



HAL
open science

Techniques d'interaction 'phygiales' pour l'exploration de systèmes informationnels complexes

Florent Cabric

► **To cite this version:**

Florent Cabric. Techniques d'interaction 'phygiales' pour l'exploration de systèmes informationnels complexes. Traitement du signal et de l'image [eess.SP]. Université Paul Sabatier - Toulouse III, 2021. Français. NNT : 2021TOU30129 . tel-03624199

HAL Id: tel-03624199

<https://theses.hal.science/tel-03624199>

Submitted on 30 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier**

**Présentée et soutenue par
Florent CABRIC**

Le 13 octobre 2021

**Techniques d'interaction 'phygiales' pour l'exploration de
systèmes informationnels complexes**

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et
Télécommunications de Toulouse**

Spécialité : **Informatique et Télécommunications**

Unité de recherche :
IRIT : Institut de Recherche en Informatique de Toulouse

Thèse dirigée par
Emmanuel DUBOIS et Christophe HURTER

Jury

M. Martin HACHET, Rapporteur
M. Thierry DUVAL, Rapporteur
M. Marcos SERRANO, Examineur
M. Pierre DRAGICEVIC, Examineur
M. Emmanuel DUBOIS, Directeur de thèse
Mme Marie-Pierre GLEIZES, Présidente

Table des matières

Chapitre 1 Introduction 1

1	Contexte.....	1
2	Objectifs de nos travaux de thèse	4
3	neOCampus : un cadre applicatif	5
4	Plan du manuscrit	6

Chapitre 2 État de l’art..... 8

1	Les dispositifs de visualisation de données situées proche d’un référent physique	9
1.1	Les systèmes de visualisation de données situées mobiles.....	10
1.1.1	Les systèmes de visualisation de données situées à l’aide de dispositifs tenus en main	11
1.1.2	Les systèmes de visualisation de données situées par réalité augmentée à l’aide de dispositifs portés sur la tête.....	13
1.2	Les systèmes de visualisation de données situées par vidéoprojection	16
2	Les techniques de sélection pour une visualisation de données situées en réalité augmentée.....	19
2.1	Sélection au contact	20
2.2	Sélection à distance.....	21
2.2.1	Sélection en l’air sans relation spatiale	22
2.2.2	Sélection en l’air avec relation spatiale.....	24
2.2.3	Sélection sans relation spatiale avec support de contact	26
2.2.4	Sélection avec relation spatiale et avec support de contact.....	28
2.3	Synthèse	29
3	Les modèles prédictifs.....	30
3.1	Les différents types de modèles pour évaluer des systèmes interactifs.....	31
3.2	Le modèle KLM de la famille des modèles GOMS	31
3.3	Extensions du modèle KLM	34
3.3.1	Les extensions de KLM pour de nouveaux dispositifs interactifs.....	34
3.3.2	Définition de nouveaux opérateurs.....	36

3.3.3	Définition des unités de temps	37
3.3.4	Les heuristiques.....	39
3.3.5	Les évaluations de modèles KLM	40
3.4	Résumé.....	40

Chapitre 3 Contexte et Outils..... 42

1	Contexte : neOCampus.....	43
1.1	Les types de données	44
1.2	Les types de référents	45
2	Outil Logiciel	46
2.1	Scénarios d’usages et exigences des maquettes interactives	47
2.1.1	Scénario 1 : Visualisation de la consommation énergétique	47
2.1.2	Scénario 2 : Maquette interactive pour l’affichage de données architecturales.....	48
2.2	Exigences pour la construction de maquette interactive.....	49
2.2.1	Modularité.....	49
2.2.2	Interactivité	49
2.3	Les maquettes physiques	50
2.3.1	Les différents types de maquettes	50
2.3.2	Propriétés des briques.....	51
2.4	Interaction et visualisation	52
2.4.1	Interaction avec la maquette à l’aide d’un lancer de rayon.....	52
2.4.2	Visualisation située, ancrée et modulaire	53
2.5	Processus de création modulaire.....	53
2.5.1	Construction physique et numérique de la maquette	54
2.5.2	Créer un modèle 3D dans Unity	55
2.5.3	Caractériser architecturalement chaque brique	55
2.5.4	Lier les données aux briques	56
2.6	Applications	56
2.6.1	Conception des techniques d’interaction pour nos applications.....	56
2.6.2	Environnement matériel et logiciel	57
2.6.3	Contrôle instantané de la consommation énergétique d’un campus ou d’un bâtiment	58
2.6.4	Support à la gestion patrimoniale	59
2.7	Synthèse	60
3	Outil Conceptuel.....	61
3.1	L’espace de conception.....	62

3.1.1	Composantes de notre espace de conception	62
3.1.2	Distance utilisateur-référent	62
3.1.3	Superposition espace moteur-visuel	63
3.1.4	Relation référent-médiateur	64
3.1.5	Localisation du feedback	65
3.2	Illustration du pouvoir descriptif de notre espace.....	66
3.2.1	Distance utilisateur-référent	66
3.2.2	Superposition de l'espace moteur-visuel	68
3.2.3	Relation référent-médiateur	68
3.2.4	Localisation du feedback	70
3.3	Illustration du pouvoir comparatif de notre espace	71
3.3.1	Interagir avec des bâtiments	72
3.3.2	Interagir avec des petits objets physiques	73
3.3.3	Interagir avec une projection sur un objet physique	75
3.3.4	Exploration de données urbaines.....	76
4	Synthèse.....	78

Chapitre 4 TouchGlass..... 80

1	Conception d'une technique de sélection depuis une vitrine	82
1.1	Problèmes et contraintes liés au lancer de rayon depuis une vitrine comme technique de sélection à distance	83
1.1.1	Effet de parallaxe.....	83
1.1.2	Manque de retour visuel	84
1.2	Facteurs influençant les techniques de sélection absolues pour la sélection d'objets distants avec une surface tactile comme support de contact	85
1.3	Synthèse	87
2	Étude informative	87
2.1	Protocole expérimental	88
2.1.1	Tâche et instructions	88
2.1.2	Participants.....	89
2.1.3	Matériel	89
2.1.4	Données recueillies	89
2.2	Résultats.....	89
2.3	Conclusion	90
3	Étude de la sélection : conception	90
3.1	Tâche et instructions	91

3.2	Conditions.....	92
3.3	Participants.....	93
3.4	Conception et procédure	93
3.5	Matériel.....	93
3.6	Données collectées et traitement des données	94
4	Résultats de l'étude de sélection	95
4.1	Taux de réussite de la tâche de sélection	95
4.2	Temps de réalisation des essais	96
4.3	Sélection au premier contact : centroïde, offset et spread	97
4.3.1	Analyse des centroïdes.....	97
4.3.2	Offset, offsetX et offsetY	98
4.3.3	Spread.....	99
4.3.4	Résumé.....	100
5	Discussion	100
5.1	Recommandations de conception	101
5.1.1	Amélioration de la sélection des cibles	101
5.1.2	Taille du rayon	101
5.1.3	Augmenter la taille virtuelle de la cible.....	102
5.2	Perspectives	102
6	Synthèse.....	102

Chapitre 5 KLM..... 104

1	Notre extension KLM pour la réalité mixte	106
1.1	Nouveaux opérateurs pour l'interaction en réalité mixte	106
1.1.1	Button Click (Bc)	107
1.1.2	Raise Hand (Rh)	107
1.1.3	Air Tap (At).....	108
1.1.4	Coarse Head Point (Chp)	108
1.1.5	Precise Head Point (Php)	109
1.2	Adaptation d'un opérateur existant : Simple Reaction (Ms) et ses heuristiques.....	109
2	Études utilisateur : définition d'une unité de temps pour les nouveaux opérateurs	111
2.1	Participants et environnement.....	111
2.2	Collecte et analyse des données.....	112
2.3	Opérateur Button Click.....	112

2.3.1	Tâche	112
2.3.2	Résultats	112
2.4	Opérateur Raise Hand	113
2.4.1	Tâche	113
2.4.2	Résultats	113
2.5	Opérateur Air Tap	114
2.5.1	Tâche	114
2.5.2	Résultats	114
2.6	Opérateur Coarse Head Point	115
2.6.1	Tâche et protocole.....	115
2.6.2	Résultats	116
2.7	Opérateur Precise Head Point	117
2.7.1	Tâche	117
2.7.2	Résultats	118
2.8	Résumé.....	119
3	Études de consolidation.....	119
3.1	Participants et environnement.....	120
3.2	Données recueillies et analyse des données.....	121
3.3	Étude de consolidation 1 : pointer hors du FoV	121
3.3.1	Plan d'étude et tâche	121
3.3.2	Modélisation de la tâche	121
3.3.3	Comparaison entre les résultats de temps mesurés expérimentalement et prédits.....	122
3.4	Étude de consolidation 2 : sélections multiples hors du FoV	123
3.4.1	Conception et tâche de l'étude	123
3.4.2	Comparaison entre les temps observés et prédits.....	124
3.5	Étude de consolidation 3 : sélections multiples dans le FoV.....	125
3.5.1	Conception de l'étude et tâche	125
3.5.2	Modélisation des tâches.....	126
3.5.3	Comparaison entre les temps observés et prédits.....	126
3.6	Résumé.....	126
4	Conclusions et Travaux Futurs.....	127

Chapitre 6 Conclusion et Perspectives..... 129

1	Résumés des contributions	130
1.1	Fonctionnelle	130

1.2	Conceptuelle	131
2	Perspectives	132
2.1	Moyen Terme.....	132
2.1.1	TouchGlass.....	132
2.1.2	KLM.....	134
2.1.3	Outils	135
2.2	Long Terme.....	135
2.3	Nouvelles solutions pour neOCampus.....	138
3	Projet de recherche	139
3.1	Support d’affichage.....	140
3.2	Objet d’intérêt	141
3.3	Instrument d’interaction.....	141
1	Gestion de la consommation énergétique.....	167
1.1	Au niveau du campus.....	167
1.2	Au niveau d’un laboratoire de recherche	168
2	Observation des données de capteurs installés dans des salles de classe.....	168
3	Bibliothèque universitaire	169
3.1	Faire prendre conscience aux étudiants du bruit.....	170
3.2	Gestion du bruit dans la bibliothèque par le personnel.....	171

