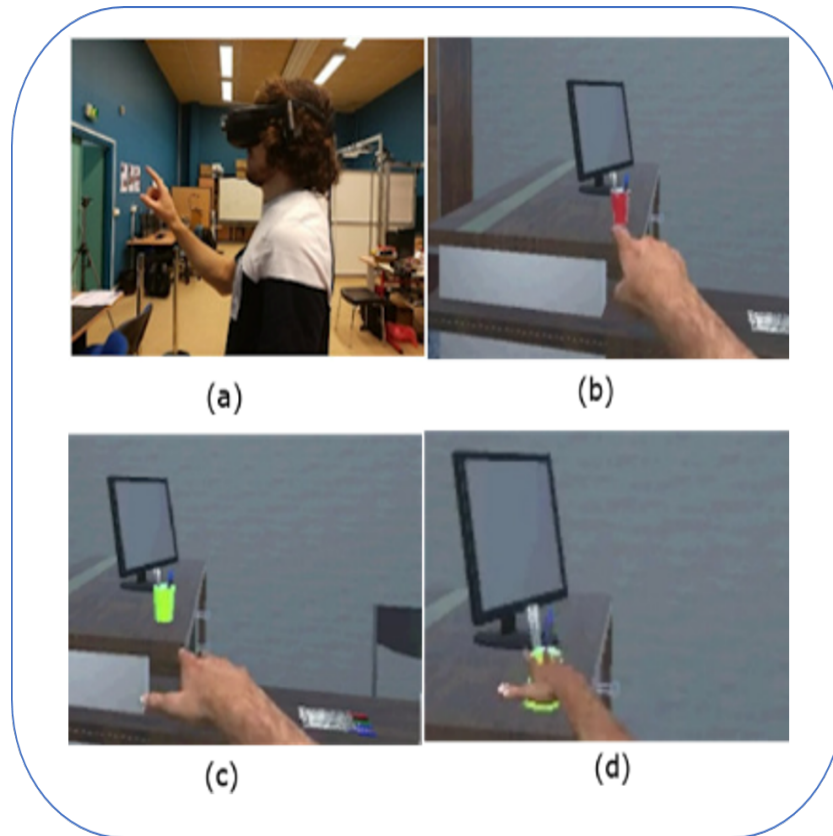


FIGURE 3.10 – Illustration du fonctionnement de la partie Zoom-sélection [3]

— Formalisation du problème

1) Calcul du facteur zoom F_z

Pour avoir l'objet distant dans la main de l'utilisateur, l'idée est de déterminer la distance entre l'objet virtuel sélectionné et la main de l'utilisateur. Après avoir obtenu la distance, le processus calcule dynamiquement un rapport de zoom appelé facteur de zoom " F_z ", ce dernier est appliqué à la caméra virtuelle. Nous pouvons présenter ce problème comme suit (Voir Figure 3.11).

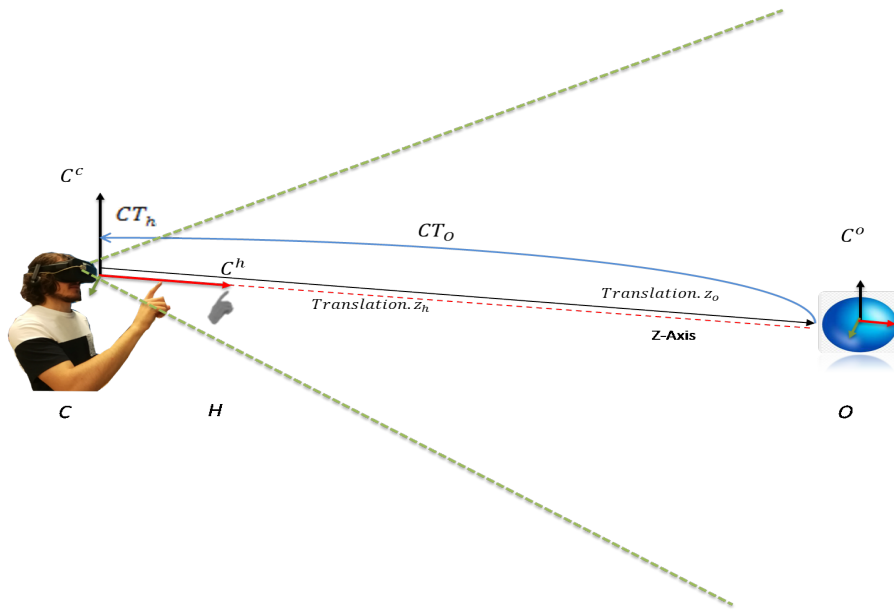


FIGURE 3.11 – Le principe du fonctionnement de la partie Zoom-sélection [3]

C^c et C^o représentent les coordonnées de la caméra C et de l'objet sélectionné O , respectivement. C^h représente les coordonnées de la main de l'utilisateur H . Toutes ces coordonnées sont par rapport au repère monde.

CT_o et CT_h représentent respectivement les matrices de transformations de l'objet sélectionné et de la main virtuelle suivie par rapport aux coordonnées de la caméra C^c .

$Translation.z_o$ et $Translation.z_h$ représentent respectivement les translations sur l'axe-Z de l'objet sélectionné obtenue à partir de la matrice CT_o et de la main virtuelle suivie à partir de la matrice CT_h .

Pour que l'objet sélectionné soit à la portée de la main de l'utilisateur, il faut passer de la distance $Translation.z_o$ à la distance $Translation.z_h$, qui est donné par le facteur " F_z "; ce facteur est calculé par la formule suivante (Voir équation 4.1).

$$F_Z = Translation.z_O.Translation.z_h^{-1} \quad (3.1)$$

2) Calcul des paramètres caméra

Une fois le facteur de zoom (rapport de zoom) calculé, les paramètres extrinsèques de la caméra sont mis à jour dynamiquement, et cela résulte par l'équation 4.4 suivante.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (R, t) \begin{pmatrix} 0 \\ 0 \\ F_Z \\ 1 \end{pmatrix} \quad (3.2)$$

$(x, y, z)'$ étant la position de la caméra dans le monde réel, (R, t) est la matrice de rotation/translation(3,4) représentant les paramètres extrinsèques de la caméra C.

e) Sélection précise

Cette partie est activée quand l'objet qu'on souhaite sélectionner se trouve dans un endroit encombré, la sélection devient dans ce cas-là très délicate et elle demande beaucoup de précision.

La précision pendant la sélection est un problème qui reste d'actualité et qui n'a pas encore été résolu.

La sélection précise est basée sur l'idée de mettre en évidence le voisinage de l'objet que nous souhaitons sélectionner, même s'il est dans un environnement encombré et entouré d'autres objets dont les formes sont similaires ou différentes de la sienne.

La technique "Zoom-fwd" offre une solution pour permettre une sélection précise de tout type d'objets (petits, fins, dans un environnement encombré etc).

— Principe de la sélection précise

La mise en évidence du voisinage de l'objet que nous voulons sélectionner s'opère en créant une sphère délimitée. Cette dernière permettra de définir l'espace de sélection (différents objets qui risquent d'être sélectionnés).

Tous les objets appartenant à cette sphère sont préalablement sélectionnés (Voir Figure 3.12.a). A l'aide d'un geste de la main de l'utilisateur, l'objet le plus accessible sera sélectionné. Si ce n'est pas l'objet souhaité, alors on déplacera légèrement la main de l'utilisateur à droite, à gauche, en haut ou même en bas pour sélectionner le bon objet (Voir Figure 3.12.b).

Si l'objet que l'on veut sélectionner est loin et qu'il n'appartient pas à cet entourage sélectionné, nous dépassons l'espace de la sphère et là un nouveau voisinage s'affiche et une nouvelle sphère est créée.

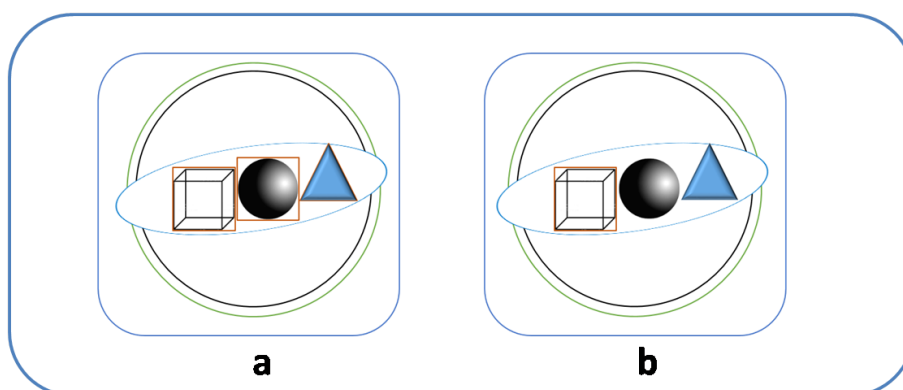


FIGURE 3.12 – principe de fonctionnement de la sélection précise : a : la sélection de tous les objets appartenant au même entourage, b : sélection de l'objet souhaité.

La figure 3.13 montre un exemple où nous avons un ensemble de crayons dans un gobelet, où l'utilisateur doit faire une sélection de l'un d'entre eux.

Cette sélection doit se faire d'une manière précise. Le crayon que l'on désire sélectionner est très fin et se trouve dans un environnement encombré collé aux autres crayons.

Une sphère est alors créée et englobe les différents crayons et le gobelet. Le premier crayon touché (Voir 3.13.a) par la main sera sélectionné (mis en avant en l'entourant en couleur orange), si ce n'est pas le bon crayon, on déplace la main de l'utilisateur vers la droite ou vers la gauche (Voir Figure 3.13.a et Figure 3.13.b) ou vers le bas si on veut sélectionner le gobelet (Voir Figure 3.13.c). Ensuite, on sélectionne le crayon souhaité.

Nous avons remarqué que cette technique permet à l'utilisateur une sélection très précise dans un environnement très encombré et avec différentes formes d'objets.

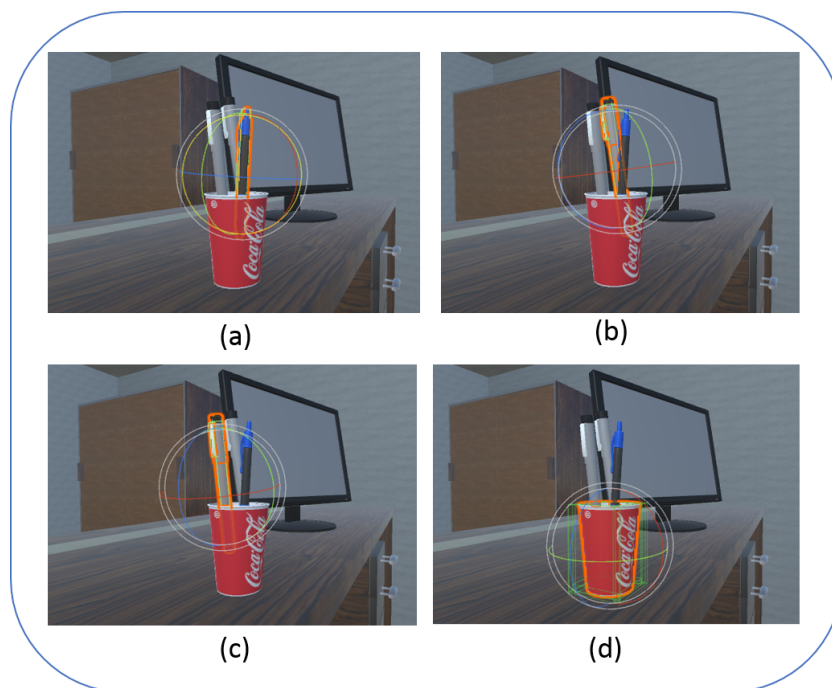


FIGURE 3.13 – Illustration de la partie sélection précise : a) sélection du premier crayon, b) sélection du deuxième crayon, c) sélection du troisième crayon, d) sélection du gobelet [3].

f) Sélection occultée

Cette partie est activée quand l'objet qu'on souhaite sélectionner se trouve occulté partiellement ou totalement. La sélection ordinaire dans ce cas-là est quasiment impossible surtout dans le cas où l'objet est totalement occulté.

La sélection des objets occultés est un problème qui reste d'actualité et qui n'a pas encore été résolu, ce problème rend la sélection très lente, voire impossible. La sélection occultée prédira s'il y a un objet caché ou occulté derrière un objet présent dans notre environnement virtuel.

La technique proposée "*Zoom-fwd*" offre une solution pour permettre de sélectionner des objets cachés, cette fonctionnalité va nous permettre d'afficher la posture ou la forme d'un objet occulté derrière chaque objet touché par l'utilisateur s'il existe (si derrière l'objet il n'y a rien alors il n'affiche rien).

L'intérêt de cette fonctionnalité est de permettre à l'utilisateur d'avoir la position exacte de l'objet caché et d'effectuer une sélection rapide. Ce dernier peut être occulté par un objet plus grand ou simplement ne peut pas être vu directement par l'utilisateur (le point de vue de l'utilisateur et/ou sa position cachent l'objet que nous souhaitons sélectionner).

La Figure 3.14 montre un exemple où nous avons un écran d'ordinateur qui cache complètement un gobelet (l'objet que nous voulons sélectionner). La sélection occultée de la technique "*Zoom-fwd*" permet d'afficher légèrement la forme du gobelet lorsque l'écran est touché par la main de l'utilisateur.

Nous avons remarqué que la sélection occultée permet à l'utilisateur de voir d'abord s'il y a des objets complètement occultés, et ensuite d'avoir leurs positions exactes ce qui nous fait gagner beaucoup de temps en termes de réalisation des tâches d'interaction 3D.

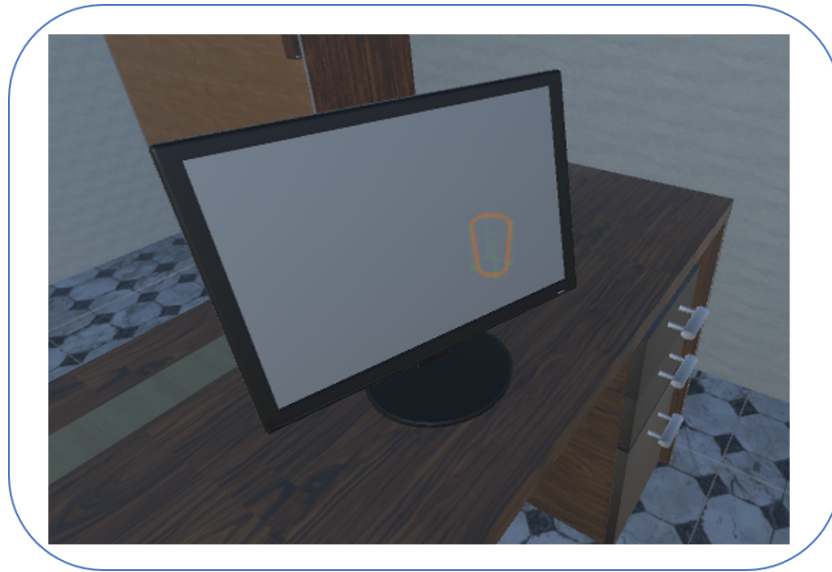


FIGURE 3.14 – Principe de fonctionnement de la partie sélection occultée[3]

Lorsque toutes les manipulations sont terminées, l'utilisateur effectue un geste par sa main (dans ce cas, nous avons choisi le geste Main-Ouverte ou Open Palm) afin de terminer l'opération. Une animation de zoom arrière (dé-zoom) est alors activée pour que la main de l'utilisateur rejoigne sa position initiale.

D'autres gestes sont également utilisés : *pointer*, le geste *L* et *pinch*(pincer).

Le geste *pointer* utilise l'index pour faire la sélection. Le geste *pinch* possède la fonction "attraper l'objet" dans le but de le manipuler par la suite.

3.3.4 Évaluation de la technique "*Zoom-fwd*"

Dans le domaine de la réalité virtuelle et spécialement dans l'interaction 3D, l'évaluation d'une technique d'interaction est une étape primordiale et nécessaire pour valider cette dernière.

Pour évaluer la technique proposée, nous devons impérativement la faire tester par un certain nombre d'utilisateurs et la comparer à une technique d'interaction existante dans les mêmes conditions.

Dans notre cas, nous avons jugé que parmi toutes les techniques d'interaction 3D (rappelées au chapitre 1 Section 1.) la technique "*HOMER*" est la plus appropriée pour permettre une sélection précise et une interaction avec les objets distants et occultés. De plus cette technique est facile et intuitive pour les utilisateurs.