Table des Illustrations

Figure 1.1 – Exemple de représentations situées (gauche) et ancrées (droite).....	2
Figure 1.2 – Représentation de données situées (gauche) et ancrées (droite) avec comme référent physique une maquette.....	3
Figure 2.1 – Exemples de visualisation de données situées à l'aide (A) d'un casque porté sur la tête (Guarese et al., 2020), (B) d'un dispositif tenu en main (N. ElSayed et al., 2015) et (C) par vidéoprojection (Millar et al., 2018)	10
Figure 2.2 – Prototype proposé par ElSayed et collab. avec (A) la fonction de filtre, (B) de recherche, (C) de classement et (D) le prototype de l'interface (N. A. M. ElSayed et al., 2016)	11
Figure 2.3 – Exemples de visualisation de données situées dans des environnements extérieurs avec (A) SiteLens (White & Feiner, 2009), (B) une visualisation sur des ruches (Engelke et al., 2016) et (C) une visualisation conçue pour des agriculteurs (Zheng & Campbell, 2019).....	12
Figure 2.4 – Différents systèmes de visualisation de données situées avec un dispositif porté à la main pour l'apprentissage de (A) circuit électrique (Beheshti et al., 2017) , (B) la conduction thermique (Knierim et al., 2020) et (C) pour faciliter la communication entre un instructeur et un étudiant (Villanueva et al., 2020)	12
Figure 2.5 – Les différentes visualisations proposées par l'AR-CANVAS (Bach et al., 2017)13	
Figure 2.6 – Exploration du « <i>information-seeking mantra</i> » pour des données situées à l'aide d'un casque HoloLens (Caggianese et al., 2019) avec (A) la vue d'ensemble et (B) la manipulation directe pour avoir des détails sur l'information	14
Figure 2.7 – Exemple applicatif pour la recherche d'information à partir d'objets présents dans l'environnement : la recherche d'une recette de cuisine avec (A) la vue finale et (B) la sélection d'un aliment à ajouter à la recette recherchée	15
Figure 2.8 – Les différentes visualisations de données situées pour faciliter l'accès à un siège dans une pièce (Guarese et al., 2020).....	15
Figure 2.9 – Différents systèmes de visualisation de données situées sur un relief physique avec (A) <i>Illuminating Clay</i> (Piper et al., 2002), (B) <i>Tangible Landscape</i> (Millar et al., 2018) et (C) une visualisation d'écoulement de l'eau et son impact sur le relief (Harmon et al., 2016)	16

Figure 2.10 – Des systèmes de visualisations situées pour (A) la consommation énergétique (Kirshenbaum et al., 2020) , (B) le bien-être (Roo et al., 2017) et (C) l’apprentissage et la découverte d’un relief (Priestnall & Cheverst, 2019).....	17
Figure 2.11 – Différents exemples de projet <i>CityScope</i> avec (gauche A) la ville d'Hanoï, (centre B) la ville d'Andorre et (droite C) le site de Volpe	19
Figure 2.12 – VirtualPads (Andujar & Argelaguet, 2007) avec (A) le concept de technique de sélection directe et indirecte et (B) le prototype avec les fenêtres et le pavé virtuel en vert ...	23
Figure 2.13 – L’espace de conception de Arpads (Brasier et al., 2020)	23
Figure 2.14 – La technique de sélection par lancer de rayon proposée par Mine (Mine, 1995)	24
Figure 2.15 – Les quatre techniques de sélection de Teather et Stuerzlinger (Teather & Stuerzlinger, 2013).....	25
Figure 2.16 – Les deux dispositifs étudiés par Pham et Stuerzlinger (Pham & Stuerzlinger, 2019).....	26
Figure 2.17 – Les trois emplacements du pavé tactile évalués dans (Gugenheimer et al., 2016)	27
Figure 2.18 – L’espace de conception des interactions de (Saidi et al., 2019)	28
Figure 2.19 – Interaction directe avec un support de contact dans un magasin avec (A) le prototype du système et (B) le système de classement des produits en fonction d’une donnée	28
Figure 2.20 – Illustration du besoin spécifique de création d’un modèle KLM pour l’interaction avec un téléphone mobile avec (A) une utilisation dans un environnement public et (B) la spécificité du téléphone mobile avec trois zones d’interaction (Holleis et al., 2007).....	35
Figure 2.21 – Illustration de l’opérateur Turn de (Schneegaß et al., 2011)	36
Figure 2.22 – Les opérateurs (a) Tapping, (b) Pointing et (c) Dragging de l’extension KLM proposée par (A. Lee et al., 2015)	38
Figure 3.1 – Les trois types de maquettes d'architecture étudiés	51
Figure 3.2 – Les deux lancers de rayon définissant un curseur dans le prototype iSith (Wyss et al., 2006).....	53

Figure 3.3 – Ensemble de processus de création à l'aide de notre outil logiciel	54
Figure 3.4 – Construction d'une maquette numérique et d'une maquette physique	54
Figure 3.5 – Importation de la maquette numérique dans Unity	55
Figure 3.6 – Illustration de notre script permettant de caractériser une brique.....	56
Figure 3.7 – Illustrations des visualisations de données situées sur une maquette avec (A) les techniques de sélection basées sur une vitrine et (B) les différentes visualisations ancrées et situées sur la maquette.....	57
Figure 3.8 – L'environnement matériel pour notre maquette augmentée	58
Figure 3.9 – (A) La maquette 2,5D avec une comparaison de deux bâtiments et (B) la maquette 3D fermée avec la température d'une salle visualisée	59
Figure 3.10 – La maquette 3D ouverte avec un dispositif lumineux sélectionné à l'aide de deux lancers de rayon.....	60
Figure 3.11 – Les quatre distances proposées par Hall (E. T. Hall, 1966). Image récupérée de la page Wikipédia - Proxemics	63
Figure 3.12 – Exemple du continuum de visualisation de données situées (Figure 10 (Willett et al., 2017))	64
Figure 3.13 – Continuum de la relation référent-médiateur.....	65
Figure 3.14 – Illustration (gauche) d'un exemple de système interactif de visualisation de données situées à une distance sociale avec un espace moteur-visuel dispersé , un médiateur ayant une relation de type extension virtuelle , un retour visuel affiché indépendamment du référent physique et des données ancrées et (droite) les propriétés de notre outil conceptuel	66
Figure 3.15 – Illustration de la distance utilisateur-référent avec (A) une distance intime , (B) une distance personnelle , (C) une distance sociale et (D) une distance publique	67
Figure 3.16 – Illustration de la superposition de l'espace moteur-visuel avec (A) un espace moteur-visuel superposé et (B) un espace moteur-visuel dispersé	68
Figure 3.17 – Illustration de la relation référent-médiateur avec (A) une relation physique pure , (B) une extension virtuelle , (C) une extension physique et (D) une relation virtuelle pure	69

Figure 3.18 – Illustration de la propriété de localisation du feedback avec (A) pas de feedback, (B) un feedback indépendant du référent physique, (C) un feedback autour du référent physique et (D) un feedback sur le référent physique	70
Figure 3.19 – Les cinq propriétés de notre outil conceptuel	71
Figure 3.20 – Les systèmes interactifs sur un bâtiment avec (A) MobiSpray (Scheible & Ojala, 2009) et (B) les travaux de Böhmer (Böhmer et al., 2011)	72
Figure 3.21 – Comparaison de systèmes interactifs sur un bâtiment avec (orange) MobiSpray (Scheible & Ojala, 2009) et (vert) les travaux de Böhmer (Böhmer et al., 2011).....	73
Figure 3.22 – Les systèmes interactifs sur des petits objets (A) TouchGlass (Cabric et al., 2019) et (B) Point-and-shake (Freeman et al., 2018)	74
Figure 3.23 – Comparaison de systèmes interactifs sur des petits objets (A) TouchGlass (Cabric et al., 2019) et (B) Point-and-Shake (Freeman et al., 2018)	74
Figure 3.24 – Les systèmes de projection sur des objets physiques avec (A) Pmomo (Zhou et al., 2016) et (B) Tangible Landscape (Millar et al., 2018).....	75
Figure 3.25 – Comparaison de systèmes de projection sur des objets avec (jaune) Tangible Landscape (Millar et al., 2018) et (rose) Pmomo (Zhou et al., 2016).....	76
Figure 3.26 – Les systèmes d’exploration de données urbaines avec (A) CityScope (Alonso et al., 2018) et (B) TangibleMap (Walker et al., 2017).....	77
Figure 3.27 – Comparaison de systèmes d’exploration de données urbaines avec (en orange) le projet CityScope (Alonso et al., 2018) et (en bleu) le projet Tangible Map (Walker et al., 2017)	77
Figure 4.1 – Exemples d’objets physiques protégés par une vitrine avec des données textuelles : (A) la synagogue de Vienne-Leopoldstadt au musée juif autrichien d’Eisenstadt et (B) la synagogue de Seesen au Städtischen Museum de Seesen (images récupérées de Wikimedia)	81
Figure 4.2 – (A) Explication et (B) exemple de parallaxe tirés de (Khamis et al., 2018) avec en rouge la zone perçue par l’utilisateur et en vert la zone réellement sélectionnée	84
Figure 4.3 – Exemples d’interaction sans retour visuel avec de gauche à droite, (A) Dynamic Shader Lamps (Bandyopadhyay et al., 2001), (B) sélection avec des pointeurs laser (Myers et al., 2002) et (c) sélection de sphères physiques en lévitation (Freeman et al., 2018).....	85

Figure 4.4 – Le concept d’HaptiCase (Corsten et al., 2015), avec (à gauche) la tâche a réalisé dans l’étude et (à droite) la présentation du dispositif avec les repères spatiaux placés derrière l’écran.....	86
Figure 4.5 – Les trois types de lancers de rayon avec (A) le rayon orienté selon l'orientation du doigt, (B) un lancer de rayon oeil-doigt et (C) un lancer orthogonal. Une vue d'ensemble du dispositif utilisé pour l'étude (D).....	88
Figure 4.6 – Pourcentage de participants ayant utilisé chaque type de lancer de rayon : lancer de rayon orienté (gauche), lancer de rayon œil-doigt (centre) et lancer de rayon orthogonal (droite).....	90
Figure 4.7 – Matériel utilisé de notre expérimentation	91
Figure 4.8 – Illustration des paramètres utilisés pour analyser les premiers contacts.....	94
Figure 4.9 – Taux de réussite moyen en % (avec IC 95 %) selon la taille de la cible et la distance à la vitrine.....	96
Figure 4.10 – Temps moyen de réalisation en ms (avec IC à 95%) en fonction de la taille de la cible et de la distance à la vitrine	97
Figure 4.11 – Distribution de tous les centroïdes de sélection au premier contact autour des cibles. La plupart d'entre eux se trouvent dans la zone supérieure gauche	97
Figure 4.12 – Offset moyen (gauche), offsetX (centre) et offsetY (droite) en cm (avec IC 95 %) en fonction de la distance de la cible par rapport au verre	98
Figure 4.13 – OffsetX (gauche) et offset Y(droite) moyens en cm (avec 95% Cis) selon la condition de référence spatiale.....	98
Figure 4.14 – Distribution des centroïdes en fonction de la référence spatiale : avec référence spatiale (gauche) ou sans référence spatiale (droite). Le point rouge correspond à la moyenne des centroïdes	99
Figure 4.15 – Spread moyen en cm (avec 95% Cis) par distance cible (à gauche) et par référence spatiale (à droite).....	100
Figure 5.1 – Illustration de l'opérateur Button Click (Bc).....	107
Figure 5.2 – Illustration de l'opérateur Raise Hand (Rh).....	107
Figure 5.3 – Illustration de l'opérateur Air Tap (At).....	108

Figure 5.4 – Illustration de l'opérateur Coarse Head Point (Chp)	108
Figure 5.5 – Illustration de l'opérateur Precise Head Point (Php)	109
Figure 5.6 – Temps de réalisation des 4 Button Click (en ms) avec les IC à 95%	113
Figure 5.7 – Temps de réalisation des 4 Air Tap (en ms) avec les IC à 95%	114
Figure 5.8 – Les différentes conditions de l'étude de l'opérateur Coarse Head Point (Chp) avec (en rouge) les cibles à 45°, (en vert) les cibles à 60°, (en rose) les cibles à 75°, (en blanc) le compteur permettant de débiter la tâche et (en bleu) le cercle inscrit dans le champ de vision de l'HoloLens	115
Figure 5.9 – Les temps de réalisation pour la tâche de <i>Coarse Head Point</i> avec (carrés bleus) la direction verticale, (triangles verts) la direction horizontale et (cercles rouges) la direction diagonale	117
Figure 5.10 – Configuration pour la tâche Precise Head Point avec le champ de vision de l'HoloLens et les cibles (en blanc)	118
Figure 5.11 – Les temps de réalisation de la tâche de Precise Head Point avec les IC à 95%	118
Figure 5.12 – Configuration d'un essai pour la tâche de sélection en dehors du champ de vision avec (à gauche) l'utilisateur sélectionne (Bc) la première cible placée au centre du champ visuel initial, (au centre) l'utilisateur tourne la tête de 45° vers la gauche (Chp) pour pointer la deuxième cible (Php) et (à droite) l'utilisateur sélectionne la cible (Bc) avec le dispositif dédié	123
Figure 5.13 – Configuration d'un essai de sélection à l'intérieur du champ visuel avec (à gauche) le participant pointe la première cible avec l'HoloLens, (au centre) le participant lève sa main et son doigt pour effectuer la validation Air Tap et (à droite) la sélection de la cinquième (dernière) cible et (à droite) la sélection de la cinquième (la dernière) cible. L'essai comprend deux cibles placées au sud-est (cible numéro deux et cible numéro trois) et deux cibles placées au sud-est (cible numéro trois)	125

Table des Tableaux

Tableau 1 – Propriétés des sélections à distance	22
Tableau 2 – Les opérateurs originaux de KLM.....	33
Tableau 3 – Temps de réalisation de la tâche Coarse Head Point (en ms) avec les IC à 95% selon la direction (d) et la distance angulaire (a) : Chp(d,a)	117
Tableau 4 – Notre Extension de KLM	119
Tableau 5 – Les opérateurs impliqués dans les 3 études de consolidation	120
Tableau 6 – Temps de réalisation en ms (TR) avec les IC à 95 %, le temps prédit par notre modèle (TP) et la différence (en ms et en pourcentage) entre le temps moyen de réalisation et le temps prédit par le modèle	122

Chapitre 1 Introduction

Contenu du Chapitre

1. Contexte.....	1
2. Objectifs de nos travaux de thèse	4
3 neOCampus : un cadre applicatif	5
4 Plan du manuscrit	6

1 Contexte

Durant cette dernière décennie, la quantité de données n'a cessé d'augmenter. D'une part, la démocratisation de capteurs dans les smartphones et les smartwatches permet de capter un nombre conséquent de données à tout moment. Par exemple, le Samsung Galaxy s10 (2019) contient 16 capteurs et la dernière Apple Watch 6 permet de mesurer le rythme cardiaque, la température ou encore le taux d'oxygène dans le sang. De plus, le nombre de ces dispositifs est en constante augmentation (respectivement 3,2 milliards d'utilisateurs de smartphone (O'Dea, 2020) et 66,5 millions de smartwatches sur le marché en 2019 (Statista Research Department, 2021)).

D'autre part, de plus en plus d'objets du quotidien sont dotés de capteurs permettant de transmettre des informations. Ces objets connectés ont permis l'émergence de l'Internet des Objects (IdO ou Internet of Things). Par exemple, le thermostat d'un appartement peut envoyer la température courante dans une pièce et l'utilisateur ou le propriétaire peuvent la consulter.

Selon l'entreprise Statista¹, 25 milliards d'objets connectés produisant des données devraient équiper nos villes, maisons ou lieux de travail à l'horizon 2030 (Holst, 2021).

Ces données permettent aux utilisateurs de comprendre et d'analyser leur santé, leur mode de vie ou encore leur consommation énergétique, et aux différentes institutions (villes, campus, promoteur immobilier) de gérer le flux de population, le trafic routier ou les pertes énergétiques dans les bâtiments administrés. La collecte et l'analyse de ces données sont donc primordiales pour toutes les administrations afin de promouvoir les villes de demain.

Un domaine spécifique qui en tire particulièrement profit est le domaine de «la ville intelligente» (terme introduit par Dupuy (Dupuy, 1992)). Depuis, plusieurs définitions sont apparues permettant de définir les besoins et les objectifs d'une ville intelligente. Chourabi et collab. (Chourabi et al., 2012) propose huit éléments permettant de définir une ville intelligente : gestion et organisation, technologie, gouvernance, contexte politique, personnes et communautés, économie, infrastructure déjà bâtie et environnement naturel. Une ville intelligente est composée de capteurs permettant de collecter des données en temps réels, contribuant ainsi à la gestion de différents services de la ville. Néanmoins entre la collecte et la décision, il est nécessaire de stocker ces données, de les traiter, et qu'un ou plusieurs acteurs organisent et manipulent ces données hétérogènes pour produire une analyse. Outre la capture, le transfert et le stockage, il est donc capital de mettre à disposition des visualisations interactives de ces données pour pouvoir faire émerger des connaissances et faciliter la prise de décision (Lavalle et al., 2020).

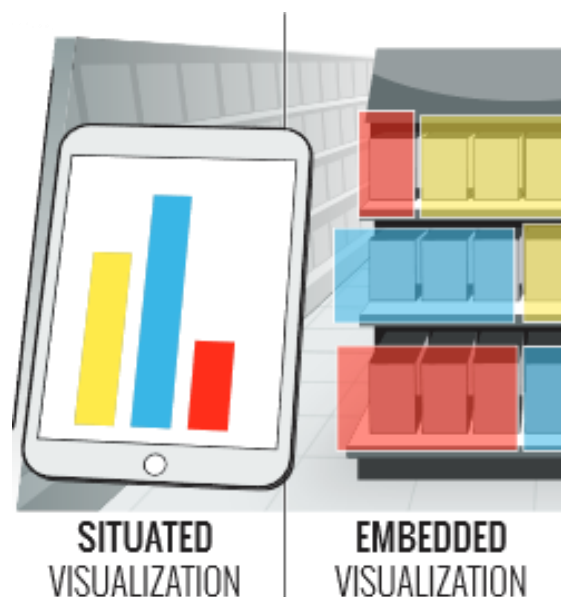


Figure 1.1 – Exemple de représentations situées (gauche) et ancrées (droite)

¹ <https://fr.statista.com/>

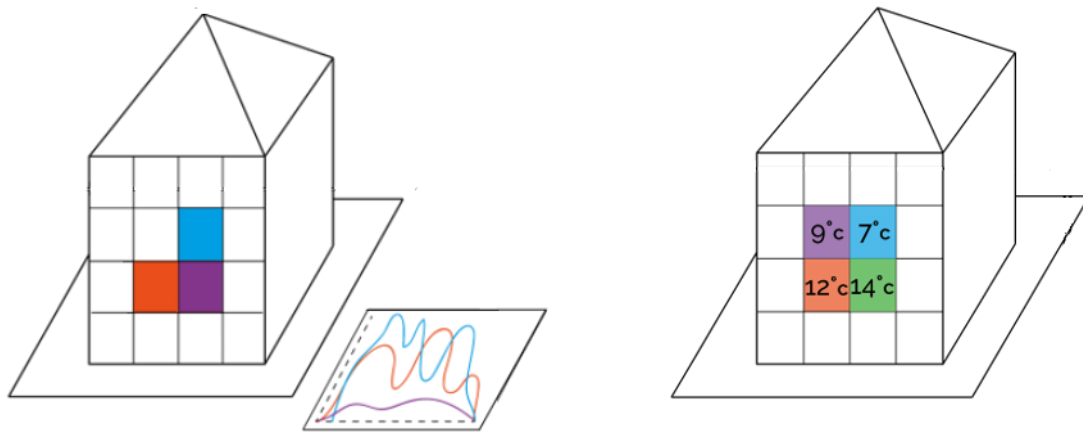


Figure 1.2 – Représentation de données situées (gauche) et ancrées (droite) avec comme référent physique une maquette

Une approche récente propose de représenter ces données proches de l’objet qui les capte ou les produit ; on parle alors de **représentation située** (Willett et al., 2017). Par exemple, visualiser la quantité de produits restant dans un stock d’un magasin sur une tablette à l’intérieur de l’entrepôt est une représentation située. Si les données sont représentées sur l’objet d’où elles proviennent (e.g. l’étagère où est stocké l’objet) on parle de **représentation ancrée** (cf. Figure 1.1 (Willett et al., 2017)). L’objet ou lieu d’acquisition des données est alors appelé **réfèrent physique**.