Les deux techniques "*Zoom-fwd*" et "*HOMER*" seront implémentées et testées dans les mêmes conditions (matériel, logiciel, utilisateurs, etc). Pour l'évaluation plusieurs étapes sont nécessaires.

3.3.4.1 Hypothèse de recherche

La première étape du processus d'évaluation c'est de définir l'hypothèse de recherche. Notre principale question est de savoir si la technique proposée améliore l'interaction 3D dans un EV. De plus, nous voulons savoir si la technique d'interaction 3D proposée influencera également l'avis subjectif des utilisateurs en termes d'utilité et de satisfaction. Par conséquent, l'hypothèse principale de notre expérience est :

H_1 : Les utilisateurs effectueront mieux les tâches de sélection-manipulation en utilisant la technique "*Zoom-fwd*".

Nous pensons que la technique "*Zoom-fwd*" permettra aux utilisateurs d'être plus rapides et plus précis dans l'exécution des tâches de sélection, car elle leur permettra d'effectuer les tâches de sélection dans les plus brefs délais. Pour étudier cette hypothèse, nous avons mené l'étude suivante auprès des utilisateurs.

3.3.4.2 Le protocole expérimental

- **Les participants** Nous avons pris un échantillon de quinze participants ($N = 15$), 8 hommes et 7 femmes, étudiants et collègues qui se sont été inscrits volontairement à cette étude. Ce sont tous des utilisateurs ayant une expérience limitée des outils de RV (Kinect, Leap-Motion et l'Oculus Rift).

L'âge moyen est de 34,5 ans, douze d'entre eux sont droitiers, tous les participants ont une vision normale ou corrigée à la normale. Neuf d'entre eux avaient une expérience antérieure des jeux vidéo (y compris sur smartphones), dont six jouaient régulièrement à des jeux vidéo. Douze d'entre eux ont déclaré avoir déjà fait l'expérience de la RV avant cette expérience.

- **Conception du protocole expérimental**

Tous les participants ont effectué la tâche expérimentale deux fois, une fois en utilisant chaque technique d'interaction. Afin de contrebalancer les effets d'apprentissage deux groupes d'utilisateurs ont été identifiés : le premier groupe de huit personnes ($N_1=8$) commencera par la technique "*HOMER*", le second groupe de sept personnes ($N_2=7$) commencera par la technique "*Zoom-fwd*".

Les observations faites étant appariées, le plan d'expérience est donc du type Within-subject (Inter-Sujets) avec une variable indépendante (technique d'interaction).

- **Le système logiciel**

Notre système utilise Leap-Motion SDK 2.0 qui fonctionne sur un processeur i7-2GHz Intel® Core avec une RAM de 6Go. L'EV qui est un bureau virtuel a été modélisé avec Blender version 2, le tout est développé sous Unity3D (version 5.7) avec des scripts en C#.

- **Les tâches expérimentales**

Pour évaluer et tester la technique de sélection et de manipulation "Zoom-fwd", nous avons défini des scénarios de sélection et de manipulation.

Nous avons défini trois scénarios de sélection de différents objets, à des emplacements différents. Ces tâches sont plus ou moins complexes, elles ont été choisies de manière à évaluer les différentes fonctionnalités de la technique "Zoom-fwd" et de voir à travers ces scénarios si les défis lancés ont été atteints ou pas. Ensuite un scénario de manipulation a été mis en place afin d'évaluer la manipulation de la technique "Zoom-fwd".

Les participants doivent effectuer ces tâches ou ces scénarios en utilisant les deux techniques d'interaction 3D. Le but à atteindre par les participants est d'effectuer ces scénarios en un minimum de temps pour chaque technique.

Scénarios de sélection :

Tâche T1 : *Sélectionner* une tablette sur le bureau (Voir Figure 3.15).

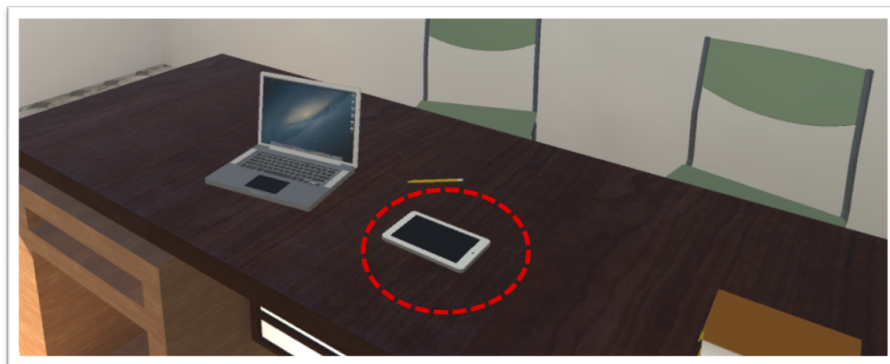


FIGURE 3.15 – Premier Scénario de sélection : sélection de la tablette [3]

Tâche T2 : *Sélectionner* un livre rouge sur l'étagère (Voir Figure 3.16).

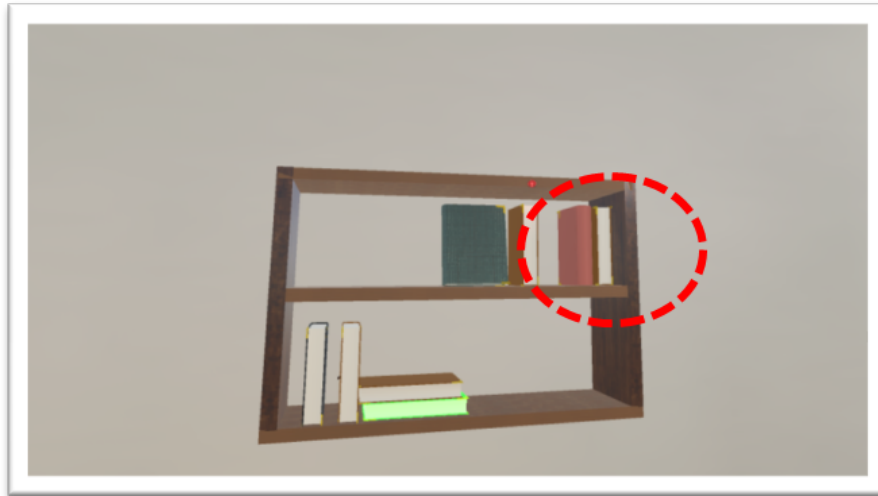


FIGURE 3.16 – Deuxième scénario de sélection : sélection du livre [3]

Tâche T3 : *Sélectionner* une statuette cachée (Voir Figure 3.17).

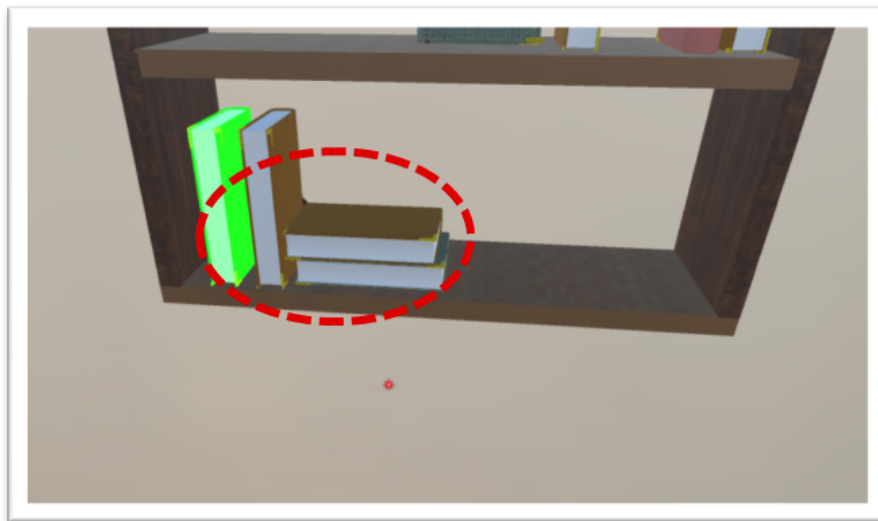


FIGURE 3.17 – Troisième scénario de sélection : sélection d'une statuette [3]

Scénario de manipulation :

Tâche T4 : *Déplacer* un livre de l'étagère au bureau et le poser sur un autre.

Comme c'est expliqué dans la section 3.3.4.2, les participants vont être divisé en deux groupes, ces derniers effectueront les tâches de sélection dans des ordres différents.

Le premier groupe (Sept participants) a été chargé d'effectuer les tâches T1, T2, T3 et T4 respectivement : en sélectionnant la tablette sur le bureau, le livre rouge sur l'étagère, la statuette cachée de déplacer le livre de l'étagère vers le bureau. Ce premier groupe commencera par la technique "*HOMER*" et enchainera par la technique "*Zoom-fwd*".

Le deuxième groupe (huit participants) est chargé quant à lui d'effectuer les tâches T2, T1, T3 et T4 respectivement : en sélectionnant d'abord un livre rouge sur l'étagère, la tablette sur le bureau, et la statuette cachée et enfin de déplacer le livre de l'étagère. Ce second groupe commencera par la technique "*Zoom-fwd*" et la technique "*HOMER*" par la suite.

Cette permutation de tâches dans les groupes évitera aux participants l'effet de l'apprentissage des différentes tâches. T1 est une tâche relativement simple, elle consiste à sélectionner une tablette sur le bureau qui est visible et accessible ; le but de T1 est de permettre aux utilisateurs de se familiariser avec l'environnement virtuel, le casque et les différents gestes.

T2 est plus complexe, les utilisateurs devront sélectionner un livre rouge sur l'étagère, se trouvant au milieu de plusieurs livres ; sa sélection nécessite beaucoup de précision. Le but de T2 est de permettre aux participants d'exercer une sélection dans un environnement encombré (plusieurs livres les uns à côté des autres). Cette sélection doit être précise.

T3 est encore plus complexe. Les utilisateurs sont chargés de sélectionner une statuette sur l'étagère, qui est cachée par les différents livres. La sélection de la statuette permet de faire face à un autre défi de la technique "*Zoom-fwd*" qui est l'occultation.

T4 est une tâche de manipulation. Les utilisateurs sont chargés de sélectionner un livre rouge sur l'étagère et de le poser sur un autre livre situé sur le bureau. Les participants n'ont pas de connaissance à priori sur la difficulté des différentes tâches, ni l'emplacement des différents objets à sélectionner ou à manipuler.

- **La procédure expérimentale**

La procédure expérimentale débute par les réponses au questionnaire démographique, contient toutes les informations concernant chaque participant (âge, droitier/gaucher..(Voir Annexe A, section A.1)). Un tutoriel détaillé consistant en une série d'étapes sous forme de courtes instructions et d'explications est présenté aux participants. Cette phase de l'expérimentation est accompagnée par des courtes phases de manipulation pour se familiariser avec le dispositif.

Toutes ces étapes peuvent être répétées si nécessaire pour permettre de comprendre progressivement chaque étape de la tâche à effectuer. Une fois le tutoriel terminé, la session expérimentale commence pour le premier groupe avec la première technique. Chaque tâche doit être effectuée deux fois pour chaque technique d'interaction. A la fin de l'expérimentation chaque participant est invité à remplir un questionnaire afin d'évaluer subjectivement les deux techniques d'interaction pour ces scénarios.

Un questionnaire de comparaison et un questionnaire d'utilisabilité du système global ont été remplis par chaque participant.

3.3.4.3 Mesures et analyse statistique des données

Pour comparer les deux techniques, des mesures objectives et subjectives sont faites. Toutes les analyses statistiques de ces mesures sont effectuées à l'aide du logiciel R (R Core Team, 2017) dans sa version 3.3.3. Les tests statistiques ont tous été effectués avec un

niveau de confiance de 95/Par conséquent, les résultats sont considérés significatifs lorsque la p-value du test (p) est inférieure 0.05 ($p < 0.005$).

TABLE 3.1 – Test de normalité

Variables	Test de Shapiro	
	W	p-values
T_1 : HOMER	0.85556	0.02084
T_1 : Zoom-fwd	0.89363	0.07611
T_2 : HOMER	0.91762	0.1772
T_2 : Zoom-fwd	0.86227	0.02603
T_3 : HOMER	0.87642	0.04197
T_3 : Zoom-fwd	0.97549	0.9294
T_4 : HOMER	0.79077	0.002806
T_4 : Zoom-fwd	0.89907	0.0921

- **Évaluation objective**

La performance du participant a été évaluée à travers le temps d'exécution des différentes tâches de sélection et de manipulation. Le calcul du temps pour chaque tâche a commencé une fois que l'utilisateur utilise le geste de navigation.

Premièrement, les données collectées ont été analysées pour déterminer le type de tests qu'on doit utiliser : paramétriques ou non paramétriques. Nous avons vérifié l'hypothèse de normalité des mesures à l'aide du test de Shapiro-Wilks. Les résultats indiquent que toutes les données ne suivent pas une distribution normale (voir la table 3.1), ce qui ne permet pas d'utiliser les tests paramétriques. Par conséquent, le test non paramétrique des rangs signés de Wilcoxon a été utilisé pour comparer les valeurs moyennes. Ce dernier montre un effet significatif de la technique sur les temps moyens de réalisation de toutes les tâches. Les statistiques du test ainsi que les p-values sont données dans la table 3.2. Elles montrent une différence significative des temps d'exécution moyens en faveur de la technique "Zoom-fwd".

Pour la première tâche de sélection [$V = 120$, p-value = 0.0007247], cette dernière a été effectuée plus rapidement avec la technique "Zoom-fwd" (69% moins de temps).

Pour effectuer la seconde tâche [$V = 120$, p-value = 0.0007229], les participants ont passés moins de temps en utilisant la technique "Zoom-fwd" (71% moins de temps).

TABLE 3.2 – Analyses descriptives et statistiques pour les données objectives [3]

Les tâches	Temps moyens		Test de Wilcoxon	
	HOMER(s)	Zoom-fwd(s)	V	p-values
Sélection de la tablette	7.71	2.23	120	0.0007247
Sélection du livre	9.25	2.66	120	0.0007229
Sélection de la statuette	3.07	0.75	120	0.0007229
Manipulation du livre	3.64	1.21	104	0.001367

Les participants ont accompli la troisième tâche de sélection [$V = 120$, p-value = .0007229] plus rapidement avec la technique "Zoom-fwd" (76% moins de temps). Les utilisateurs ont mis moins de temps pour manipuler l'objet [$V = 104$, p-value = .001367] et ils ont mis 67% moins de temps en utilisant la technique proposée "Zoom-fwd". Les temps moyens d'exécution des différentes tâches sont montrés dans la table 3.2. Les participants ont mis beaucoup plus de temps pour réaliser T1 en utilisant la technique "HOMER" (7.17 s) que la technique proposée (2.23 s).

Le temps d'exécution de T2 est nettement inférieure avec la technique "Zoom-fwd" (2.66 s) qu'avec la technique "HOMER" (9.25 s).

Pour T3 les temps sont respectivement de (0.75 s) "Zoom-fwd" et de (3.07 s) "HOMER".

Pour T4, le temps mis par les participants à utiliser la technique "Zoom-fwd" (3.64 s) est inférieur au temps mis à utiliser la technique "HOMER" (1.21 s).

Nous avons remarqué une diminution importante des temps d'exécution des deux dernières tâches, et ce pour les deux techniques, qui s'explique par l'effet de l'apprentissage des gestes de la main et des techniques par les participants.

La Figure 3.18 (Boxplots) illustre toutes ces conclusions. Elle montre de plus la plus grande homogénéité des temps d'exécution, aussi bien pour les différentes tâches que pour les différents participants de la méthode "Zoom-fwd".

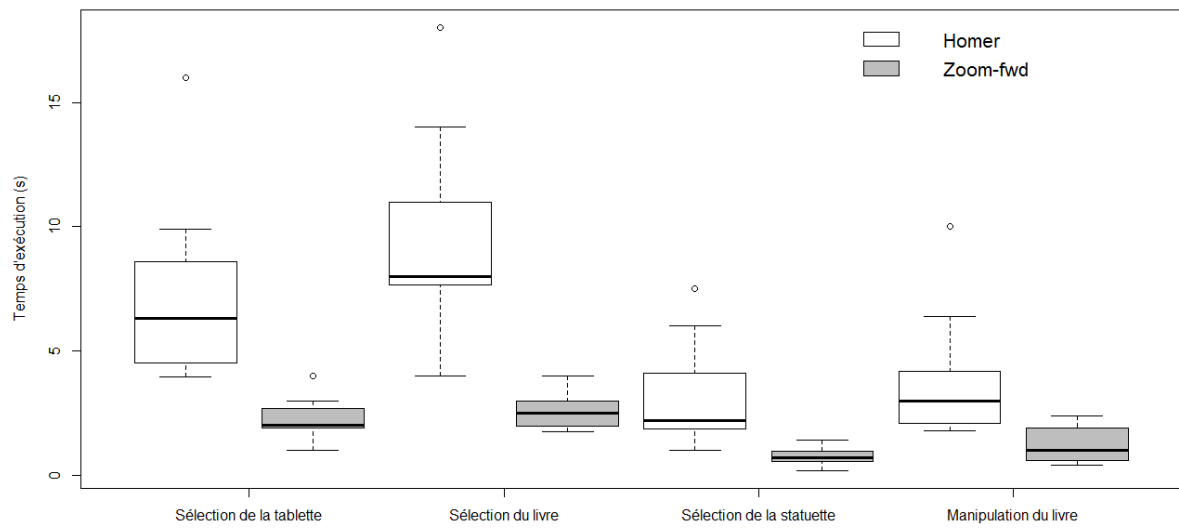


FIGURE 3.18 – Le temps d'exécution des différentes tâches d'interaction [3]

Nous avons également étudié l'influence de la complexité de la tâche par rapport au temps moyen d'exécution.

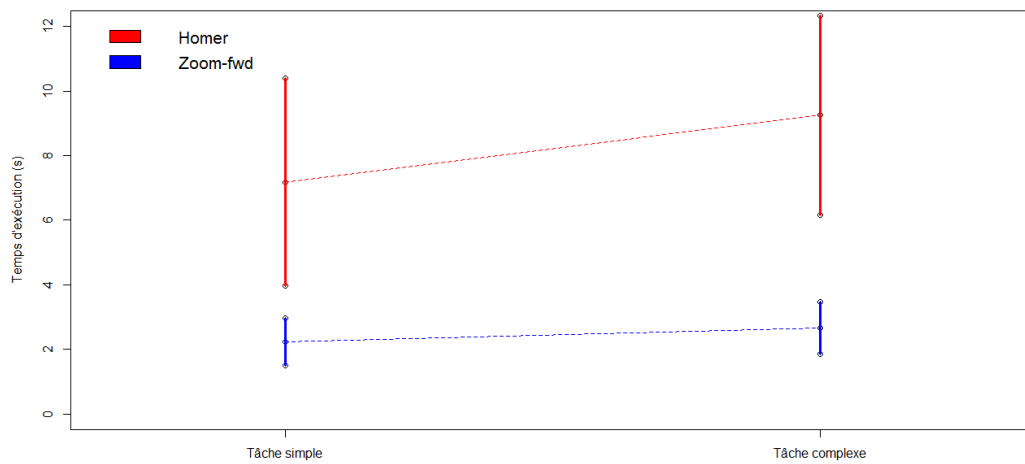


FIGURE 3.19 – Comparaison entre une tâche complexe et une tâche simple : premier cas

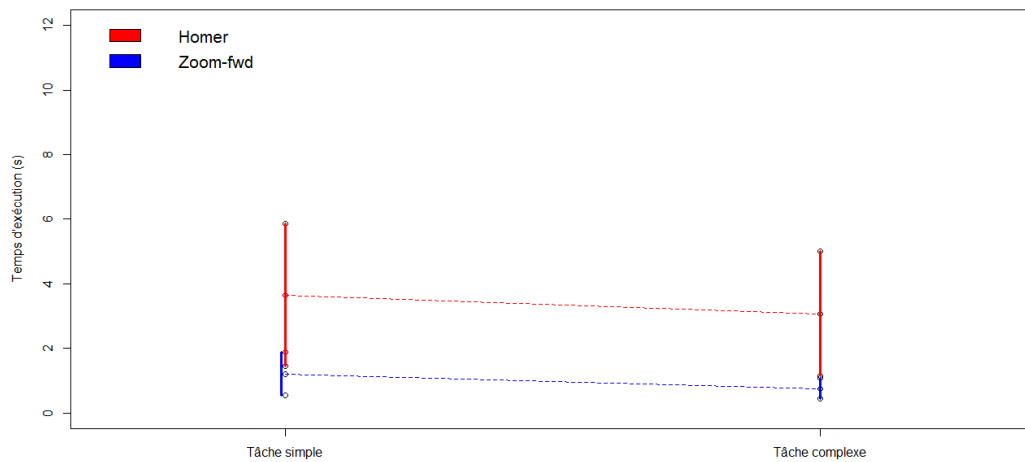


FIGURE 3.20 – Comparaison entre une tâche complexe et une tâche simple : deuxième cas

Le graphe des moyennes de temps illustrés dans les figures 3.19 et 3.20, montre que la technique "*Zoom-fwd*" nécessite beaucoup moins de temps pour ces deux types de tâches simples et complexes.

On relève un effet significatif de la complexité de la tâche sur le paramètre temps d'exécution. Nous avons également remarqué qu'entre les différents participants le temps d'exécution avec la technique "*Zoom-fwd*" varie uniformément qu'avec la technique "*HOMER*" ce qui signifie que la technique "*Zoom-fwd*" est une technique plus stable.

- **Évaluation subjective**

Une fois les expériences terminées, nous avons demandé aux participants de remplir deux types de questionnaires, afin d'évaluer notre technique et de la comparer avec la technique "*HOMER*", d'un point de vue subjectif (avis de l'utilisateur).

TABLE 3.3 – Les questions post-expérimental et les moyennes des notes données pour chaque technique

Catégorie	ID	Question	Valeurs moyennes		Test de Wilcoxon	
			Homer	Zoom-fwd	V	P-values
Utilité	Q1	La technique est-elle-utile ?	2.46	3.66	6	0.01621
	Q2	La technique me fait elle-gagner du temps ?	3.00	4.20	2.5	0.006746
	Q3	La technique répond -elle- à mes attentes ?	2.73	4.66	0	0.0005644
Facilité d'apprentissage	Q4	La technique est-elle-facile à utiliser ?	3.66	3.46	6	0.1489
	Q5	La technique est-elle-rapide à apprendre ?	4.20	4.06	27	0.6078
Satisfaction	Q6	Ses étapes sont-elles-faciles à remémorer ?	3.46	3.06	24.5	0.07076
	Q7	La technique est-elle-nécessaire ?	3.73	3.86	13.5	0.5297
	Q8	La technique est-elle-satisfaisante ?	2.60	4.20	0	0.0008042
	Q9	La technique est-elle agréable à l'utiliser ?	2.33	3.66	0	0.001212

1. Questionnaire "USE"

Le premier questionnaire concerne l'utilisabilité du système. Pour cela, nous avons utilisé le questionnaire "USE" (Usefulness, Satisfaction and Ease) [150]. Ce dernier est largement utilisé pour l'évaluation des systèmes de réalité augmentée/virtuelle et de l'interaction 3D. Il comporte 30 questions regroupées en 4 catégories : l'utilité (Usefulness), la satisfaction (Satisfaction), la facilité de l'utilisation (Ease of Use) et la facilité de l'apprentissage (Ease of Learning). Le participant peut répondre aux questions à travers une échelle de Likert de 5 points, allant de "1 : fortement en désaccord" à "5 : fortement d'accord".

Une version française de ce questionnaire peut être trouvée en Annexe (Annexe A, section A.2.) Les questions comprenaient différents critères : utilité, satisfaction et apprentissage (difficulté d'utiliser la technique).

Les questions (Q1-Q9) (Voir Tableau 3.3) sont extraites de ce dernier questionnaire USE [150]. Une fois les questionnaires remplis par les participants, nous avons calculé la moyenne de chaque catégorie du questionnaire USE. Les résultats obtenus sont illustrés par les graphes de la figure(Figure 3.21).

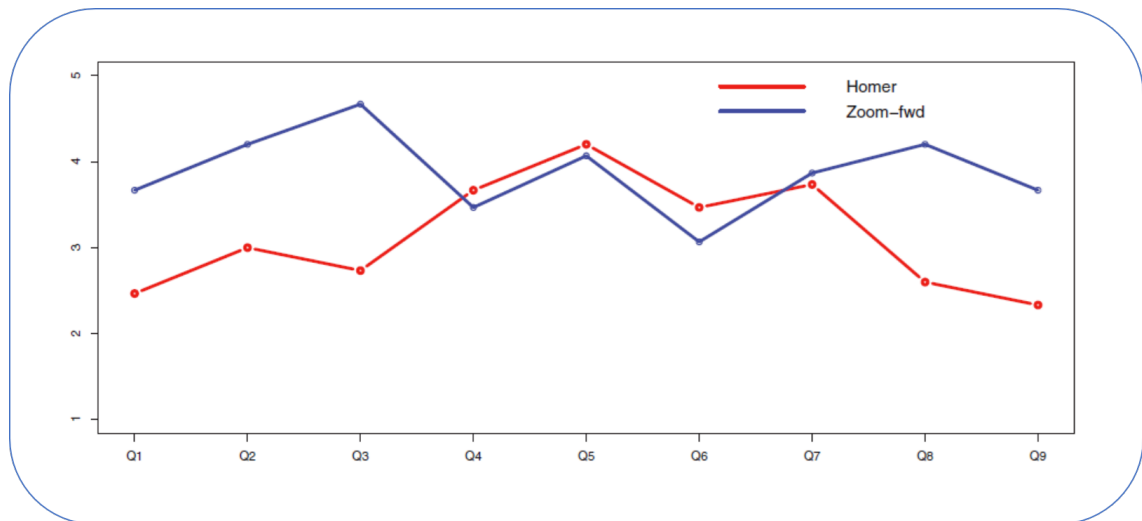


FIGURE 3.21 – Scores pour chaque question [3]

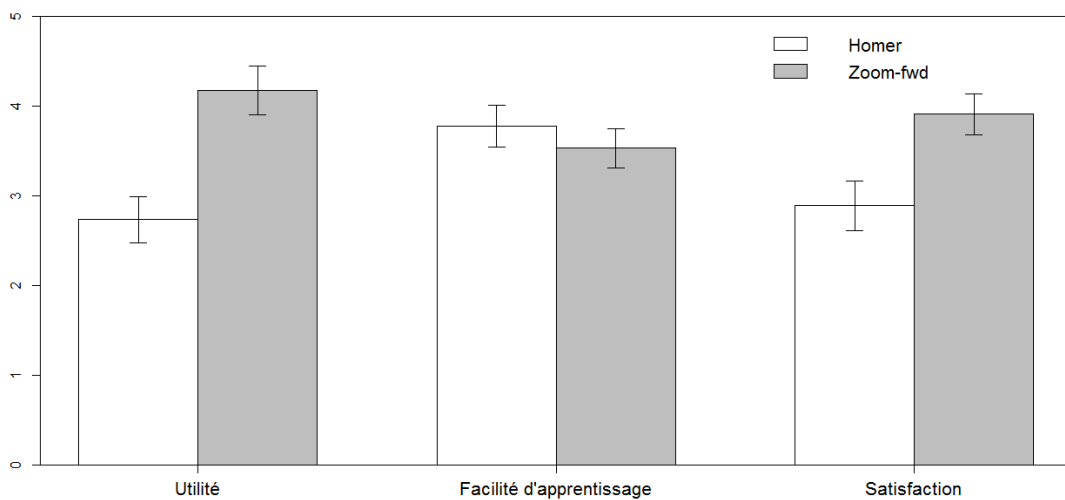


FIGURE 3.22 – Résultats subjectifs groupés par critères [3]

2. Questions de comparaison

Nous avons aussi posé des questions aux participants et on a leur demandé d'indiquer quelle technique était plus facile à utiliser, avait une meilleure performance, laquelle était spécifiquement préférée pour effectuer toutes les tâches.

3. Questionnaire SUS

Enfin, un troisième questionnaire a été posé aux participants, ce dernier concerne l'utilisabilité du système. Le questionnaire SUS (System Usability Scale)[151], est un questionnaire normalisé se composant essentiellement de 10 questions, il a été proposé afin de déterminer le niveau de satisfaction des utilisateurs d'un service (un logiciel, une application, un site internet, etc.) ou d'un système (système de RV, Système de RA, etc.).

L'analyse des questions groupées par les tests des rangs signés de Wilcoxon montre un effet significatif de la technique "*Zoom-fwd*" pour les critères : utilité et satisfaction (voir 3.22), mais le même test montre un effet non significatif de cette technique sur le critère de facilité de l'apprentissage. Nous avons remarqué que la plupart des participants ont apprécié l'utilité de la technique "*Zoom-fwd*" (4.17) par rapport à la technique "*HOMER*" (2.73). Cela est dû à la précision qu'offre la technique "*Zoom-fwd*" lors de la sélection et de la manipulation des objets distants et dans un milieu encombré par rapport à la technique "*HOMER*". En revanche, les participants ont estimé que la technique "*HOMER*" est légèrement plus facile à apprendre (3.77) par rapport à la technique "*Zoom-fwd*" (3.52). Quant au critère de la satisfaction, la technique "*Zoom-fwd*" présente une moyenne de (3.91) pour une moyenne de (2.88) pour la technique "*HOMER*".

Pour les questions de comparaison, les résultats ont montré que 62% des participants ont trouvé la technique "*Zoom-fwd*" plus facile à utiliser, 55% des participants préfèrent utiliser la technique "*Zoom-fwd*" et sont plus performants avec elle et préfèrent terminer toutes les tâches avec cette technique.

Grâce à l'étude de Bangor [152], il est ainsi possible de déterminer, à partir du score au SUS, si un système est bon, excellent, acceptable ou mauvais (Voir Figure 3.23). Le score SUS indique une valeur moyenne de 70,22%, ce qui correspond à une appréciation "Acceptable" mais qui est plutôt proche de l'appréciation "Bonne".

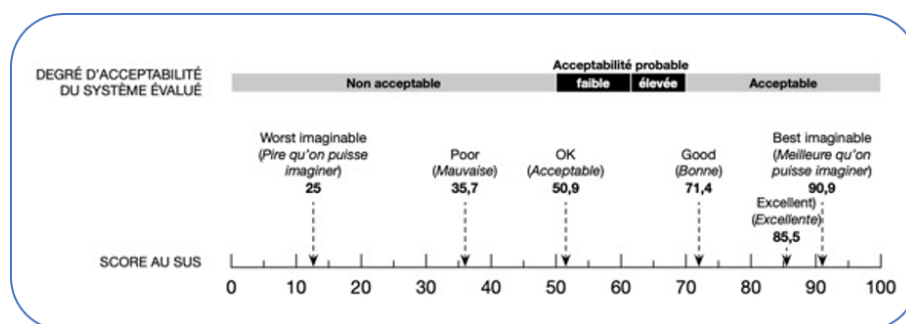


FIGURE 3.23 – Signification du score au SUS [152]

3.4 Discussion

Dans cette section, on va discuter les résultats de notre étude expérimentale. Les résultats de l'étude des mesures objectives et subjectives montrent que l'utilisation de la technique "Zoom-fwd" facilite et améliore la sélection et la manipulation des objets dans l'environnement virtuel. Ceci est montré par le temps moyen mis par les utilisateurs à effectuer chaque tâche, le temps d'exécution de ces dernières est nettement plus réduit avec la technique proposée "Zoom-fwd" qu'avec la technique "HOMER".

Pour réaliser la première tâche (T1) qui est une tâche relativement simple, les participants ont mis beaucoup moins de temps en utilisant la technique proposée que la technique existante.

Pour la deuxième tâche (T2) qui consiste à sélectionner un livre se trouvant au milieu des autres dans un environnement encombré, cette tâche nécessite beaucoup de précision.

La sélection du livre doit activer la partie « sélection précise » de la technique "*Zoom-fwd*" ce qui a fait gagner beaucoup de temps aux différents utilisateurs. Cela nous montre que la technique "*Zoom-fwd*" est plus recommandée lorsque la sélection et/ou manipulation précise est nécessaire dans un environnement encombré ou non. La troisième tâche (T3) étant la tâche la plus complexe, cette dernière consiste en la sélection d'une statuette complètement caché par d'autres objets. Cette tâche a été effectuée par la partie « sélection occulté » de la technique "*Zoom-fwd*", une fois de plus le temps d'exécution de T3 en utilisant la technique proposée est inférieur à celui utilisé par la technique déjà existante. Ce qui nous mène à conclure que la technique "*Zoom-fwd*" est plus adaptée à ce type de situations (objets complètement ou partiellement cachés ou occultés).