À l’échelle d’une ville ou d’un campus, représenter de manière située des données revient à représenter les données à proximité ou sur une représentation miniaturisée du référent physique ; on parle alors de *facsimile* (Willett et al., 2017 cf. Figure 1.2). Ce domaine prometteur permet aux utilisateurs de mieux comprendre ces données (ElSayed et al., 2015; Lobo & Christophe, 2020; S. White et al., 2006).

Cependant, ce domaine est encore assez récent et il reste de nombreux défis à explorer. Un des défis majeurs identifiés par Thomas et collab. (Thomas et al., 2018) est le développement et la conception de techniques d’interaction pour faciliter l’analyse de données situées. Il apparaît donc primordial de compléter cet essor des représentations de données situées par des solutions interactives permettant leur exploitation. La présence d’un référent physique proche de la représentation de ces données ouvre des perspectives de solutions multiples. En effet, rendre interactif ce référent constituerait un nouvel espace physique d’interaction étendant les capacités d’interaction disponibles pour manipuler, explorer ou déclencher une représentation de données numériques.

2 Objectifs de nos travaux de thèse

Nos travaux de thèse se situent donc à l'intersection de trois espaces : 1) l'espace numérique des données, 2) l'espace physique du référent physique et 3) l'espace des possibilités d'interaction constitué par les capacités d'interaction de l'utilisateur et des dispositifs disponibles.

Dans ce contexte, l'objectif de mes travaux de thèse est de concevoir et d'évaluer des solutions interactives tirant profit du référent physique pour interagir avec des données numériques situées. Ces travaux se focalisent sur une des interactions fondamentales de l'Interaction Humain-Machine : la sélection. Cette sélection se trouvant à la fois dans le monde physique et le monde numérique nous parlons alors de sélection **phygitale**².

En effet, parmi toutes les interactions possibles, la sélection d'un objet reste une tâche fondamentale qui prévaut sur toutes les autres (Argelaguet & Andujar, 2013). Dans le domaine de la représentation de données situées, cette sélection peut être répartie entre des actions sur la représentation numérique des données et d'autres sur le référent physique. Par exemple, l'utilisateur peut sélectionner l'étage d'une maquette physique en le touchant physiquement ou en le pointant à distance, déclencher la visualisation de la température, modifier la température souhaitée à l'étage en glissant son doigt sur l'étage en question, ou modifier le mode d'affichage (instantané ou historique). L'utilisateur peut aussi vouloir interagir plus précisément avec les données affichées dans l'environnement numérique. La sélection d'une des données peut alors se faire dans cette représentation numérique des données par un geste en l'air.

Les travaux présentés dans ce manuscrit ont donc pour but de concevoir et d'évaluer des moyens d'interaction pour sélectionner un élément soit sur le référent physique soit dans les données numériques associées. Cela donne lieu à deux objectifs de thèse :

1. D'une part, interagir avec un référent physique nécessite de prendre en compte les contraintes dans lesquelles ce référent physique, généralement inaccessible au toucher, est présenté aux utilisateurs (p. ex. une maquette d'aéroport ou une pièce de musée sous une vitrine). Notre premier objectif est donc de concevoir des solutions permettant d'interagir à distance avec ces objets physiques.

² Mot-valise dérivé de l'anglais *physical* (physique) et *digital* (numérique).

2. D'autre part, concevoir et évaluer des systèmes immersifs permettant d'interagir avec des données numériques situées est coûteux en temps. Pouvoir prédire le temps de réalisation d'une tâche est une approche bien connue dans le domaine de l'Interaction Humain-Machine (Card et al., 1980). Ces modèles permettent de prédire, à partir d'une séquence d'action des utilisateurs, le temps total d'une tâche. Il est alors possible de comparer plusieurs solutions d'interaction pour évaluer laquelle est la plus utilisable en amont de tout processus de développement. Notre deuxième objectif est donc de concevoir un modèle prédictif pour estimer le temps de complétion d'une tâche de sélection de données numériques situées.

3 neOCampus : un cadre applicatif

Cette thèse, financée conjointement par l'opération neOCampus et la région Occitanie au travers du projet IDPHYx, trouve un ancrage naturel dans les problématiques liées à cette opération. En effet, l'opération neOCampus repose largement sur la collecte et l'exploitation de données. Ce projet multidisciplinaire, lancé en juin 2013 à l'Université de Toulouse, implique 11 laboratoires de recherche de différents domaines. L'objectif du projet est d'améliorer le confort de la vie quotidienne des utilisateurs de l'université (étudiants, professeurs et personnel), tout en diminuant son empreinte écologique en réduisant les ressources de fonctionnement (eau, électricité, etc.). Le projet encourage le travail de recherche, en offrant sur le campus de l'université une plate-forme pour des expériences innovantes, réalisées à grande échelle et *in situ*. Le projet vise à atteindre ses objectifs en tirant parti de la prolifération de dispositifs connectés peu coûteux. L'approche consiste à créer un « campus intelligent » qui permettrait de connecter non seulement des capteurs et des dispositifs intelligents à des positions fixes dans l'université, mais aussi les dispositifs mobiles personnels des usagers (smartphones, tablettes, etc.). Les données produites par une telle approche sont hétérogènes et les visualiser et interagir avec elles proches de leurs référents physiques nécessitent des outils d'interaction et de conception adaptées.

Ce cadre a donc largement contribué à l'émergence de nos contributions et nous servira donc tout au long de l'exposé de nos contributions comme support à illustrations.

4 Plan du manuscrit

Pour répondre aux objectifs de la thèse, nous commençons par présenter le cadre théorique de cette thèse (Chapitre État de l'art). Nous présentons dans un premier temps les différents dispositifs permettant de visualiser des données situées selon la technologie utilisée. Nous mettons en exergue deux technologies particulièrement utilisées : la réalité augmentée basée sur des dispositifs mobiles (smartphones ou casques) et la réalité augmentée spatiale. Nous détaillons au travers d'exemples, différents contextes d'utilisation et les contraintes liées à chaque technologie (p. ex. manque de mobilité). Nous présentons ensuite les techniques d'interaction pour sélectionner un élément d'un système de visualisation de données situées (c'est-à-dire soit le référent physique soit les données numériques). Nous nous concentrons en particulier sur les interactions à distance, car elles permettent de pallier certains problèmes des interactions au contact (p. ex. l'occultation). Ces interactions à distance sont présentées selon deux propriétés : la zone où est effectuée l'action de l'utilisateur et la présence d'une relation spatiale entre le dispositif et le référent physique ou les données numériques. Malheureusement, pour concevoir et évaluer des systèmes interagissant avec un modèle physique doit prendre en compte de nombreux paramètres. Une approche possible pour réduire la complexité de la conception et de l'évaluation de ces systèmes est l'utilisation de modèles permettant de prédire le temps de réalisation d'une tâche. Nous étudions, dans la dernière partie de l'état de l'art, les modèles prédictifs facilitant la conception et l'évaluation de systèmes interactifs pour interagir avec des visualisations de données situées.

Par la suite (Chapitre Contexte et Outils), nous présentons différents outils technologique et conceptuel que nous avons pu mettre en place pour étudier les caractéristiques d'un système de visualisation de données situées. Nous présentons dans un premier temps un outil technologique basé sur Unity et des logiciels tiers (Lego Digital Designer³ et LeoCad⁴), opérationnalisant le lien entre des données numériques, un modèle 3D numérique et un modèle 3D physique. Dans un deuxième temps nous proposons un cadre conceptuel permettant de caractériser un système interactif avec un objet physique servant de référent. Ce cadre conceptuel se concentre sur

³ <https://www.lego.com/en-us/ldd>

⁴ <https://www.leocad.org/>

l'interaction avec ce référent en explorant plusieurs dimensions comme, par exemple, la proximité entre l'utilisateur et le référent.

Ces deux outils nous ont permis de cerner les différentes composantes d'une technique d'interaction en faisant notamment émerger différentes problématiques d'interaction inhérentes à l'environnement dans lequel le référent physique et ses données sont placés. En effet, la technique d'interaction doit s'adapter à plusieurs facteurs tels que, la technologie utilisée pour afficher les données numériques, l'outil permettant de faire le lien entre le référent physique intrinsèquement non interactif et l'utilisateur, etc.

Dans le chapitre suivant (Chapitre TouchGlass), nous présentons une technique d'interaction permettant d'interagir avec un référent physique au travers d'une surface tactile semi-transparente pour visualiser des données situées à côté de l'objet par vidéoprojection. Nous soulignons l'importance d'avoir une interaction : 1) « *walk-up and use* » (c'est-à-dire qui ne demande pas d'être apprise et qui permet d'interagir efficacement avec le système) afin que l'utilisateur puisse réussir sa tâche de sélection le plus précisément et rapidement possible et 2) adaptée aux contraintes de la présence d'un référent physique placé dans un contexte public. Au travers de deux évaluations (*in situ* et en laboratoire), nous évaluons la précision d'une technique de sélection basée sur une vitrine selon plusieurs distances et tailles d'objet. Nous présentons les résultats selon la précision, le temps de sélection et les erreurs de sélection. Ces résultats nous permettent de faire émerger des conseils de conception pour des systèmes interactifs permettant de visualiser des données situées dans un contexte public.

Enfin, dans le chapitre suivant nous présentons une étude sur l'interaction avec des données situées (Chapitre KLM). Pour pallier le problème de la conception et de l'évaluation avec des utilisateurs d'un système de visualisation de données situées, la contribution originale présentée dans ce chapitre KLM consiste en l'élaboration d'un modèle de prédiction du temps de réalisation de tâche de sélection dans les applications de réalité mixte. Ce modèle est conçu dans le but d'aider les futurs concepteurs à évaluer le temps que devrait mettre l'utilisateur final pour réaliser sa tâche. Pour créer ce modèle plusieurs études utilisateur ont été conduites pour définir les différents paramètres de ce modèle et le valider.

Nous concluons ce manuscrit par un résumé des contributions, puis nous élaborons plusieurs perspectives de travaux futurs permettant d'étendre ces travaux.

Chapitre 2 État de l'art

Contenu du Chapitre

1. Les dispositifs de visualisation de données situées proche d'un référent physique	9
1.1 Les systèmes de visualisation de données situées mobiles	10
1.2 Les systèmes de visualisation de données situées par vidéoprojection	16
2 Les techniques de sélection pour une visualisation de données situées en réalité augmentée	19
2.1 Sélection au contact	20
2.2 Sélection à distance	21
2.3 Synthèse	29
3 Les modèles prédictifs	30
3.1 Les différents types de modèles pour évaluer des systèmes interactifs	31
3.2 Le modèle KLM de la famille des modèles GOMS	31
3.3 Extensions du modèle KLM	34
3.4 Résumé	40

La sélection d'un élément dans une visualisation de données situées, donc proche d'un objet physique utilisé comme référent, est une tâche primordiale à étudier. En effet, cette sélection peut permettre de déclencher de nouvelles visualisations, de changer le mode d'affichage des données (p. ex., historique ou temps réel) ou encore permettre d'avoir une vue plus détaillée de certaines données. Dans cet état de l'art, nous présentons en premier lieu une vue d'ensemble des différentes technologies permettant de visualiser des données situées proches d'un référent physique. Cette partie de l'état de l'art permet de faire ressortir deux technologies majeures pour visualiser des données situées : 1) la réalité augmentée par vidéoprojection (ou réalité augmentée spatiale) et 2) la réalité augmentée mobile (c'est-à-dire soit à l'aide d'un dispositif

tenu en main soit un dispositif porté sur la tête). Nous détaillons les problématiques soulevées par chacune de ces approches.

Nous présentons ensuite les techniques d'interaction possibles pour interagir avec un système de visualisation de données situées (c'est-à-dire soit avec le référent physique soit avec les données numériques), en particulier les interactions à distance. Les avantages et limitations nous amènent à ajouter une troisième dimension à notre état de l'art. En effet, concevoir, évaluer et comparer un ou plusieurs systèmes interactifs de visualisation de données situées est complexe. Au travers d'une revue de la littérature des évaluations basées sur des modèles de prédiction du temps d'interaction, nous mettons en avant le potentiel de cette approche. Cette approche permet, en amont, de décrire et évaluer des systèmes et ainsi déceler des problèmes d'utilisabilité, en particulier sur les temps de réalisation des tâches.

1 Les dispositifs de visualisation de données situées proche d'un référent physique

La conception des systèmes de visualisation de données situées proche d'un référent physique nécessite de concevoir des solutions se trouvant à la fois dans la réalité physique et dans la réalité numérique. Cet espace entre physique et numérique est conceptualisé par Milgram et collab. en 1995 (Milgram et al., 1995) pour caractériser les différentes formes de rendu sur un dispositif d'affichage : la Réalité Mixte, terme utilisé depuis pour désigner aussi les interactions en entrée se situant entre les deux réalités. Depuis, de nombreuses approches se réapproprient et affinent ce concept : Virtualité Augmentée (Choo et al., 2019; Dubois et al., 2002), Réalité Remixée (Lindlbauer & Wilson, 2018), Réalité Augmentée (Azuma, 1997), Réalité diminuée (Kawai et al., 2016) ou Réalité Étendue (J. Kim et al., 2019). Au travers d'une revue de la littérature et d'entrevues avec 68 experts, sept aspects ont été identifiés pour définir un système comme un système de réalité mixte (Speicher et al., 2019) : le nombre d'environnements, le nombre d'utilisateurs, le degré d'immersion, le degré de virtualité, le degré d'interactivité, le type d'interaction en entrée et le type de représentation en sortie. Par exemple, le degré de virtualité propose de définir la réalité mixte, en partie, par la présence de contenu numérique et de contenu physique dans l'application. Le cas de la visualisation de données situées entre précisément dans ce cadre, car il présente des informations numériques dans un environnement physique.

Néanmoins, différentes technologies peuvent être utilisées pour faire de la réalité mixte. Azuma et collab. (Azuma et al., 2001) identifient trois technologies (illustrés en Figure 2.1) permettant de faire de la réalité augmentée et donc pouvant être utilisées pour des outils de visualisations de données situées :

- Les casques portés sur la tête (Head-Mounted Displays ou Head-Worn Displays, p. ex. (Guarese et al., 2020)),
- Les dispositifs tenus en main (Handheld, p. ex. (ElSayed et al., 2015)),
- La réalité augmentée par vidéoprojection (réalité augmentée spatiale ou Spatial Augmented Reality, p. ex. (Millar et al., 2018)).

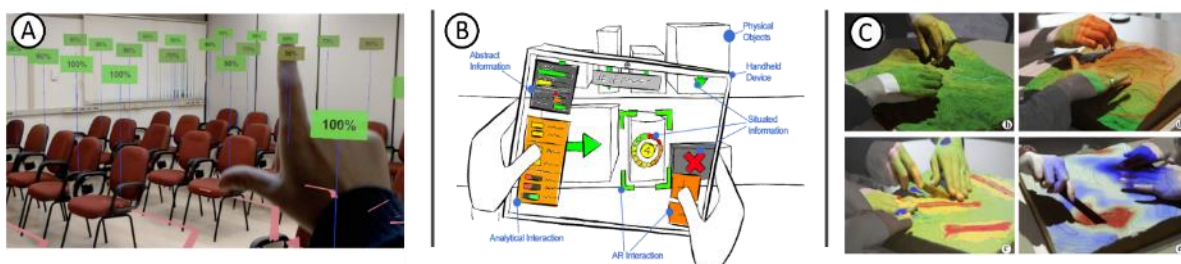


Figure 2.1 – Exemples de visualisation de données situées à l'aide (A) d'un casque porté sur la tête (Guarese et al., 2020), (B) d'un dispositif tenu en main (ElSayed et al., 2015) et (C) par vidéoprojection (Millar et al., 2018)

Ces technologies n'ont cependant pas les mêmes propriétés d'interaction et de visualisation. En effet, les systèmes de visualisation de données situées utilisant des dispositifs mobiles (portés sur la tête ou tenus en main) ne voient l'espace numérique qu'au travers de l'écran présent sur le dispositif. Les systèmes de réalité augmentée par vidéoprojection sont eux dépendants des outils de vidéoprojection qui ne sont pas fournis avec des outils interactifs.

Cette section de l'état de l'art est donc décomposée en deux parties : 1) les systèmes de visualisation de données situées par réalité augmentée mobile et 2) les systèmes de visualisation de données situées par vidéoprojection.

1.1 Les systèmes de visualisation de données situées mobiles

Les systèmes de visualisation de données situées utilisant des dispositifs mobiles n'affichent l'espace numérique qu'au travers de l'écran présent sur le dispositif. L'utilisateur déplace alors l'écran pour parcourir et visualiser l'entièreté de l'espace numérique. Nous détaillons dans un premier temps les systèmes de visualisation de données situées à l'aide de dispositifs tenus en main puis les systèmes de visualisation de données situées à l'aide de dispositifs portés sur la tête.

1.1.1 Les systèmes de visualisation de données situées à l'aide de dispositifs tenus en main

Pour les dispositifs tenus en main comme les smartphones ou les tablettes, la visualisation se fait généralement à l'aide de l'écran du dispositif et de la caméra placée à l'arrière permettant un rendu vidéo en direct. De nombreux contextes ont utilisé la visualisation de données situées par réalité augmentée avec un smartphone ou une tablette.

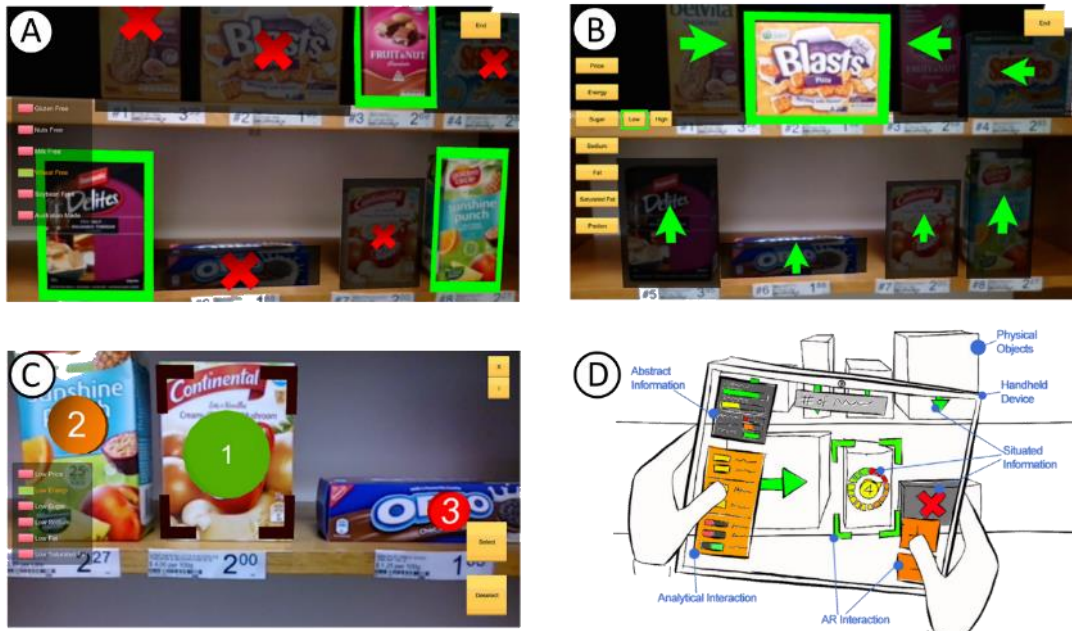


Figure 2.2 – Prototype proposé par ElSayed et collab. avec (A) la fonction de filtre, (B) de recherche, (C) de classement et (D) le prototype de l'interface (N. A. M. ElSayed et al., 2016)

ElSayed et collab. (ElSayed et al., 2016) visualisent des données dans un contexte d'achat de produits dans un magasin. Les utilisateurs peuvent scanner des produits à l'aide de la caméra de la tablette pour obtenir diverses informations comme le prix, le taux de sucre ou la note moyenne des utilisateurs. Afin de faciliter la visualisation des données et de permettre à l'utilisateur de trouver plus facilement ce qu'il recherche, l'utilisateur peut aussi : 1) filtrer les référents physiques, c'est-à-dire atténuer ou mettre en évidence certains produits en rayon (cf. Figure 2.2 A), 2) trouver un référent physique parmi d'autres (cf. Figure 2.2 B), ou encore 3) classer les référents physiques en fonction des données associées (c'est-à-dire un nombre représentant le classement selon un ou plusieurs critères apparaît sur le produit cf. Figure 2.2 C).