Pour la tâche de manipulation(T4), le temps mis par les participants en utilisant la technique "*Zoom-fwd*" est également plus court que le temps passé en utilisant la technique "*HOMER*".

En outre, les résultats montrent que le temps de réalisation des différentes tâches en utilisant la technique proposée est très proche pour tous les utilisateurs, ce qui signifie que la technique "*Zoom-fwd*" est une technique stable, qui s'adapte aux différents scénarios et qui nous donne des résultats plus ou moins proches pour tous les utilisateurs. D'après ces résultats, on a tiré des observations :

- La technique "*Zoom-fwd*" est une technique qui permet une interaction rapide avec les différents types d'objets et s'adapte aux différents scénarios ;

- La technique "*Zoom-fwd*" apporte des solutions à différents défis rencontrés dans le domaine des interactions 3D tels que :
 1. interaction avec des objets petits et distants, en introduisant la zoom-sélection ;
 2. interaction précise et exacte dans un environnement encombré, en introduisant la sélection-précise ;
 3. interaction avec des objets complètement ou partiellement cachés ou occultés, en utilisant la sélection-occultée.

- les résultats des questionnaires montrent que 62% des participants ont trouvé que la technique "*Zoom-fwd*" est une technique facile à utiliser et 55% d'entre eux préférant l'utiliser que d'utiliser la technique "*HOMER*". Cela permet de conclure que la technique "*Zoom-fwd*" est plus facile à utiliser.

La facilité d'utilisation et d'apprentissage des deux techniques provient de l'utilisation de la reconnaissance des gestes. Ceci, à son tour, permet de conclure que la technique "*Zoom-fwd*" offre une manière naturelle, intuitive et rapide d'interagir avec l'environnement virtuel.

Par conséquent, H1 (Les utilisateurs effectueront mieux les tâches de sélection-manipulation en utilisant la technique "*Zoom-fwd*") est validée par l'étude expérimentale.

Enfin, en termes d'utilisabilité, le score SUS rapporte une valeur moyenne de 70,22%, ce qui encourage les choix faits pour la conception de cette technique d'interaction 3D, à savoir la technique "*Zoom-fwd*".

3.5 Conclusion

Nous avons présenté dans ce chapitre les différentes étapes entreprises pour concevoir une nouvelle technique d'interaction 3D appelée "*Zoom-fwd*". Cette technique utilise

la reconnaissance de gestes pour les tâches d'interaction 3D (navigation, sélection, manipulation et contrôle d'applications).

La technique "*Zoom-fwd*" permet une interaction rapide et précise avec des objets distants et occultés. Notre objectif principal était de proposer une technique d'interaction 3D qui tentera d'éviter les problèmes rencontrés en utilisant les différentes interfaces physiques. De plus, la technique proposée utilise la reconnaissance des gestes, ce qui la rend très naturelle et facile à utiliser.

Les résultats de notre étude expérimentale montrent que la technique proposée offre une meilleure performance dans l'accomplissement des tâches. Cette augmentation de performance a été plus particulièrement observée lors des tâches qui nécessitent de la précision et sur des objets occultés (sélection et manipulation de livres, sélection de statuettes).

Il a été conclu que ce travail a une valeur ajoutée dans le domaine de l'interaction 3D, en offrant une nouvelle technique d'interaction 3D qui permet de gagner du temps et de la précision.

Dans le prochain chapitre, nous allons présenter ce que nous avons réalisé dans la deuxième partie de notre travail et qui concerne le travail collaboratif ou plus précisément l'interaction 3D collaborative dans un environnement virtuel partagé.

Chapitre

4

Système 3D collaboratif

4.1 Introduction

Nous avons présenté dans le chapitre précédent les différentes contributions relatives à l'interaction 3D en RV en proposant une nouvelle technique d'interaction 3D "*Zoom-fwd*". Cette technique utilise la reconnaissance de gestes pour les tâches d'interaction 3D (navigation, sélection, manipulation et contrôle d'applications).

Comme déjà mentionné auparavant, les Environnements Virtuels Collaboratifs (EVCs) sont des environnements complexes dans lesquels plusieurs utilisateurs partagent les mêmes objets pour agir ensemble. La complexité de ces environnements est liée à l'interaction du groupe d'utilisateurs avec les éléments partagées.

Dans un monde partagé, plusieurs contraintes apparaissent, notamment celles qui sont liées à la coordination et à la communication des utilisateurs ainsi qu'à la gestion de l'interaction d'un utilisateur dans le monde virtuel vis à vis des objets et des autres utilisateurs. Malheureusement, comme nous l'avons vu dans le premier chapitre, les interactions 3D mono et multi-utilisateur(s) en RV sont actuellement loin de procurer des solutions satisfaisantes. En effet, l'interaction 3D souffre d'un manque de modèles et de formalismes permettant de gérer et de contrôler les actions et les intentions des utilisateurs dans l'environnement virtuel.

Un autre aspect à prendre en considération est l'hétérogénéité des plateformes de RV entre les participants ainsi que la disparité entre les utilisateurs afin de proposer des modèles et des systèmes permettant le travail collaboratif au sein d'un EVC.

Dans ce qui suit, nous présenterons la deuxième partie de nos travaux concernant l'interaction 3D collaborative en RV. Nous présenterons un formalisme d'interaction 3D collaborative basé sur des modèles existants, ce formalisme permettra de générer les actions d'un groupe d'utilisateurs. Ce dernier facilitera la coordination des interactions mono-utilisateurs pour permettre aux différents utilisateurs de partager un environnement et des objets le constituant. Nous présenterons par la suite une proposition d'implémentation informatique et nous terminerons par les résultats obtenus.

4.2 Système 3D collaboratif

4.2.1 Modélisation de la collaboration

4.2.1.1 Modèle des "3C"

Pour modéliser un environnement virtuel collaboratif une panoplie de modèles et de formalismes existent en littérature [153], nous nous sommes basés sur le modèle "3C" [154]. Ces derniers ont proposé trois fonctionnalités que chaque EVC devra couvrir : Communication, Collaboration et Coordination d'où l'appellation "3C". La collaboration dans [154] est une activité qui peut être synchrone ou asynchrone, au même lieu ou dans des lieux différents. Comme nous pouvons le voir sur la Figure 4.1.

- L'espace de communication concerne la communication entre les individus. Par exemple, le courrier électronique et les média-spaces (systèmes conçus pour supporter la communication synchrone par ordinateurs interposés).
- L'espace de coordination codifie les protocoles qui gouvernent la façon dont les tâches sont gérées par les groupes d'utilisateurs. Ces protocoles peuvent être purement sociaux, comme par exemple la règle sociale qui spécifie qu'une seule personne doit parler à la fois, ou peuvent être formellement spécifiés à travers un système de workflow.
- L'espace de collaboration comprend les artefacts partagés qui sont manipulés de façon coopérative pour réaliser les différentes tâches, comme par exemple les documents ou les données partagées.

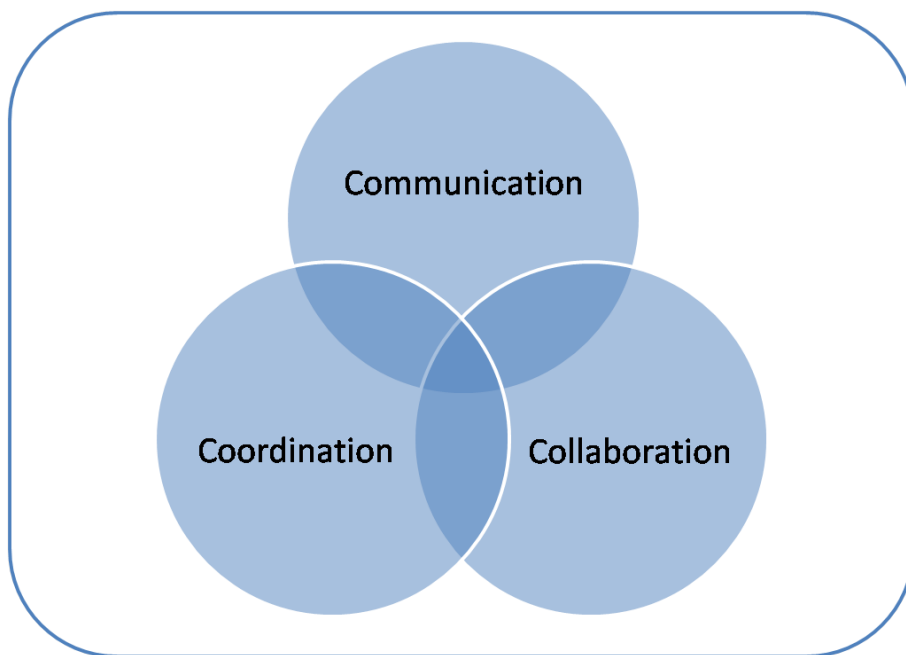


FIGURE 4.1 – Le modèle 3C par Ellis [154]

Dans ce qui suit, nous présentons un modèle d'EVC qui basé essentiellement sur le modèle des "3C" [154], ayant la particularité d'avoir une couche supplémentaire qui est "Interaction 3D Collaborative" (Voir Figure 4.1).

4.2.1.2 Modèle 3D collaboratif

Ce modèle séparera l'aspect interaction des autres. L'interaction 3D collaborative est un module qui va pouvoir fournir un formalisme de collaboration entre les différents utilisateurs qui partagent un environnement tout en assurant une interaction 3D collaborative cohérente.

Le formalisme de collaboration doit être suffisamment générique pour pouvoir interfacier d'autres applications développées avec des standards différentes (nombre d'utilisateurs, outils d'interaction, environnement partagé, etc) (Voir Figure 4.2

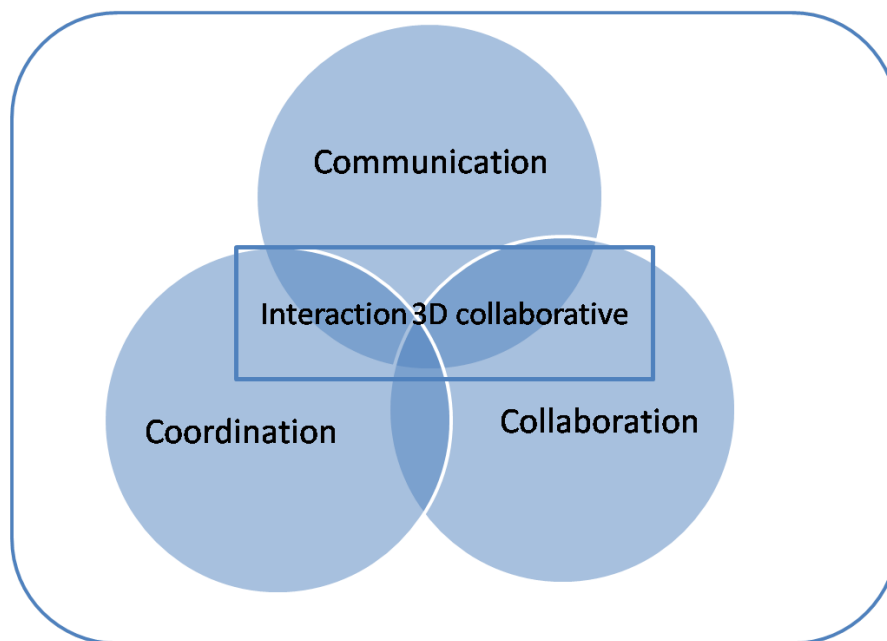


FIGURE 4.2 – Le modèle représentant un EVC avec une interaction 3D collaborative

— Communication

La collaboration nécessite un moyen de communication entre les différents utilisateurs, la communication se repose sur un échange de différentes informations (différentes tâches à entreprendre, les différents objets 3D ainsi que leur positions et orientations, etc.) La communication peut être verbale, par message selon le type d'application et le type de scénario.

- **Coordination** Après avoir échanger entre eux , les utilisateurs mettent en place un plan d'actions a entreprendre pour permettre une collaboration cohérente (les techniques d'interactions 3D des différents utilisateurs, etc.).
- **Collaboration** On retrouve dans ce module les données a partager entre les différents utilisateurs afin d'effectuer les différentes tâches .
- **Interaction 3D collaborative** Ce module regroupera tous ce qui est lié a l'interaction collaborative (les techniques d'interactions 3D du groupe, les résultats du travail collaboratif réalisé sur les données partagés, etc.). Nous allons reprendre la classification des techniques d'interaction 3D présenté lors du premier chapitre (classification en quatre tâches), pour représenter l'interaction collaborative on définira les mêmes tâches (Navigation collaborative, sélection collaborative, manipulation collaborative et contrôle d'application).

1. Navigation 3D collaborative

Dans un environnement collaboratif chaque utilisateur utilise une technique de navigation pour se déplacer dans l'environnement virtuel partagé ; la navigation collaborative représente l'action du groupe d'utilisateurs pour se déplacer en groupe. Nous avons donc besoins de définir une technique pour naviguer en groupe.

N_c représente l'ensemble des techniques de navigations collaboratives, N_i représente la technique de navigation du i ème utilisateur, F la fonction

définissant la relation entre les différentes techniques de navigations illustré dans l'équation 4.1.

$$N_c = F(N_{u_i})_{i = 1..n} \quad (4.1)$$

$$N_{u_i} = D_{u_i}(pos, ori), \quad (4.2)$$

$$D = (x, y, z, O_x, O_y, O_z). \quad (4.3)$$

D étant le déplacement 3D de l'utilisateur dans l'environnement virtuel partagé.

2. Sélection Collaborative

Dans un environnement collaboratif chaque utilisateur utilise une technique de sélection pour sélectionner un ou plusieurs objets partagés ; la sélection collaborative va permettre aux groupes d'utilisateurs de sélectionner un ou plusieurs objets partagés qui vont le/les manipuler dans la prochaine étape.

S_c représente l'ensemble des techniques de sélections collaboratives, S_i la technique de sélection du i ème utilisateur et G la fonction définissant la relation entre les différentes techniques de sélections illustrées dans l'équation 4.4.

$$S_c = G(S_{u_i})_{i = 1..n} \quad (4.4)$$

$$S_{u_i} = P_{u_i}(pos, ori) \quad (4.5)$$

$$P = (x, y, z, O_x, O_y, O_z) \quad (4.6)$$

P étant la position 3D de l'objet à sélectionner.

3. Manipulation Collaborative Lors de la manipulation collaborative sur des objets partagés, deux modes de manipulation sont possible à savoir (synchrone et asynchrone), qui ont été bien expliqués dans le premier chapitre. Chaque utilisateur utilise une technique de manipulation 3D pour changer les propriétés (position et /ou la rotation) d'un ou plusieurs objet(s) partagé(s). Dans le cas asynchrone : le système choisit un utilisateur et lui permet de manipuler l'objet partagé en utilisant sa technique d'interaction 3D, à la fin de cette tâche le deuxième utilisateur doit à son tour manipuler l'objet partagés.

Dans le cas synchrone : nous avons deux scénarios, le premier permettra une manipulation simultanée sur le même objet en séparant ses propriétés (degrés de liberté : position et rotation) en les assignant à des différents utilisateurs. Le second scénario va permettre une manipulation simultanée sur les mêmes propriétés de l'objet partagé.

M_c étant l'ensemble des techniques de manipulation collaborative, M_i la technique de manipulation du i ème utilisateur et H étant la fonction définissant la relation entre les différentes techniques de manipulations illustrées dans l'équation ci-dessous 4.7.

$$M_c = H(M_{u_i})_{i = 1, \dots, n} \quad (4.7)$$

$$M_{u_i} = C_{u_i}(pos, ori, pos', ori') \quad (4.8)$$

$$C = (x', y', z', O'_x, O'_y, O'_z) \quad (4.9)$$

C étant la nouvelle position 3D de l'objet manipulé.

Afin de valider ce formalisme, on va prendre l'exemple de deux utilisateurs, ces deux derniers utiliseront la reconnaissance de gestes pour interagir avec l'EVC, ils partageront un environnement et un ensemble d'objets.

4.3 Exemple informatique du système collaboratif 3D

4.3.1 Système d'apprentissage collaboratif pour la maintenance industrielle

4.3.1.1 Conception du Setup

Dans cette section, nous allons présenter notre système collaboratif pour la maintenance industrielle, il comprend deux utilisateurs : le premier aura le rôle de l'expert dans le domaine de maintenance industrielle et le deuxième un technicien. Le scénario de cette application est de permettre à l'expert d'accompagner le technicien dans les différentes tâches de maintenance. Le setup utilisé par les deux utilisateurs est présenté ci-dessous (voir Figure 4.3), une Kinect est fixée au dessus des mains du premier utilisateur(l'expert), un Leap-Motion est fixé sur la table pour le deuxième utilisateur(technicien). Le retour visuel des reproductions des gestes et de l'environnement partagé seront affichés à travers un visio-casque (casque-vidéo) pour l'expert et un écran pour le technicien.