Le smartphone ou la tablette étant des dispositifs facilement transportables et toujours accessibles, les chercheurs ont aussi utilisé la visualisation de données situées pour des environnements en extérieur. Par exemple, le projet SiteLens (Sean White & Feiner, 2009) est conçu dans le but de visualiser des données de monoxyde de carbone en fonction de la position de l'utilisateur dans la rue (cf. Figure 2.3 A). Engelke et collab. (Engelke et al., 2016), proposent

un système de visualisation de données sur des ruches pour analyser les abeilles qui les composent (cf. Figure 2.3 B). Enfin, Zheng et Campbell (Zheng & Campbell, 2019) proposent d'aider les agriculteurs pour fournir une assistance directement dans les champs et les terrains (cf. Figure 2.3 C).

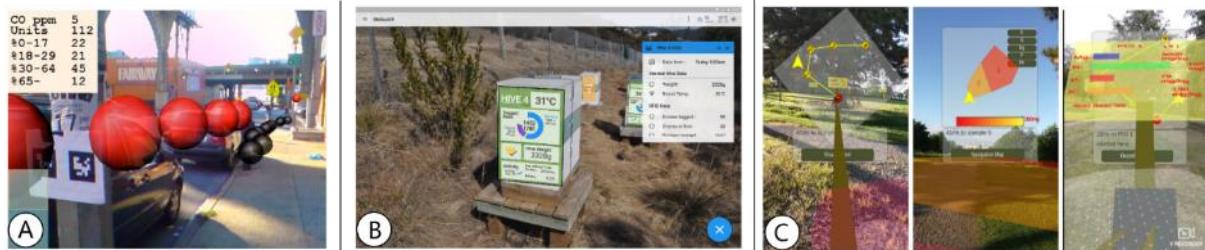


Figure 2.3 – Exemples de visualisation de données situées dans des environnements extérieurs avec (A) SiteLens (White & Feiner, 2009), (B) une visualisation sur des ruches (Engelke et al., 2016) et (C) une visualisation conçue pour des agriculteurs (Zheng & Campbell, 2019)

Le cout relativement faible des smartphones et des tablettes a permis de les faire entrer dans nos vies quotidiennes que ce soit à la maison, au musée ou à l'école. Les chercheurs se sont donc intéressés à l'utilisation de la réalité augmentée tenue en main pour la visualisation de données situées dans des contextes d'apprentissage à l'école ou dans des musées. Beheshti et collab. (Beheshti et al., 2017) étudient l'apprentissage et la construction de circuits électriques à l'aide d'une visualisation de données situées. Des groupes enfant-parent peuvent construire un circuit électrique à l'aide d'objets tangibles et ensuite visualiser les ions et les électrons se déplacer à l'intérieur du circuit à l'aide de la tablette (cf. Figure 2.4 A). Knierim et collab. (Knierim et al., 2020) utilisent un système de visualisation de données situées avec un smartphone pour aider l'apprentissage de la conduction de chaleur dans des métaux (cf. Figure 2.4 B). Enfin, le projet Meta-AR (Villanueva et al., 2020) comprend un système de visualisation de données situées pour aider la discussion entre un instructeur et un étudiant (cf. Figure 2.4 C).

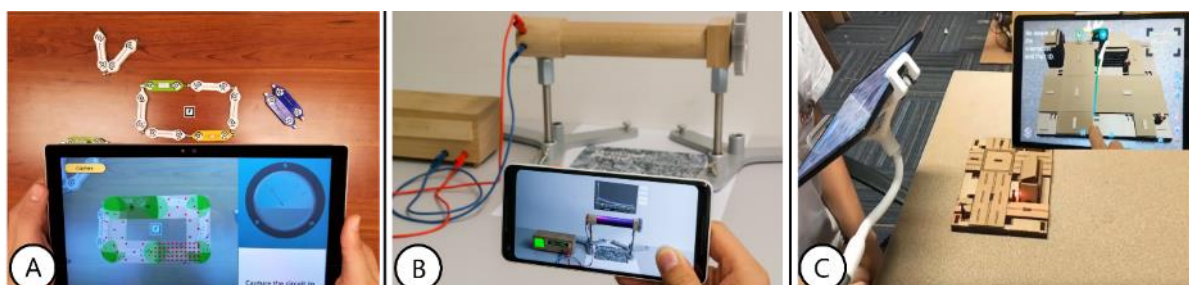


Figure 2.4 – Différents systèmes de visualisation de données situées avec un dispositif porté à la main pour l'apprentissage de (A) circuit électrique (Beheshti et al., 2017) , (B) la conduction thermique (Knierim et al., 2020) et (C) pour faciliter la communication entre un instructeur et un étudiant (Villanueva et al., 2020)

Les systèmes de visualisation de données situées avec un dispositif tenu en main offrent plusieurs avantages. En effet, les tablettes et les smartphones sont des outils très répandus et facilement transportables. Leur utilisation dans des contextes avec plusieurs utilisateurs est donc pertinente. De plus, le dispositif est équipé directement d'un écran et d'une caméra permettant un rendu vidéo en direct. Tous ces avantages permettent d'explorer de nombreux contextes d'utilisation (apprentissage, découvertes d'informations dans un magasin, etc.) et d'environnements (p. ex. à l'intérieur ou à l'extérieur). Cependant, la visualisation ne se fait qu'au travers d'une petite fenêtre (c'est-à-dire l'écran du dispositif). Les utilisateurs doivent donc explorer l'environnement pour trouver l'information qu'ils souhaitent. L'utilisateur effectue alors d'amples mouvements avec le bras tenant le dispositif mobile pour explorer l'environnement ce qui peut être source de fatigue (Hincapié-Ramos, Guo, Moghadasian, et al., 2014).

1.1.2 Les systèmes de visualisation de données situées par réalité augmentée à l'aide de dispositifs portés sur la tête

Depuis quelques années, la démocratisation de nouveaux casques (p. ex. HoloLens, Meta 2) de réalité augmentée permettant de cartographier l'environnement en temps réel a ouvert de nombreuses perspectives de recherche. Ces nouveaux dispositifs permettent de visualiser à la fois du contenu numérique au travers d'un écran semi-transparent et de l'insérer ou l'ancrer dans le monde physique. Contrairement aux tablettes ou aux smartphones, l'écran est placé sur le casque et ne requiert pas d'être tenu en main. Explorer l'environnement ne demande donc pas de larges mouvements du bras de l'utilisateur ce qui permet donc de réduire la fatigue de l'exploration de l'environnement numérique en comparaison des visualisations à l'aide de dispositif tenu en main. De plus, les casques étant transportables, les applications peuvent être développées à la fois dans des environnements extérieurs et dans des environnements intérieurs.

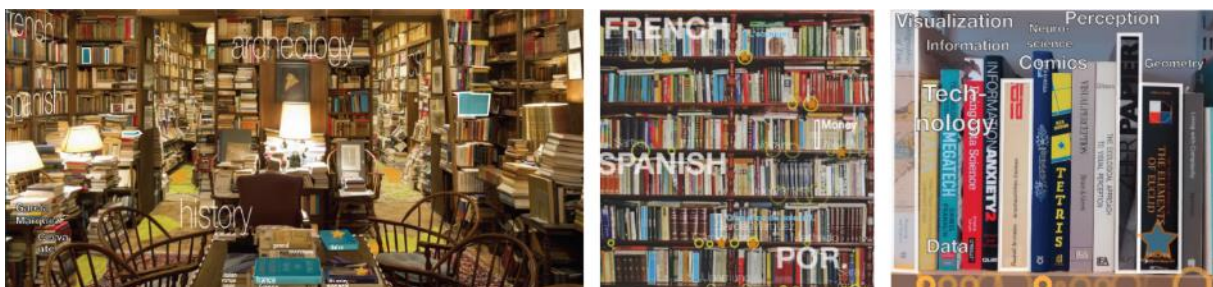


Figure 2.5 – Les différentes visualisations proposées par l'AR-CANVAS (Bach et al., 2017)

À l'intérieur d'une bibliothèque, Bach et collab. (Bach et al., 2017) conçoivent une application pour afficher des informations sur des objets visibles et non visibles dans le champ de vision (p. ex. visualiser où se trouve une section spécifique de la bibliothèque cf. Figure 2.5).

L'utilisateur peut alors voir les différentes sections de la bibliothèque ou des avis d'autres usagers sur des livres (c'est-à-dire une note moyenne). HypAR est un système de visualisation de plusieurs types de données géologiques représentant des minéraux et des matériaux de l'Ouest australien (Engelke et al., 2019). Ce dispositif est présenté dans l'Australian Resources Research Centre⁵. Les données numériques relatives aux minéraux et matériaux sont superposées à un relief 2D à l'aide d'un HoloLens. Ces visualisations de données situées à l'intérieur d'un bâtiment public permettent d'augmenter un environnement pour des visiteurs et faciliter la compréhension ou l'accès à certaines données. Cependant, les casques de réalité augmentée restent des outils coûteux et il n'est pas toujours possible d'équiper tous les visiteurs de cette technologie.