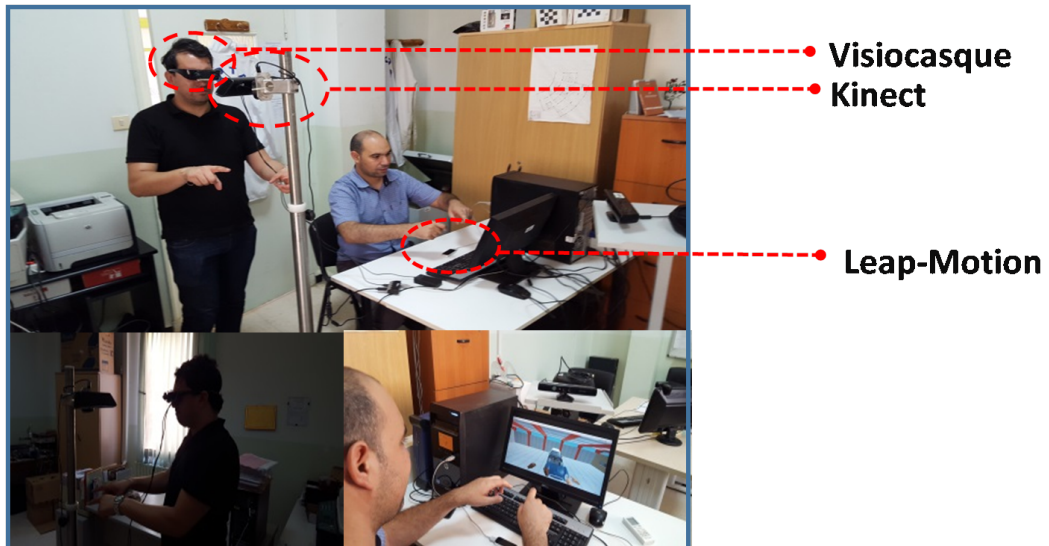


FIGURE 4.3 – Dispositif matériel utilisé.

L'environnement virtuel consiste en un garage de voitures avec des différentes pièces (capots, roues, moteurs etc.)(Voir Figure 4.4)



FIGURE 4.4 – Environnement virtuel développé.

4.3.2 Conception du système global

Cette étape consiste à identifier le déroulement des opérations de notre système global en identifiant les différents modules, leurs rôles ainsi que les données et les informations qui circulent entre eux (Voir Figure 4.5).

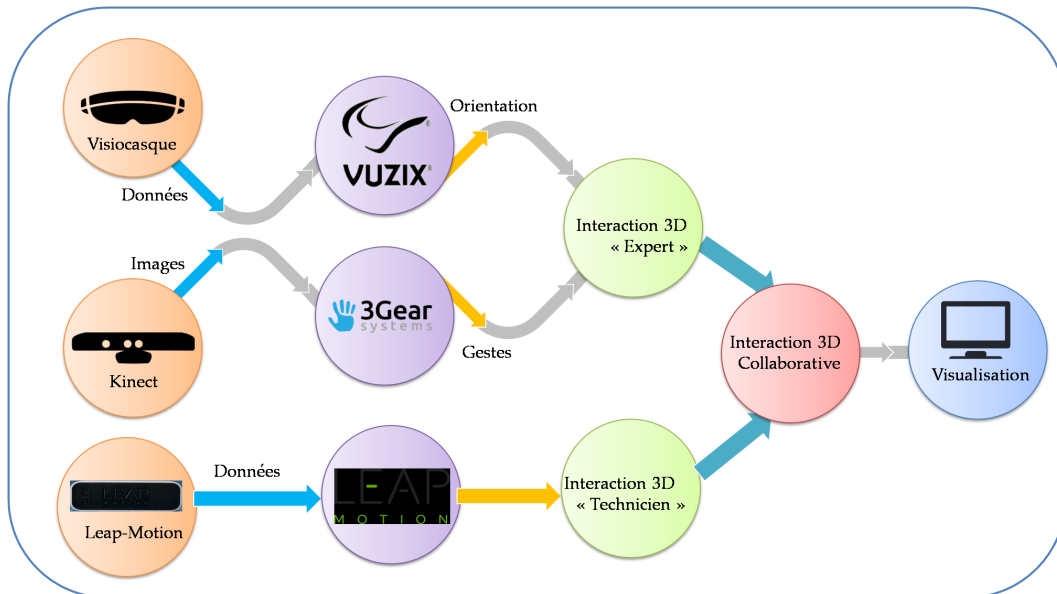


FIGURE 4.5 – L'architecture du système global

4.3.2.1 Interaction 3D

Ces modules collectent les différentes données, selon les outils utilisés par les utilisateurs (les positions 3D des différentes articulations de la main ainsi que les orientations de la tête). Après avoir collecté ces données, il applique l'interaction 3D mono-utilisateur pour chaque utilisateur, selon la métaphore ou la technique utilisée. Pour ce cas, l'expert utilisera la technique de la "Main virtuelle simple" [1], le technicien a son tour utilisera la technique "HOMER". Ces modules utiliseront l'architecture client/serveur, l'expert jouera le rôle de serveur dans celle-ci et le

technicien le rôle du client via un protocole réseau. Les résultats seront envoyés au module "Interaction 3D collaborative"

4.3.2.2 Interaction 3D Collaborative

Ce module décide de l'action collaborative à entreprendre selon le formalisme présenté dans la section 1 ainsi que les différents scénarios de collaboration qu'on veut mettre en œuvre. Dans cette partie on trouve les différentes tâches : navigation collaborative, sélection collaborative, manipulation collaborative et contrôle d'application.

— **Navigation 3D collaborative**

Pour la navigation collaborative nous avons choisi la technique de navigation de l'expert afin que le technicien puisse suivre son déplacement dans le monde virtuel partagé. La technique de navigation de l'expert combine la reconnaissance des gestes à la direction du regard à l'aide d'un HMD.

$$N_c = F(N_{u1}, N_{u2}) \quad (4.10)$$

$$N_c = N_{u1} \quad (4.11)$$

— **Sélection 3D collaborative**

La sélection collaborative permettra la sélection d'un objet appartenant au monde virtuel partagé afin qu'il soit manipulé dans la prochaine étape. Pour une sélection de groupe, nous avons le choix entre la technique de sélection de l'expert (Main Virtuelle Simple) ou celle du technicien (Ray-casting) selon le scénario.

$$S_c = G(S_{u1}, S_{u2}) \quad (4.12)$$

$$S_c = (S_{u1} \cup S_{u2}) - (S_{u1} \cap S_{u2}) \quad (4.13)$$

— **Manipulation 3D collaborative**

La manipulation collaborative assurera la manipulation de l'objet appartenant au monde virtuel partagé préalablement sélectionné lors de l'étape précédente. Pour une manipulation collaborative, nous avons développé les deux modes (synchrone et asynchrone).

Manipulation collaborative asynchrone L'expert manipulera l'objet partagé en premier dans le but de montrer et d'accompagner le technicien, par la suite le technicien le manipulera à son tour (Voir Figure 4.6).

$$M_c = H(M_{u1}, M_{u2}) \quad (4.14)$$

$$M_c = M_{u1} \quad (4.15)$$

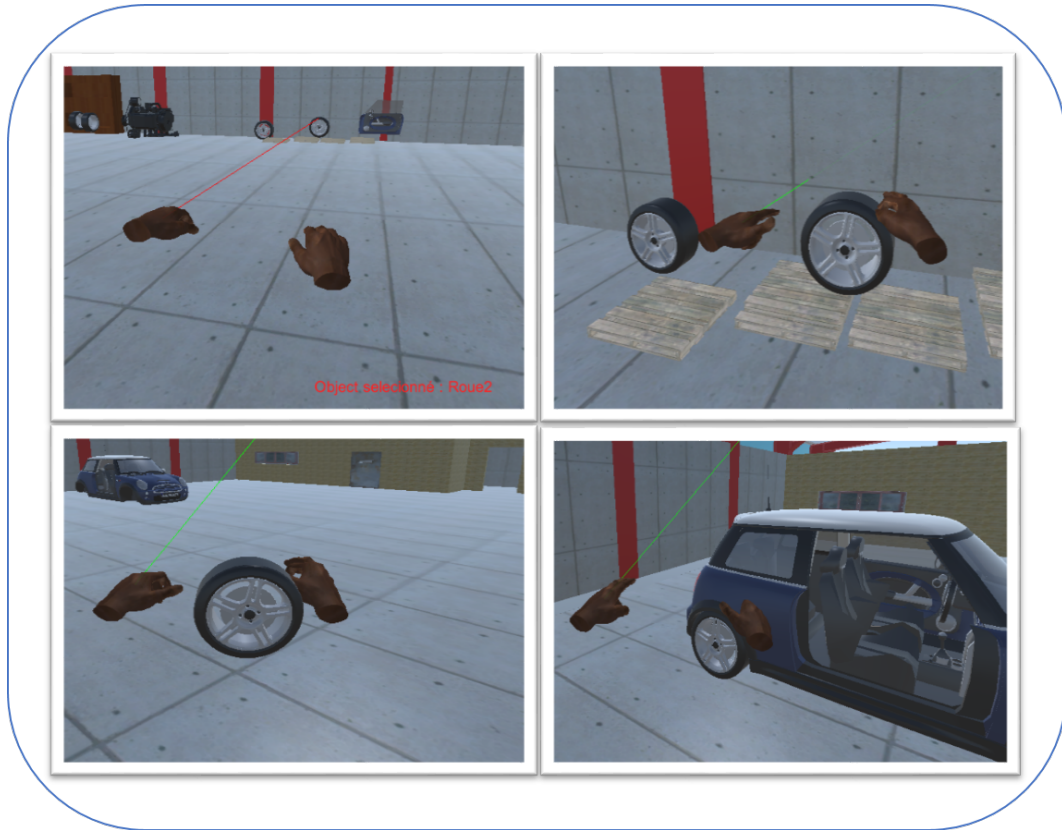


FIGURE 4.6 – Manipulation collaborative asynchrone

Manipulation collaborative synchrone

Les deux utilisateurs manipuleront en même temps l'objet partagé, le premier scénario étant une manipulation collaborative synchrone en séparant les degrés de liberté (position et orientation). L'expert change la position de l'objet partagé et le technicien change à son tour l'orientation du même objet en même temps (Voir Figure 4.7). Le deuxième scénario est une manipulation synchrone avec priorité, ou cette dernière est calculé par une fonction. Dans ce cas l'expert a une priorité pour manipuler l'objet partagé, cette priorité englobe la translation et l'orientation, on a introduit un pourcentage(coefficient d'influence) de 70%. Le but est de donner une priorité à l'expert pour mieux expliquer et accompagner le technicien (Voir Figure 4.8).

$$M_c = H(M_{u1}, M_{u2}) \quad (4.16)$$

Manipulation collaborative simultanée en séparant les degrés de liberté :

$$\pi_{p_i} = \begin{cases} t_1 pos_{u1} \theta_2 ori_{u2} \end{cases} \quad (4.17)$$

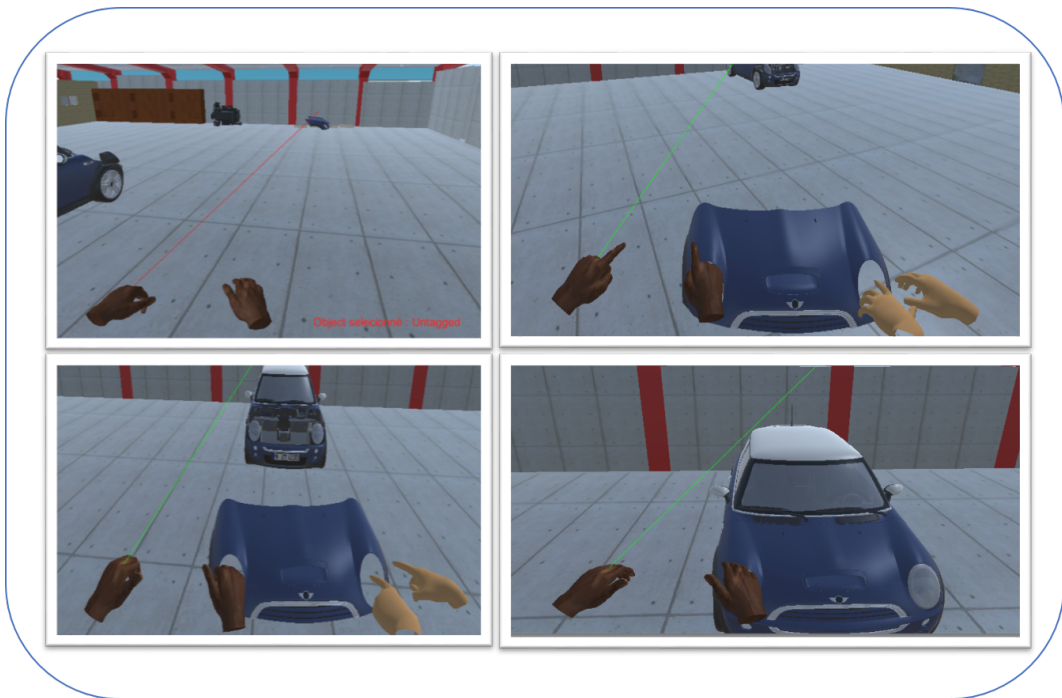


FIGURE 4.7 – Manipulation collaborative simultanée en séparant les degrés de liberté

Manipulation collaborative simultanée en utilisant une priorité :

$$\pi_{p_i} = \begin{cases} t_1 pos_{u1} + t_2 pos_{u2} \quad \theta_1 ori_{u1} + \theta_2 ori_{u2} & t_1 = 70\%t_2 \quad \text{et} \quad \theta_1 = 70\%\theta_2. \end{cases}$$

70% est le coefficient d'influence.

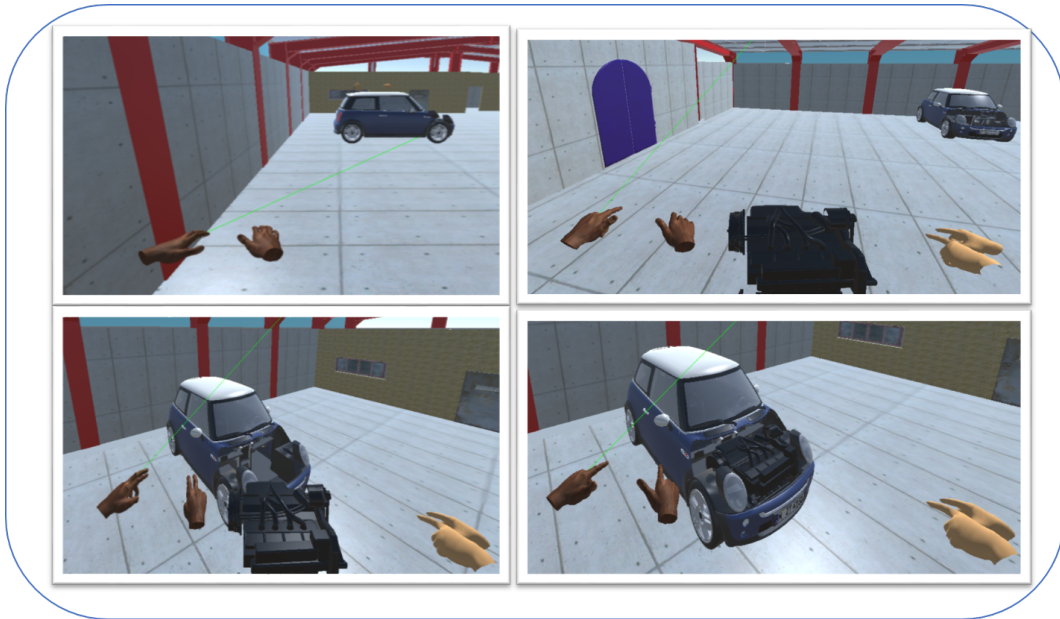


FIGURE 4.8 – Manipulation collaborative simultan e en utilisant une priorit e

Tous ces r sultats seront envoy es au module "Visualisation"

4.3.2.3 Visualisation

Ce module visualisera l'environnement partag e ainsi que l'interaction 3D collaborative.

4.4  valuation du syst me 3D collaboratif

Dans le domaine de la r alit e virtuelle et des environnements collaboratifs, l' valuation d'un syst me collaboratif ou d'une technique d'interaction collaborative est une  tape primordiale et n cessaire pour valider ce dernier ou cette derni re. Pour  valuer le syst me collaboratif propos e, on doit imp rativement le faire tester par un certain nombre d'utilisateurs. Le but ultime de notre  tude est de valider le formalisme de collaboration qui permet   partir de plusieurs techniques d'interactions

mono-utilisateur de produire une technique d'interaction 3D collaborative. Par manque de temps, nous avons pas pu réaliser l'évaluation, nous allons donc proposer le protocole expérimental.

Un ensemble d'étapes est nécessaire.

4.4.1 Hypothèse de recherche

La première étape du processus d'évaluation est de définir les différentes hypothèses de recherches sur lesquelles on va se baser pour concevoir notre protocole expérimental. Ces dernières seront des réponses aux différentes questions de recherche qu'on aurait défini au préalable.

4.4.1.1 Questions de recherche

- Techniques "*Main virtuelle simple*" et "*HOMER*" sont elles aussi bien efficaces dans un contexte collaboratif que mono-utilisateur ?
- Système d'apprentissage collaboratif pour la maintenance industrielle apporte-t-il un plus aux systèmes existants en intégrant le formalisme de l'interaction collaborative 3D ?
- Quel est le meilleur scénario d'interaction 3D collaborative ?

4.4.1.2 Hypothèses de recherche

Après avoir présenté les différentes questions de recherches, les hypothèses de recherches seront :

- H_1 **Les utilisateurs effectueront les tâches collaboratives mieux en utilisant le formalisme d'interaction 3D collaborative ;**

- H_2 Les utilisateurs effectueront les tâches collaborative facilement et rapidement ;
- H_3 La manipulation simultanée par séparation des degrés liberté est meilleure que la manipulation simultanée par priorité.

4.4.2 Le protocole expérimental

4.4.2.1 Les participants

Pour répondre à nos hypothèses nous prévoyons un échantillon de plus de 24 participants, âgé de plus de dix huit ans, entre hommes et femmes. Le nombre minimum de participants a été calculé selon la méthode de Cohen [155], dans le but d'obtenir une étude statistique satisfaisante. Le recrutement vise des personnes ayant ou non une expérience préalable des technologies de RV ou des jeux vidéo. Les sujets qui ne se sentent pas à l'aise avec le HMD(visio-casque), ou qui présentent des problèmes de motricité au niveau des membres supérieurs ne peuvent pas être inclus dans cette étude.

4.4.2.2 Conception du protocole expérimental

Tous les participants auront à effectuer des tâches expérimentales, soit en utilisant le formalisme d'interaction 3D collaborative ou le modèle "3C". Un premier groupe de douze personnes ($N_1=12$), ces derniers formeront des groupes de deux personnes qui effectueront les différentes tâches en utilisant le formalisme d'interaction 3D collaborative, et un deuxième groupe de douze personnes ($N_2=12$) qui utiliserons le modèle "3C". Les observations faites étant non appariées, le plan d'expérience sera donc du type "Between-subject". Par conséquent, tous les participants vont effectuer les tâches expérimentales trois fois.

4.4.2.3 Tâches expérimentales

Les tâches qui seront effectuées par les participants consistent à manipuler à deux des différents objets par des différents scénarios (synchrone, asynchrone).

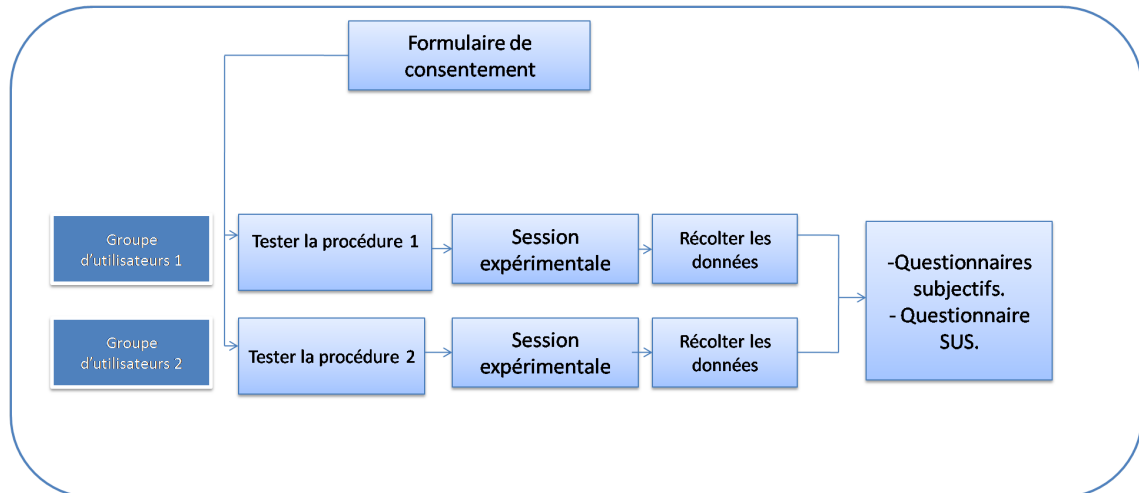


FIGURE 4.9 – Le protocole expérimental

4.4.3 Procédure expérimentale

La procédure expérimentale sera divisée en deux sessions (voir Figure 4.9). La procédure commence par la signature d'un formulaire de consentement et d'une fiche d'information (Âge, droitier/gaucher, portant des lunettes etc (Voir Annexe A, section A.1)).

Ensuite, un tutoriel détaillé consistant en une série d'étapes sous forme de courtes instructions et d'explications sur la manière d'effectuer les différentes tâches, d'utilisation du Leap-Motion, la Kinect et la présentation des différents gestes de la mains, est présenté aux participants. Cette phase d'expérimentation est accompagnée par de courtes phases de manipulation pour se familiariser avec les dispositifs.

Toutes les étapes pouvaient être répétées si nécessaire, et permettaient de comprendre progressivement chaque étape de la tâche à effectuer. La dernière étape consistait à effectuer toute une procédure de sélection/manipulation, pour garantir que le participant avait bien compris la tâche expérimentale. Une fois le tutoriel terminé, la session expérimentale commence pour le premier groupe avec la première technique. Chaque tâche doit être effectuée deux fois pour chaque scénario. En dernier, chaque participant remplira un questionnaire afin d'évaluer subjectivement les différentes tâches effectués.

4.4.4 Mesures et analyses statistiques des données

Les mesures de performance sont le temps de réalisation de la tâche (ms) et la précision du placement des objets(mm). En plus des mesures quantitatives, les participants sont invités à répondre au questionnaire comparatif pour chaque couple de techniques d'interaction et à la fin un questionnaire d'utilisabilité du système (SUS).

4.5 Conclusion et perspectives

L'objectif principal de la recherche dans le domaine de l'interaction 3D multi-utilisateurs est de permettre aux utilisateurs d'évoluer en groupe dans un environnement partagé, et d'interagir simultanément sur des objets communs. Dans ce domaine, un manque important de formalisme de coordination entre les différents acteurs de ce monde est ressenti. Nous avons présenté dans ce travail un système d'interaction 3D collaboratif qui permet de coordonner les actions des différents utilisateurs co-localisés, le

formalisme de l'interaction 3D collaborative assure une interaction simultanément cohérente sur les différents objets partagés.

Nous prévoyons en premier temps d'évaluer l'apprentissage collaboratif développé pour la maintenance industrielle par le biais du protocole expérimental.

Deuxièmement, nous envisageons l'implémentation et l'évaluation de la technique *Zoom-fwd* dans un environnement virtuel collaboratif, ceci nous permettra de vérifier si elle est si efficace et précise que dans un environnement mono-utilisateur.



Conclusion générale

Nous avons présenté dans cette thèse l'ensemble de nos travaux. Notre objectif était d'une part de proposer une technique d'interaction 3D basé sur la reconnaissance de gestes, permettant une interaction naturelle, intuitive, rapide et précise notamment avec les objets distants, occultés et se trouvant dans un milieu encombré, et d'autre part de formaliser les différentes actions collaboratives dans un environnement multi-utilisateurs.

Nous nous sommes intéressés à :

- L'amélioration de l'interaction 3D dans un environnement virtuel,
- La coordination des différentes actions des utilisateurs pour garantir une cohérence dans un environnement virtuel partagé.

Afin de répondre à cela, nous avons réalisé un état de l'art sur les travaux existants dans ce domaine. Ceci nous a permis d'orienter le reste de notre travail, en déterminant les parties où nous devons contribuer, les techniques que nous devons utiliser et les adaptations à envisager. Dans ce sens, nous avons dégagé deux axes principaux à savoir : l'interaction 3D mono et multi utilisateurs et les techniques de reconnaissances de gestes.

Dans un premier temps, nous avons travaillé sur l'interaction 3D qui est une composante motrice de la réalité virtuelle. Nous nous sommes donc penché sur les différentes techniques d'interactions, en relevant les verrous existant et en discutant

les éventuelles améliorations. A cet effet, une technique d'interaction 3D nommée "*Zoom-fwd*" basée sur la reconnaissance de gestes a été proposée. "*Zoom-fwd*" est une technique de sélection qui se base sur le principe du zoom, ce dernier permet à l'utilisateur de sélectionner et/ou manipuler des objets virtuels sans avoir à se déplacer quand la nécessité n'y est pas. L'utilisation de la reconnaissance de gestes permet d'avoir une interaction naturelle, intuitive et précise.

D'un autre côté, nous avons étudié les environnements virtuels collaboratifs, les interactions 3D collaboratives et les différents modes d'interactions sur des informations partagées. Le manque de formalisme de coordinations entre les différents utilisateurs dans un environnement partagé nous a poussé à présenter un formalisme d'interaction 3D collaborative, ce dernier coordonne les actions d'un groupe d'utilisateurs (cas expérimental de deux utilisateurs) sur un environnement partagé tout en gardant la cohérence de ce dernier.

Malgré l'émergence des approches d'intelligence artificielle : l'utilisation du deep learning pour assurer l'immersion et l'interaction 3D naturelle ou des machines learning pour optimiser les algorithmes de reconnaissances de gestes, nos travaux et contributions restent d'actualité.

Bien que les travaux aient donné des résultats satisfaisants lors des évaluations effectuées, des améliorations peuvent être envisagées à court terme :

1. Malgré les efforts fournis pour réaliser l'évaluation, celle-ci peut encore être améliorée : augmenter le nombre d'utilisateurs (un échantillon de 15 personnes reste petit), augmenter la complexité des tâches, etc ;
2. On prévoit de tester la technique "*Zoom-fwd*" dans une application médicale (application qui nécessite beaucoup de précision) ;
3. Réaliser l'évaluation du formalisme de l'interaction 3D collaborative.

Les améliorations à prévoir à moyen terme :

1. Évaluer d'autres paramètres tels que l'effet du casque sur les utilisateurs (zoomer et dé-zoomer peut provoquer des malaises chez certains utilisateurs);
2. Le suivi de la main dans les interfaces tels-que l'Oculus-Quest n'est pas encore aussi stable que Leap-Motion, mais il serait très intéressant de tester la technique "*Zoom-fwd*" sur ce dernier. Selon les résultats, la technique "*Zoom-fwd*" est stable, ce qui améliorera sûrement la reconnaissance de gestes;
3. Utiliser les approches de l'intelligence artificielle tel-que le deep learning pour assurer une meilleure reconnaissance de gestes;
4. Implémenter et tester la technique "*Zoom-fwd*" dans un environnement collaboratif partagé entre plusieurs utilisateurs co-localisés ou distants.

Beaucoup de verrous subsistent encore, autant sur le plan technologique que scientifique. En effet, bien que, les caméras, les casques de visualisation, les dispositifs d'interactions aient évolué ces dernières années (casques oculus, HTC Vive, dispositifs de reconnaissance de geste Leap-Motion, Realsense, etc.), ces technologies n'ont pas encore atteint leurs maximums. Aussi, les recherches en termes de techniques d'interaction mono ou multi-utilisateurs doivent être approfondies en bénéficiant des technologies d'aujourd'hui qui ouvrent un autre regard sur le monde de demain qui sera peut-être un monde virtuel.



Bibliographie

- [1] A MESSACI et al. “3d interaction techniques using gestures recognition in virtual environment”. In : *2015 4th International Conference on Electrical Engineering (ICEE)*. IEEE. 2015, p. 1-5.
- [2] A MESSACI, N ZENATI et M BELHOCINE. “3D Interaction Technique for Navigation, Selection and Manipulation”. In : *2018 The 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)*. ACM. 2018, p. 1-5.
- [3] Assia MESSACI et al. “Zoom-fwd : Efficient technique for 3D gestual interaction with distant and occluded objects in virtual reality”. In : *Computer Animation and Virtual Worlds* (2021), e2033.
- [4] Jaron LANIER. “A vintage virtual reality interview”. In : (1988).
- [5] Philippe FUCHS. *Les interfaces de la réalité virtuelle*. éditeur AJIIMD, 1996.
- [6] Stephen R ELLIS. “What are virtual environments?” In : *IEEE Computer Graphics and Applications* 14.1 (1994), p. 17-22.
- [7] Grigore BURDEA et Philippe COIFFET. “La rCalit6 virtuelle”. In : *Hermes* (1993).
- [8] Steve AUKSTAKALNIS et David BLATNER. “Silicon mirage : The art and science of virtual reality”. In : (1992).

- [9] Philippe FUCHS, Guillaume MOREAU et Alain BERTHOZ. “Le traité de la réalité virtuelle(Vol. 1, Fondements et interfaces comportementales)”. In : *Sciences mathématiques et informatique* (2003).
- [10] Dominik P. KÄSER et al. “The Making of Google Earth VR”. In : Association for Computing Machinery, 2017.
- [12] VR OSSO. “AAOS News : Osso VR to Showcase First Virtual Reality Training Module for Robotics”. In : (2020).
- [13] Samir BENBELKACEM et al. “Lung infection region quantification, recognition, and virtual reality rendering of CT scan of COVID-19”. In : *Multimodal Image Exploitation and Learning 2021*. T. 11734. International Society for Optics et Photonics. 2021, p. 117340I.
- [17] Ronald T AZUMA. “A survey of augmented reality”. In : *Presence : teleoperators & virtual environments* 6.4 (1997), p. 355-385.
- [19] Nazim HAOUCHINE et al. “Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery”. In : *2013 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE. 2013, p. 199-208.
- [20] Danica STANIMIROVIC et al. “Engineering and pharmacology of blood–brain barrier-permeable bispecific antibodies”. In : *Advances in Pharmacology* 71 (2014), p. 301-335.
- [21] Samir BENBELKACEM et al. “Augmented reality for photovoltaic pumping systems maintenance tasks”. In : *Renewable energy* 55 (2013), p. 428-437.

- [22] Nadia ZENATI et al. "E-maintenance for photovoltaic power system in Algeria". In : *2015 IEEE International Conference on Industrial Technology (ICIT)*. IEEE. 2015, p. 2594-2599.
- [23] Paul MILGRAM et Fumio KISHINO. "A taxonomy of mixed reality visual displays". In : *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), p. 1321-1329.
- [24] Mike HEIM Featherstone et Roger BURROWS. *Cyberspace/cyberbodies/cyberpunk : Cultures of technological embodiment*. Sage, 1995.
- [25] Roy S KALAWSKY. "Exploiting virtual reality techniques in education and training : Technological issues". In : <http://www.man.ac.uk/MVC/SIMA/vrtech/title.html> (1996).
- [28] Ivan POUPYREV et Tadao ICHIKAWA. "Manipulating objects in virtual worlds : Categorization and empirical evaluation of interaction techniques". In : *Journal of Visual Languages & Computing* 10.1 (1999), p. 19-35.
- [29] Douglas A BOWMAN. *Interaction techniques for common tasks in immersive virtual environments : design, evaluation, and application*. Georgia Institute of Technology, 1999.
- [30] William GIBSON. "Neuromancer. New York, Ace Science Fiction, 1984". In : *All subsequent quotations from the novel will refer to this edition and will be paginated in my text* (1989).
- [32] Laurent AGUERRECHE, Thierry DUVAL et Bruno ARNALDI. "Analyse de techniques de coopération en environnements virtuels 3D". In : *Revue des Sciences et Technologies de l'Information-Série TSI : Technique et Science Informatiques* 28.6-7 (2009), p. 767-797.