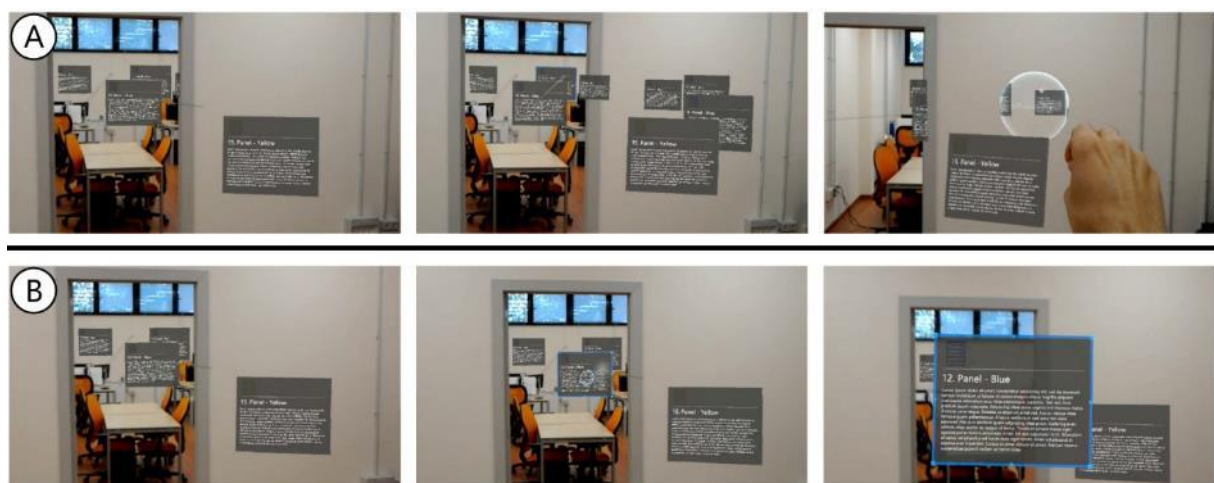


Figure 2.6 – Exploration du « *information-seeking mantra* » pour des données situées à l'aide d'un casque HoloLens (Caggianese et al., 2019) avec (A) la vue d'ensemble et (B) la manipulation directe pour avoir des détails sur l'information

Dans un contexte privé (c'est-à-dire à l'intérieur d'une maison ou de son lieu de travail), Caggianese et collab. (Caggianese et al., 2019) explorent l'*information-seeking mantra* (les principes de bases pour la conception de système de visualisation de données interactives (Shneiderman, 1996)) avec des données situées dans une pièce. Dans leurs travaux, les auteurs proposent des techniques de visualisation et d'interaction pour avoir : 1) une vue d'ensemble des informations et 2) une manipulation directe pour accéder aux détails de certaines données (cf. Figure 2.6). Büschel et collab. (Büschel et al., 2018) proposent un cadre pour explorer la recherche d'information située en fonction des différents objets présents dans l'environnement. Ce cadre est exploré à l'aide de deux applications : 1) la recherche d'images correspondant à un objet et 2) la recherche de recettes de cuisine. Dans cette deuxième application, les utilisateurs voient le nom de chaque aliment présent dans la pièce à l'aide de labels disposés

⁵ <https://www.rrc.net.au/>

au-dessus, peuvent les sélectionner et, en fonction des aliments sélectionnés, sélectionner une des recettes proposées (cf. Figure 2.7).

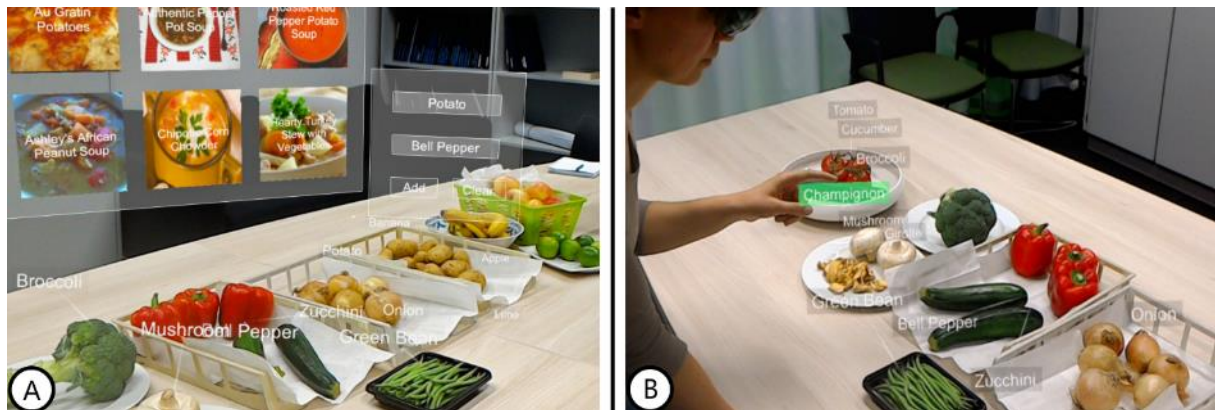


Figure 2.7 – Exemple applicatif pour la recherche d'information à partir d'objets présents dans l'environnement : la recherche d'une recette de cuisine avec (A) la vue finale et (B) la sélection d'un aliment à ajouter à la recette recherchée

Enfin, Guarese et collab. (Guarese et al., 2020), utilisent les casques de réalité augmentée pour aider les utilisateurs à se positionner dans une pièce (p. ex. lors d'une conférence). Des informations telles que la circulation des personnes, l'accès réservé aux personnes souffrant de handicap ou la température sont visualisées à l'aide d'un HoloLens (cf. Figure 2.8). Ces systèmes permettent de visualiser des informations situées dans leur environnement ce qui facilite à la fois la compréhension et le traitement de celles-ci (p. ex. évaluer la meilleure place dans une salle de conférence ou trouver une recette à partir de plusieurs aliments).

Ces dispositifs, bien que moins transportables que les smartphones ou les tablettes, sont des dispositifs pouvant s'utiliser en contexte de mobilité (p. ex. à l'extérieur). Rausch et collab. (Rausch et al., 2021) utilisent la réalité augmentée pour visualiser des informations présentes dans les villes intelligentes. Par exemple, les auteurs proposent de visualiser la silhouette d'une voiture approchant, même si la voiture est occultée par un mur ou un panneau de signalisation.

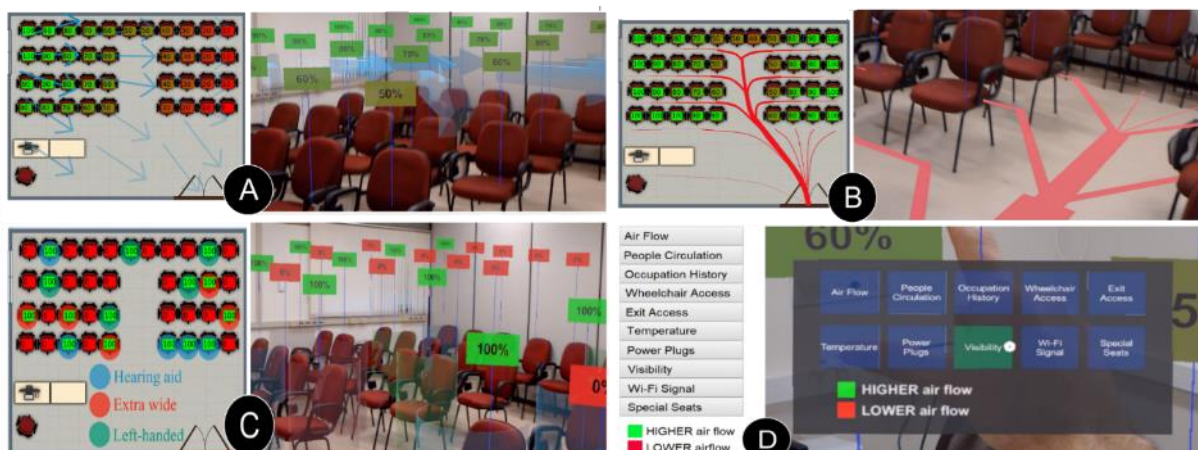


Figure 2.8 – Les différentes visualisations de données situées pour faciliter l'accès à un siège dans une pièce (Guarese et al., 2020)

En synthèse, nous retenons que dans ces systèmes de visualisation de données situées mobiles (tenus en main ou portés sur la tête), les données numériques ne peuvent se voir qu'au travers d'un écran d'une taille réduite par rapport au champ de vision des yeux humains. La visualisation des données situées est donc dépendante de la taille de la fenêtre au travers de laquelle l'utilisateur regarde. Il n'est donc pas toujours possible de visualiser toutes les données en même temps dans le même espace. De plus, pour un utilisateur, l'espace numérique est vu uniquement de son point de vue et selon l'orientation et la position dans lesquelles l'utilisateur tient ou porte le dispositif de réalité augmentée, même si plusieurs dispositifs sont connectés ensemble. Malgré de récents travaux prometteurs (Piumsomboon et al., 2018, 2017; Rühmann et al., 2018), il est encore assez difficile d'avoir la même visualisation pour différents utilisateurs. Dans un contexte où l'objet est présenté dans un lieu public (p. ex. un relief dans un musée, une maquette d'un aéroport, ou la maquette d'un bâtiment architectural placé dans son hall), il est important que toutes les personnes puissent voir et interagir avec le système de la même manière. Les systèmes de visualisation par vidéoprojection permettent d'afficher des informations numériques dans le même environnement que le référent physique (c'est-à-dire, les informations ne s'affichent pas sur un écran séparé).

1.2 Les systèmes de visualisation de données situées par vidéoprojection

Après avoir vu les systèmes de visualisation de données situées mobiles, nous synthétisons dans cette partie les approches exploitant la vidéoprojection. L'utilisation de la réalité augmentée par vidéoprojection permet d'afficher des informations sur l'objet physique. Cette technologie est déjà régulièrement utilisée pour visualiser des données situées, en particulier des données topologiques (p. ex. projeter sur un modèle réduit d'un relief ou d'une ville). Par exemple, en 2002, Piper et collab. (Piper et al., 2002) conçoivent *Illuminating Clay*, un système permettant à partir d'un relief modélisé en argile de visualiser des données relatives à ce relief. Le relief était analysé et capturé à l'aide d'un scanner laser. Un vidéoprojecteur placé au plafond projette



Figure 2.9 – Différents systèmes de visualisation de données situées sur un relief physique avec (A) *Illuminating Clay* (Piper et al., 2002), (B) *Tangible Landscape* (Millar et al., 2018) et (C) une visualisation d'écoulement de l'eau et son impact sur le relief (Harmon et al., 2016)

des données situées sur le relief (la représentation de l'écoulement de l'eau et de l'érosion) ainsi que des données situées complémentaires à ses côtés (cf. Figure 2.9 A).