- [33] Nassima OURAMDANE, Samir OTMANE et Malik MALLEM. "Interaction 3D en Réalité Virtuelle-Etat de l'art". In : *Technique et Science Informatiques* 28.8 (2009), p. 1017-1049.
- [34] Chris HAND. "A survey of 3D interaction techniques". In : *Computer graphics forum*. T. 16. 5. Wiley Online Library. 1997, p. 269-281.
- [35] Joseph J LAVIOLA JR et al. *3D user interfaces : theory and practice*. Addison-Wesley Professional, 2017.
- [36] Shumin ZHAI, William BUXTON et Paul MILGRAM. "The "Silk Cursor" investigating transparency for 3D target acquisition". In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1994, p. 459-464.
- [37] Doug A BOWMAN et al. *3D user interfaces : theory and practice*. Addison-Wesley, 2005.
- [38] Doug A BOWMAN, Joseph L GABBARD et Deborah HIX. "A survey of usability evaluation in virtual environments : classification and comparison of methods". In : *Presence : Teleoperators & Virtual Environments* 11.4 (2002), p. 404-424.
- [39] Jacek JANKOWSKI et Martin HACHET. "A survey of interaction techniques for interactive 3D environments". In : *Eurographics 2013-STAR*. 2013.
- [40] Mark MINE. "ISAAC : A virtual environment tool for the interactive construction of virtual worlds". In : (1995).
- [41] Hirokazu KATO et al. "Virtual object manipulation on a table-top AR environment". In : *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. Ieee. 2000, p. 111-119.

- [42] Richard H JACOBY, Mark FERNEAU et Jim HUMPHRIES. “Gestural interaction in a virtual environment”. In : *Stereoscopic Displays and Virtual Reality Systems*. T. 2177. International Society for Optics et Photonics. 1994, p. 355-364.
- [43] Ferran ARGELAGUET et Carlos ANDUJAR. “A survey of 3D object selection techniques for virtual environments”. In : *Computers & Graphics* 37.3 (2013), p. 121-136.
- [44] Krzysztof PIETROSZEK et Chao Cheng LIN. “Univresity : Face-to-face class participation for remote students using virtual reality”. In : *25th ACM symposium on virtual reality software and technology*. 2019, p. 1-2.
- [45] Ivan POUPYREV et al. “The go-go interaction technique : non-linear mapping for direct manipulation in VR”. In : *Proceedings of the 9th annual ACM symposium on User interface software and technology*. 1996, p. 79-80.
- [46] Eliane BALAA et al. “Classification of interaction techniques in the 3d virtual environment on mobile devices”. In : *International Conference on Virtual, Augmented and Mixed Reality*. Springer. 2014, p. 3-13.
- [47] Doug A BOWMAN et Larry F HODGES. “An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments”. In : *Proceedings of the 1997 symposium on Interactive 3D graphics*. 1997, 35-ff.
- [48] Scott FREES et G Drew KESSLER. “Precise and rapid interaction through scaled manipulation in immersive virtual environments”. In : *IEEE Proceedings. VR 2005. Virtual Reality, 2005*. IEEE. 2005, p. 99-106.

- [49] Lode VANACKEN, Tovi GROSSMAN et Karin 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*. IEEE. 2007.
- [50] Richard A BOLT. “?Put-that-there? Voice and gesture at the graphics interface”. In : *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*. 1980, p. 262-270.
- [51] Raffaele DE AMICIS, Michele FIORENTINO et Andre STORK. “Parametric interaction for cad application in virtual reality environment”. In : *International Conference on Design Tools and Methods in Industrial Engineering, pages D. T. 3*. 2001.
- [52] Jeffrey S PIERCE et al. “Image plane interaction techniques in 3D immersive environments”. In : *Proceedings of the 1997 symposium on Interactive 3D graphics*. 1997, 39-ff.
- [53] Jang Han LEE et al. “Experimental application of virtual reality for nicotine craving through cue exposure”. In : *CyberPsychology & Behavior* 6.3 (2003), p. 275-280.
- [54] Alex Olwal Steven FEINER. “The flexible pointer : An interaction technique for selection in augmented and virtual reality”. In : *Proc. UIST. T. 3*. 2003, p. 81-82.
- [55] Jiandong LIANG et Mark GREEN. “JDCAD : A highly interactive 3D modeling system”. In : *Computers & graphics* 18.4 (1994), p. 499-506.
- [56] Andrew FORSBERG, Kenneth HERNDON et Robert ZELEZNIK. “Aperture based selection for immersive virtual environments”. In : *Proceedings of the*

- 9th annual ACM symposium on User interface software and technology*. 1996, p. 95-96.
- [57] Ferran ARGELAGUET et Carlos ANDUJAR. “Efficient 3D pointing selection in cluttered virtual environments”. In : *IEEE Computer Graphics and Applications* 29.6 (2009), p. 34-43.
- [58] Tovi GROSSMAN et Ravin 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*. 2006, p. 3-12.
- [59] Richard STOAKLEY, Matthew J CONWAY et Randy PAUSCH. “Virtual reality on a WIM : interactive worlds in miniature”. In : *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1995, p. 265-272.
- [60] Chadwick A WINGRAVE, Yonca HACIAHMETOGLU et Doug A BOWMAN. “Overcoming world in miniature limitations by a scaled and scrolling WIM”. In : *3D User Interfaces (3DUI'06)*. IEEE. 2006, p. 11-16.
- [61] Jeffrey S PIERCE, Brian C STEARNS et Randy PAUSCH. “Voodoo dolls : seamless interaction at multiple scales in virtual environments”. In : *Proceedings of the 1999 symposium on Interactive 3D graphics*. 1999, p. 141-145.
- [62] Jeffrey S PIERCE et Randy PAUSCH. “Comparing voodoo dolls and HOMER : exploring the importance of feedback in virtual environments”. In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2002, p. 105-112.

- [63] Adalberto L SIMEONE. “Indirect touch manipulation for interaction with stereoscopic displays”. In : *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE. 2016, p. 13-22.
- [64] Mark R MINE, Frederick P BROOKS JR et Carlo H SEQUIN. “Moving objects in space : exploiting proprioception in virtual-environment interaction”. In : *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 1997, p. 19-26.
- [65] Andreas SIMON. “First-person experience and usability of co-located interaction in a projection-based virtual environment”. In : *Proceedings of the ACM symposium on Virtual reality software and technology*. 2005, p. 23-30.
- [66] Jaemoon JUNG et al. “A review on interaction techniques in virtual environments”. In : *Proc. 2014 international conference on industrial engineering and operations management*. 2014, p. 1582-1590.
- [67] Daniel MENDES et al. “A survey on 3d virtual object manipulation : From the desktop to immersive virtual environments”. In : *Computer graphics forum*. T. 38. 1. Wiley Online Library. 2019, p. 21-45.
- [68] Curtis WILKES et Doug A BOWMAN. “Advantages of velocity-based scaling for distant 3D manipulation”. In : *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*. 2008, p. 23-29.
- [69] Regis KOPPER, Felipe BACIM et Doug A BOWMAN. “Rapid and accurate 3D selection by progressive refinement”. In : *2011 IEEE symposium on 3D user interfaces (3DUI)*. IEEE. 2011, p. 67-74.

- [70] Jeffrey CASHION, Chadwick WINGRAVE et Joseph J LAVIOLA JR. “Dense and dynamic 3d selection for game-based virtual environments”. In : *IEEE transactions on visualization and computer graphics* 18.4 (2012), p. 634-642.
- [71] Sascha GEBHARDT et al. “Extended pie menus for immersive virtual environments”. In : *IEEE transactions on visualization and computer graphics* 19.4 (2013), p. 644-651.
- [72] Kevin ARTHUR et al. “Designing and building the pit : a head-tracked stereo workspace for two users”. In : *2nd International Immersive Projection Technology Workshop*. 1998, p. 11-12.
- [74] Kevin CHOW et al. “Challenges and design considerations for multimodal asynchronous collaboration in VR”. In : *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), p. 1-24.
- [75] Chris GREENHALGH et al. “Applications of temporal links : Recording and replaying virtual environments”. In : *Proceedings IEEE Virtual Reality 2002*. IEEE. 2002, p. 101-108.
- [76] Tiago Joao GUERREIRO et al. “Beyond Post-It : Structured Multimedia Annotations for Collaborative VEs.” In : *ICAT-EGVE*. Citeseer. 2014, p. 55-62.
- [77] Márcio S PINHO, Doug A BOWMAN et Carla MDS FREITAS. “Cooperative object manipulation in immersive virtual environments : framework and techniques”. In : *Proceedings of the ACM symposium on Virtual reality software and technology*. 2002, p. 171-178.
- [78] Marcio S PINHO, Doug A BOWMAN et Carla M FREITAS. “Cooperative object manipulation in collaborative virtual environments”. In : *Journal of the Brazilian Computer Society* 14.2 (2008), p. 53-67.

- [79] Roy A RUDDLE, Justin CD SAVAGE et Dylan M JONES. “Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments”. In : *ACM Transactions on Computer-Human Interaction (TOCHI)* 9.4 (2002), p. 285-308.
- [80] Thierry DUVAL, Anatole LÉCUYER et Sébastien THOMAS. “Skewer : a 3d interaction technique for 2-user collaborative manipulation of objects in virtual environments”. In : *3D User Interfaces (3DUI'06)*. IEEE. 2006, p. 69-72.
- [81] Thierry DUVAL et Christian LE TENIER. “Interactions 3D coopératives en environnements virtuels avec OpenMASK pour l’exploitation d’objets techniques”. In : *Mechanics & Industry* 5.2 (2004), p. 129-137.
- [82] Thierry DUVAL et Aurélien FENALS. “Faciliter la perception de l’interaction lors de manipulations coopératives simultanées en environnements virtuels 3d”. In : *annex of the Proceedings of IHM 2002* (2002), p. 29-32.
- [83] Leonardo Pavanatto SOARES, Márcio Sarroglia PINHO et Regis KOPPER. “Design and preliminary evaluation of an EGO-exocentric technique for cooperative manipulation”. In : *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE. 2017, p. 203-204.
- [84] Jerônimo Gustavo GRANDI et al. “Design and evaluation of a handheld-based 3d user interface for collaborative object manipulation”. In : *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, p. 5881-5891.
- [85] Jonathan WIELAND et al. “Separation, Composition, or Hybrid? Comparing Collaborative 3D Object Manipulation Techniques for Handheld Augmented Reality”. In : *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2021, p. 403-412.

- [86] “The Psychology of Human-Computer Interaction”. In : *CRC Press*. Taylor. 1983.
- [87] Philippe FUCHS. *Le traité de la réalité virtuelle*. T. 2. Presses des MINES, 2006.
- [88] Johnny Chung LEE. “Hacking the nintendo wii remote”. In : *IEEE pervasive computing* 7.3 (2008), p. 39-45.
- [89] Joshua BLAKE. “The natural user interface revolution. Natural User Interfaces in .NET”. In : *Manning Publications Co* (2012).
- [90] Kenton O’HARA et al. “On the naturalness of touchless : putting the interaction back into NUI”. In : *ACM Transactions on Computer-Human Interaction (TOCHI)* 20.1 (2013), p. 1-25.
- [91] Thomas BAUDEL et Michel BEAUDOUIN-LAFON. “Charade : remote control of objects using free-hand gestures”. In : *Communications of the ACM* 36.7 (1993), p. 28-35.
- [92] Susan GOLDIN-MEADOW et Martha Wagner ALIBALI. “Gesture’s role in speaking, learning, and creating language”. In : *Annual review of psychology* 64 (2013), p. 257-283.
- [93] Claude CADOZ. “Le geste canal de communication homme/machine : la communication" instrumentale"”. In : *Revue des Sciences et Technologies de l’Information-Série TSI : Technique et Science Informatiques* 13.1 (1994), p. 31-61.
- [94] Joseph J LAVIOLA JR. “A survey of hand posture and gesture recognition techniques and technology”. In : (1999).

- [95] Ng Yong Yi KEVIN, S RANGANATH et D GHOSH. “Trajectory modeling in gesture recognition using CyberGloves/sup/spl reg//and magnetic trackers”. In : *2004 IEEE Region 10 Conference TENCN 2004*. IEEE. 2004, p. 571-574.
- [97] Nicholas KATZAKIS et al. “INSPECT : extending plane-casting for 6-DOF control”. In : *Human-centric Computing and Information Sciences* 5.1 (2015), p. 1-22.
- [98] Saiwen WANG et al. “Interacting with soli : Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum”. In : *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 2016, p. 851-860.
- [99] Jaime LIEN et al. “Soli : Ubiquitous gesture sensing with millimeter wave radar”. In : *ACM Transactions on Graphics (TOG)* 35.4 (2016), p. 1-19.
- [100] Li DU et al. “Airtouch : A novel single layer 3D touch sensing system for human/mobile devices interactions”. In : *2016 53rd ACM/EDAC/IEEE Design Automation Conference (DAC)*. IEEE. 2016, p. 1-6.
- [103] F LEITNER et P CINQUIN. “From snake and splines to snake-splines”. In : *Workshop on Geometric Reasoning for Perception and Action*. 1993, p. 264-281.
- [104] Shanon X JU et al. “Analysis of gesture and action in technical talks for video indexing”. In : *Proceedings of IEEE computer society conference on computer vision and pattern recognition*. IEEE. 1997, p. 595-601.
- [105] Lars BRETZNER, Ivan LAPTEV et Tony LINDBERG. “Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering”. In : *Proceedings of fifth IEEE international conference on automatic face gesture recognition*. IEEE. 2002, p. 423-428.

- [106] Robert Y WANG et Jovan POPOVIĆ. “Real-time hand-tracking with a color glove”. In : *ACM transactions on graphics (TOG)* 28.3 (2009), p. 1-8.
- [107] Abdelkader BELLARBI et al. “Hand gesture interaction using color-based method for tabletop interfaces”. In : *2011 IEEE 7th International Symposium on Intelligent Signal Processing*. IEEE. 2011, p. 1-6.
- [108] Ahmet BIRDAL et Reza HASSANPOUR. “Region based hand gesture recognition”. In : (2008).
- [109] Kouichi MURAKAMI. “Gesture Recognition Using Recurrent Neural Networks”. In : *Proceedings of CHI'91 Human Factors in Computing Systems*. 1991, p. 237-242.
- [110] Ekaterini STERGIOPOULOU et Nikos PAPAMARKOS. “Hand gesture recognition using a neural network shape fitting technique”. In : *Engineering Applications of Artificial Intelligence* 22.8 (2009), p. 1141-1158.
- [111] Haitham HASAN et S ABDUL-KAREEM. “RETRACTED ARTICLE : Static hand gesture recognition using neural networks”. In : *Artificial Intelligence Review* 41.2 (2014), p. 147-181.
- [112] Christian VOGLER et Dimitris METAXAS. “Handshapes and movements : Multiple-channel american sign language recognition”. In : *International Gesture Workshop*. Springer. 2003, p. 247-258.
- [113] Liu YUN, Zhang LIFENG et Zhang SHUJUN. “A hand gesture recognition method based on multi-feature fusion and template matching”. In : *Procedia Engineering* 29 (2012), p. 1678-1684.

- [114] Jinjun RAO et al. “Low cost hand gesture learning and recognition system based on hidden markov model”. In : *2009 Second International Symposium on Information Science and Engineering*. IEEE. 2009, p. 433-438.
- [115] Yun LIU, Yanmin YIN et Shujun ZHANG. “Hand gesture recognition based on HU moments in interaction of virtual reality”. In : *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics*. T. 1. IEEE. 2012, p. 145-148.
- [116] S Padam PRIYAL et Prabin Kumar BORA. “A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments”. In : *Pattern Recognition* 46.8 (2013), p. 2202-2219.
- [117] Yen-Ting CHEN et Kuo-Tsung TSENG. “Multiple-angle hand gesture recognition by fusing SVM classifiers”. In : *2007 IEEE International Conference on Automation Science and Engineering*. IEEE. 2007, p. 527-530.
- [118] Deng-Yuan HUANG, Wu-Chih HU et Sung-Hsiang CHANG. “Gabor filter-based hand-pose angle estimation for hand gesture recognition under varying illumination”. In : *Expert Systems with Applications* 38.5 (2011), p. 6031-6042.
- [119] Cem KESKIN, Ali Taylan CEMGIL et Lale AKARUN. “DTW based clustering to improve hand gesture recognition”. In : *International Workshop on Human Behavior Understanding*. Springer. 2011, p. 72-81.
- [120] Cem KESKIN et al. “Randomized decision forests for static and dynamic hand shape classification”. In : *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE. 2012, p. 31-36.

- [121] Xian ZHAO et al. "Real-time hand gesture detection and recognition by random forest". In : *Communications and information processing*. Springer, 2012, p. 747-755.
- [122] Wei-Syun LIN et al. "A study of real-time hand gesture recognition using SIFT on binary images". In : *Advances in Intelligent Systems and Applications-Volume 2*. Springer, 2013, p. 235-246.
- [123] Yi YAO et Chang-Tsun LI. "Real-time hand gesture recognition for uncontrolled environments using adaptive SURF tracking and hidden conditional random fields". In : *International Symposium on Visual Computing*. Springer. 2013, p. 542-551.
- [124] Youdong DING et al. "Recognition of hand-gestures using improved local binary pattern". In : *2011 International Conference on Multimedia Technology*. IEEE. 2011, p. 3171-3174.
- [125] Abdelkader BELLARBI et al. "Hand gesture recognition using contour based method for tabletop surfaces". In : *2013 10th IEEE INTERNATIONAL CONFERENCE ON NETWORKING, SENSING AND CONTROL (ICNSC)*. IEEE. 2013, p. 832-836.
- [126] Nadia ZENATI-HENDA et al. "Augmented reality system based on hand gestures for remote maintenance". In : *2014 International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE. 2014, p. 5-8.
- [127] Tony HEAP et David HOGG. "Towards 3D hand tracking using a deformable model". In : *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*. Ieee. 1996, p. 140-145.

- [128] Carlo TOMASI, Slav PETROV et Arvind SASTRY. “3D Tracking= Classification+ Interpolation.” In : *ICCV*. T. 3. 2003, p. 1441.
- [129] Ying WU, John LIN et Thomas S HUANG. “Analyzing and capturing articulated hand motion in image sequences”. In : *IEEE transactions on pattern analysis and machine intelligence* 27.12 (2005), p. 1910-1922.
- [130] Vassilis ATHITSOS et Stan SCLAROFF. “Estimating 3D hand pose from a cluttered image”. In : *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*. T. 2. IEEE. 2003, p. II-432.
- [131] Akihiro IMAI, Nobutaka SHIMADA et Yoshiaki SHIRAI. “3-d hand posture recognition by training contour variation”. In : *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*. IEEE. 2004, p. 895-900.
- [132] Robert WANG, Sylvain PARIS et Jovan POPOVIĆ. “6D hands : markerless hand-tracking for computer aided design”. In : *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 2011, p. 549-558.
- [133] Sameh KHAMIS et al. “Learning an efficient model of hand shape variation from depth images”. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, p. 2540-2548.
- [134] Jonathan TAYLOR et al. “Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences”. In : *ACM Transactions on Graphics (TOG)* 35.4 (2016), p. 1-12.
- [135] Chen QIAN et al. “Realtime and robust hand tracking from depth”. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, p. 1106-1113.

- [136] Stan MELAX, Leonid KESELMAN et Sterling ORSTEN. “Dynamics based 3D skeletal hand tracking”. In : *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. 2013, p. 184-184.
- [137] Abdelkader BELLARBI et al. “A 3d interaction technique for selection and manipulation distant objects in augmented reality”. In : *2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B)*. IEEE. 2017, p. 1-5.
- [138] Clement FARABET et al. “Learning hierarchical features for scene labeling”. In : *IEEE transactions on pattern analysis and machine intelligence* 35.8 (2012), p. 1915-1929.
- [139] Jonathan TOMPSON et al. “Real-time continuous pose recovery of human hands using convolutional networks”. In : *ACM Transactions on Graphics (ToG)* 33.5 (2014), p. 1-10.
- [140] Markus OBERWEGER, Paul WOHLHART et Vincent LEPETIT. “Hands deep in deep learning for hand pose estimation”. In : *arXiv preprint arXiv :1502.06807* (2015).
- [141] Ao TANG et al. “A real-time hand posture recognition system using deep neural networks”. In : *ACM Transactions on Intelligent Systems and Technology (TIST)* 6.2 (2015), p. 1-23.
- [142] Di WU et al. “Deep dynamic neural networks for multimodal gesture segmentation and recognition”. In : *IEEE transactions on pattern analysis and machine intelligence* 38.8 (2016), p. 1583-1597.

- [144] Felipe BACIM, Mahdi NABIYOUNI et Doug A BOWMAN. “Slice-n-Swipe : A free-hand gesture user interface for 3D point cloud annotation”. In : *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE. 2014, p. 185-186.
- [145] Minseok KIM et Jae Yeol LEE. “Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality”. In : *Multimedia Tools and Applications* 75.23 (2016), p. 16529-16550.
- [146] Giuseppe CAGGIANESE, Luigi GALLO et Pietro NERONI. “An investigation of leap motion based 3D manipulation techniques for use in egocentric viewpoint”. In : *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer. 2016, p. 318-330.
- [147] Matthew M DAVIS et al. “Depth-based 3D gesture multi-level radial menu for virtual object manipulation”. In : *2016 IEEE Virtual Reality (VR)*. IEEE. 2016, p. 169-170.
- [148] Parth Rajesh DESAI et al. “A review paper on oculus rift-a virtual reality headset”. In : *arXiv preprint arXiv :1408.1173* (2014).
- [150] Arnold M LUND. “Measuring usability with the use questionnaire12”. In : *Usability interface* 8.2 (2001), p. 3-6.
- [151] John BROOKE. “SUS : a “quick and dirty” usability scale”. In : *Usability evaluation in industry* 189.194 (1996), p. 4-7.
- [152] Aaron BANGOR, Philip KORTUM et James MILLER. “Determining what individual SUS scores mean : Adding an adjective rating scale”. In : *Journal of usability studies* 4.3 (2009), p. 114-123.

- [153] Nader CHEAIB, Samir OTMANE et Malik MALLEM. “A machine-machine collaboration formalism based on web services for groupware tailorability”. In : *Proceedings of the 2011 15th international conference on computer supported cooperative work in design (CSCWD)*. IEEE. 2011, p. 238-245.
- [154] Clarence ELLIS et Jacques WAINER. “A conceptual model of groupware”. In : *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. 1994, p. 79-88.
- [155] Jacob COHEN. “Quantitative methods in psychology : A power primer”. In : *Psychological bulletin*. Citeseer. 1992.



Webographie

- [11] SHARECARE. *Getting started with the body*. 2020. Consulté 06/2023. URL : https://www.oculus.com/experiences/quest/5090936304309796/locale=fr_FR.
- [14] TRUESCALE. *Getting started with the architecture*. 2017. Consulté : 06/2023. URL : <https://www.viveport.com/apps/b6a97756-4fce-49aa-a0c5-a111b122ba48>.
- [15] FLYINSIDE. *Getting started with the fly*. 2019. Consulté 09/2021. URL : <https://flyinside-fsx.com/>.
- [16] VRSUTZ. *Getting started with the militaryVR*. 2021. Consulté : 02/2022. URL : <https://flyinside-fsx.com/>.
- [143] David HOLZ, Kyle HAY et Michael BUCKWALD. *Electronic sensor*. 2015. Consulté : 06/2023. URL : <https://www.ultraleap.com/product/leap-2/>.
- [149] DAVIS. *Getting started with the Leap Motion SDK*. 2014. Consulté : 12/2022. URL : <http://blog.leapmotion.com/getting-started-leap-motion-sdk/>.



Annexe

4.6 Questions démographique

1. Numéro du participant : * Informations générales : *
2. Âge : *
3. Sexe : *
4. Main dominante : *
5. Portez-vous des lentilles/lunettes correctrices ? *
6. Si oui, les portez-vous pendant l'expérience ? Expérience avec les jeux vidéo
7. Avez-vous déjà joué à des jeux vidéo ? *
8. Si oui, il y a combien d'années avez-vous commencé à jouer à des jeux vidéo ?
9. Si oui, a quelle fréquence jouez-vous aux jeux vidéo ?
10. Si oui, quand était la dernière fois ? Expérience avec la réalité virtuelle
11. Avez-vous déjà utilisé des environnements virtuels 3D ? *
12. Si oui, a quelle fréquence utilisez-vous des environnements virtuels 3D ?
13. Si oui, quand était la dernière fois ?
14. et dans quelle occasion était-il ?
15. Avez-vous déjà utilisé des casques de réalité virtuelle ? *
16. Si OUI, a quelle fréquence utilisez-vous des casques de réalité virtuelle ?
17. Si OUI, pendant combien de temps les portez-vous à chaque occasion ?

4.7 Questionnaires d'évaluation

4.7.1 Questionnaire d'évaluation USE

1. Utilité

- Cela m'aide à être plus efficace.
- Cela m'aide à être plus productif.
- C'est utile.
- Elle rend les choses que je veux accomplir plus facile à faire.
- Cela m'économise du temps quand je l'utilise.
- Elle répond à mes besoins.
- Elle fait tout ce que je m'attends à ce qu'il fasse.

2. Facilité de l'utilisation

- C'est facile à utiliser.
- Elle est simple à utiliser.
- Elle est conviviale.
- Elle exige le moins d'étapes possible pour accomplir ce que je veux faire.
- Elle est flexible.
- Son utilisation nécessite le moindre d'effort.
- Je peux l'utiliser sans instructions écrites.
- Je ne remarque aucune incohérence au moment où je l'utilise.

- Je peux me remettre rapidement des erreurs.
- Je peux l'utiliser avec succès à chaque fois.

3. Facilité d'apprentissage

- J'ai appris à l'utiliser rapidement.
- Je me souviens facilement comment l'utiliser.
- Elle est facile d'apprendre à l'utiliser.
- Je deviens rapidement habile avec elle.

4. Satisfaction

- Elle est nécessaire.
- J'en suis satisfait.
- Je la recommande à un ami.
- C'est amusant de l'utiliser.
- Elle fonctionne de la façon dont je veux qu'elle fonctionne.
- C'est merveilleux.
- Je sens que je dois l'avoir.
- Elle est agréable à utiliser.

4.7.2 Questionnaire d'évaluation sur l'Utilisabilité du Système SUS

- Numéro du participant : *
- Je pense que je vais utiliser ce système fréquemment. *

- Je pense que ce système est facile à utiliser. *
- Je trouve ce système inutilement complexe. *
- Je pense que j'aurais besoin d'un support technique pour pouvoir utiliser ce système. *
- J'ai trouvé que les différentes fonctionnalités de ce système sont bien intégrées.
*
- Je pense qu'il y a trop d'incohérence dans ce système. *
- J'imagine que la plupart des gens serait capable d'apprendre à utiliser ce système très rapidement. *
- J'ai trouvé ce système très lourd à utiliser. *
- Je me sentais très en confiance en utilisant ce système. *
- J'ai besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système.
*

Résumé

La réalité virtuelle est une discipline qui se situe à la croisée de plusieurs domaines tels que l'Interaction Homme Machine (IHM), la Vision par Ordinateur et le Travail Collaboratif Assisté par Ordinateur. L'interaction 3D est la composante motrice de la réalité virtuelle. Elle permet à l'homme d'être un acteur capable de changer le cours des événements dans un environnement synthétique et ainsi d'interagir avec des entités virtuelles. Il existe différentes techniques d'interaction 3D qui sont dédiées aux différentes tâches d'interaction 3D à savoir la navigation, la sélection, la manipulation et le contrôle d'application. Dans la plupart des cas, les techniques d'interaction 3D traitent l'interaction d'une manière générale et ne répondent pas aux exigences de certaines applications telles que la précision en particulier avec des objets distants, occultés et se trouvant dans un milieu encombré.

Les environnements collaboratifs quant à eux, sont des mondes peuplés, dans lesquels les utilisateurs interagissent entre eux pour réaliser des objectifs ensemble. La complexité de ces environnements est liée à l'interaction d'un groupe d'utilisateurs avec des entités partagées.

Dans ce travail, nous avons développé dans un premier temps une technique d'interaction 3D "Zoom-fwd" basée sur la reconnaissance de gestes, cette dernière permet à l'utilisateur de sélectionner et/ou manipuler précisément des objets virtuels distants et occultés sans avoir à se déplacer. Dans un second temps, nous avons présenté un formalisme d'interaction 3D collaborative, ce dernier coordonne les actions d'un groupe d'utilisateurs sur un environnement partagé tout en gardant la cohérence de ce dernier. Notre technique d'interaction 3D peut être considérée comme un moyen de contourner les limitations de certaines techniques d'interaction 3D classiques afin de les rendre utilisables dans des contextes différents.

Mots clés : Réalité virtuelle, Interaction 3D, Techniques d'Interaction 3D, Interaction 3D Collaborative, Travail Collaboratif, Environnements Virtuels Collaboratifs.

Abstract

Virtual Reality is a research domain that is situated at the crossing of several domains as Human Computer Interaction (HCI), computer aided design and collaborative work, etc. The 3D interaction is a core component of virtual reality. It allows the user to be an actor able to interact with virtual entities. There exists different 3D interaction techniques that are dedicated to the different tasks of 3D interaction that are navigation, selection, manipulation and command control. In most cases, the 3D interaction techniques are only used to perform one of the 3D interaction tasks without taking into account the requirements of some applications such as precision especially with distant, occluded objects in a crowded environment.

The collaborative environments are populated worlds, in which users interact between them to accomplish goals together. The complexity of these environments is related to the interaction of a group of users with shared entities.

In this work, the "Zoom-fwd" which is an efficient 3D interaction technique is presented.

The proposed technique uses gesture recognition for different 3D interaction tasks like selection and manipulation. This new approach allows an efficient interaction with distant and occluded objects, while providing a precise selection, even when the environment is crowded. On the other hand, we presented a collaborative 3D interaction formalism, which coordinates the group of user's actions on a shared environment.

Our 3D interaction assistance system is considered as a mean to bypass the limitations of some 3D interaction classical techniques in order to make them usable in different contexts.

Key Words : Virtual Reality, 3D interaction, 3D Interaction Techniques, 3D collaborative Interaction, Collaborative work, Collaborative Virtual Environments.

ملخص

الواقع الافتراضي هو تخصص يسمح للإنسان بالتفاعل مع الحاسوب (IHM)، وكذا خلق رؤية حاسوبية من أجل العمل التعاوني بمساعدة الكمبيوتر.

التفاعل ثلاثي الأبعاد هو المكون الرئيسي للواقع الافتراضي، حيث يسمح للإنسان أن يكون عنصراً فاعلاً قادراً على تغيير مسار الأحداث في بيئة اصطناعية وبالتالي التفاعل مع الكيانات الافتراضية. للتفاعل ثلاثية الأبعاد عدة تقنيات، وهي مختلفة من حيث المهام على سبيل المثال (التنقل والاختيار والمعالجة والتحكم في التطبيق). في معظم الحالات، تتعامل تقنيات التفاعل ثلاثي الأبعاد مع التفاعل بطريقة عامة ولا تفي بمتطلبات تطبيقات معينة مثل الدقة، لا سيما مع الكائنات البعيدة والمخفية وكذا الموجودة في بيئة مزدحمة.

من ناحية أخرى، فإن البيئات التعاونية هي عوالم مأهولة يتفاعل فيها المستخدمون مع بعضهم البعض لتحقيق الأهداف معاً. يرتبط تعقيد هذه البيئات بتفاعل مجموعة من المستخدمين مع الكيانات المشتركة. في هذا العمل، قمنا أولاً بتطوير تقنية تفاعل ثلاثية الأبعاد "Zoom-fwd" تعتمد على التعرف على الإيماءات كمرحلة أولى، والتي تسمح للمستخدم بتحديد و / أو التعامل بدقة مع الكائنات الافتراضية البعيدة والمخفية دون الحاجة إلى التحرك. في المرحلة الثانية، قدمنا تشكيلات تفاعلية ثلاثية الأبعاد، يقوم هذا الأخير بتنسيق إجراءات مجموعة من المستخدمين في بيئة مشتركة مع المحافظة على اتساق الأخيرة.

يمكن اعتبار أسلوب التفاعل ثلاثي الأبعاد الخاص بنا كوسيلة للتحايل على قيود المرتبطة ببعض تقنيات التفاعل ثلاثية الأبعاد الكلاسيكية، وهذا من أجل جعلها قابلة للاستخدام في سياقات مختلفة.

الكلمات الرئيسية: الواقع الافتراضي، التفاعل ثلاثي الأبعاد، تقنيات التفاعل ثلاثية الأبعاد، التفاعل ثلاثي الأبعاد التعاوني، العمل التعاوني، البيئات الافتراضية التعاونية.