La visualisation de données situées par vidéoprojection a ouvert la porte aux chercheurs pour construire, développer et expérimenter la visualisation de données situées sur des maquettes physiques dans de nombreux buts ou contextes d'utilisation. Millar et collab. (Millar et al., 2018) utilisent Tangible Landscape, un relief augmenté, pour l'apprentissage de différentes propriétés topographiques (hydrologie, géomorphologie, etc.). Les élèves peuvent modifier le modèle pour voir l'impact du changement du relief sur les données visualisées. Par exemple, les élèves sculptent le modèle, dessinent des routes ou encore positionnent des marqueurs pour indiquer des zones végétales (cf. Figure 2.9 B). Au travers d'une étude avec 16 participants durant plusieurs semaines, Millar et collab. mettent en évidence le potentiel de l'utilisation d'un système de visualisation de données situées avec une maquette pour l'apprentissage des données topographiques. Harmon et collab. (Harmon et al., 2016), proposent un système identique pour étudier les effets de l'écoulement de l'eau sur un relief 3D (cf. Figure 2.9 C). Au travers d'une étude, les auteurs comparent leur dispositif de visualisation de données situées et un logiciel entièrement numérique. Les auteurs observent de meilleures performances avec le dispositif de visualisation de données situées.

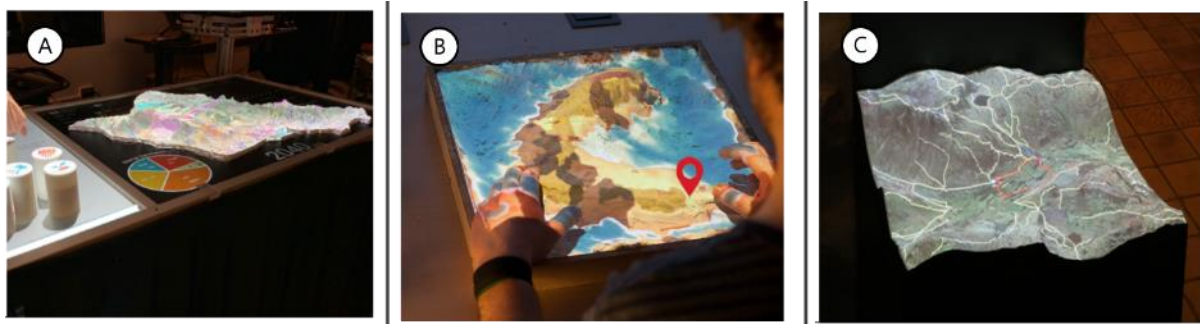


Figure 2.10 – Des systèmes de visualisations situées pour (A) la consommation énergétique (Kirshenbaum et al., 2020) , (B) le bien-être (Roo et al., 2017) et (C) l'apprentissage et la découverte d'un relief (Priestnall & Cheverst, 2019)

Ces dispositifs peuvent être aussi utilisés pour visualiser des données de consommation énergétique. Kirshenbaum et collab. (Kirshenbaum et al., 2020) utilisent un dispositif de visualisation de données situées sur un modèle 3D d'une île d'Hawaï pour le projet HAVEN ayant pour but de construire divers scénarios susceptibles de propulser Hawaï vers son objectif de 100 % d'énergie renouvelable (cf. Figure 2.10 A). L'utilisation d'un système de visualisation de données situées sur le relief physique semble offrir un meilleur engagement des utilisateurs qu'une représentation sur un plan 2D. Les visualisations de données situées sur une maquette physique permettent donc de mieux comprendre des données numériques de consommation énergétique projetées dessus ou à côté. D'autres informations peuvent être projetées sur un

relief physique. Le dispositif Inner Garden (Roo et al., 2017) présente un bac à sable augmenté pour faciliter la prise de conscience du bien être personnel (cf. Figure 2.10 B). Une visualisation s'adapte au relief créé par l'utilisateur avec ses mains. La visualisation peut évoluer en fonction de différentes données physiologiques (rythme cardiaque et pulmonaire) récupérées à l'aide de capteurs. Grâce à une évaluation menée avec des personnes pratiquant de la méditation, les chercheurs montrent qu'un système de visualisation de données situées sur un relief 3D miniature permet de véhiculer un sentiment de bien-être basé sur des données physiologiques. Ces systèmes de projection de visualisation de données situées sur des modèles 3D peuvent aussi être installés et étudiés dans des environnements publics. Priestnall et Cheverest (Priestnall & Cheverst, 2019) étudient l'impact d'un modèle réduit en relief présenté dans un pub pour attirer et orienter les visiteurs dans le paysage rural d'une vallée isolée au cœur de la région des lacs en Angleterre (cf. Figure 2.10 C). Les utilisateurs peuvent interagir en entrant en contact avec le modèle pour sélectionner des éléments. Pendant trois jours, 167 personnes ont remarqué le dispositif, 68% y ont porté leur attention et 47% sont venus interagir avec le dispositif.

L'utilisation de systèmes de visualisation de données situées n'est pas réservée à des reliefs 3D topographiques. En effet, des systèmes permettant de visualiser et d'appréhender des données relatives aux villes intelligentes ont aussi été étudiés par la communauté. Le projet CityScope (Larson et al., 2013) du MIT Media Lab⁶ vise à développer une série de plateformes tangibles et numériques destinées à résoudre les problèmes de conception spatiale et d'aménagement urbain dans différentes villes du monde. Par exemple, la ville et l'université d'Hanoï (cf. Figure 2.11 A) se sont dotées d'un CityScope pour aider la ville à la gestion de l'eau et du système d'irrigation du delta du Fleuve Rouge (Grignard et al., 2020). Les utilisateurs peuvent manipuler des tuiles représentant certaines informations afin de mettre à jour la visualisation. La ville d'Andorre (cf. Figure 2.11 B) a utilisé un système CityScope pour comprendre et analyser le comportement des visiteurs de la ville (Grignard et al., 2018). Enfin, le site de Volpe aux États-Unis d'Amérique (cf. Figure 2.11 C) visualisent des données comme la densité, la diversité, la proximité ou encore l'énergie consommée par les bâtiments (Alonso et al., 2018).

Les systèmes de vidéoprojection s'adaptent à de nombreux contextes. En effet, plusieurs utilisateurs ont le même point de vue sur les données et sur le référent physique et peuvent alors se coordonner dans la prise de décision. Cependant, ces systèmes ne sont que peu mobiles,

⁶ <https://www.media.mit.edu/>



Figure 2.11 – Différents exemples de projet CityScope avec (gauche A) la ville d'Hanoï, (centre B) la ville d'Andorre et (droite C) le site de Volpe

l'utilisation d'une projection lumineuse limite leur utilisation dans des contextes trop lumineux et une calibration de chaque vidéoprojecteur doit être faite selon la position, l'orientation ou les propriétés intrinsèques du vidéoprojecteur.

Dans cette section, nous avons présenté différents systèmes de visualisation de données situées au travers de différentes approches. Dans un premier temps, nous avons présenté les systèmes mobiles (tenus en main ou portés à la tête) ainsi que leurs avantages et leurs limitations. L'utilisation d'un dispositif externe équipé de caméra et d'écran permet de concevoir des applications aussi bien en intérieur qu'en extérieur. Cependant, la visualisation des données situées ne se fait qu'au travers de l'écran. La taille de l'espace numérique visible est alors fortement liée à la taille de l'écran. La taille de ces écrans est encore trop réduite et l'utilisateur doit donc effectuer des mouvements d'exploration (avec la tête ou le bras) pour visualiser l'ensemble de l'espace numérique. Ces mouvements peuvent alors induire de la fatigue. Dans un deuxième temps, nous avons présenté les systèmes par vidéoprojection. Ces dispositifs permettent à tous les utilisateurs d'avoir le même point de vue de l'ensemble des informations. Cependant, il est complexe de concevoir des systèmes par vidéoprojection dans des environnements moins contrôlés ou trop lumineux (p. ex. en extérieur).

Dans la suite de l'état de l'art, nous présentons des techniques d'interaction potentielles pour interagir avec un système de visualisation de données situées (c'est-à-dire avec le référent physique ou avec les données numériques). Nous focalisons notre revue de la littérature sur les techniques d'interaction permettant de sélectionner un élément physique ou numérique.

2 Les techniques de sélection pour une visualisation de données situées en réalité augmentée

La sélection dans un système de réalité augmentée pour la visualisation de données situées peut avoir lieu à la fois avec les données (p. ex. obtenir des détails d'une valeur) et à la fois avec le

réfèrent physique (p. ex. déclencher des visualisations, changer le type de données). La sélection demande à l'utilisateur d'indiquer une zone, une donnée, un objet physique pour ensuite pouvoir les modifier, les manipuler ou leur appliquer des commandes. La sélection est alors une interaction qui précède toutes les autres. Cette sélection peut prendre deux formes : 1) **au contact** en touchant directement ce que l'utilisateur souhaite sélectionner ou 2) **à distance** où un geste de l'utilisateur est interprété par un système qui transforme ce mouvement physique en une action de désignation dans l'espace numérique (p. ex. l'utilisateur bouge une souris avec son bras et le système transforme ce mouvement en coordonnées pour le pointeur).

Dans cette section, nous détaillons les spécificités de la sélection au contact et à distance à l'aide d'une revue de la littérature de techniques de sélection en environnement immersif (réalité augmentée ou réalité virtuelle).

2.1 Sélection au contact

La sélection au contact demande à l'utilisateur d'entrer physiquement en contact soit avec le réfèrent physique soit avec les données numériques (projetées, visualisées à l'aide d'un casque ou visualisées à l'aide d'un dispositif tenu en main).

La visualisation de données situées par vidéoprojection permet d'afficher des données numériques dans l'environnement du réfèrent physique (c'est-à-dire il n'y a pas d'écran entre le réfèrent physique et les données numériques). L'interaction pour sélectionner des données numériques ou le réfèrent physique est alors similaire. La sélection au contact est très répandue pour les systèmes de vidéoprojection (cf. section 1.2). Par exemple, Piper et collab. (Piper et al., 2002) proposent plusieurs modes de visualisation pour afficher des données topologiques. L'utilisateur peut alors modifier le réfèrent physique et les données vont pouvoir s'actualiser en fonction des modifications. Ce type d'interaction est aussi présente dans d'autres projets d'interaction avec un relief topologique (Harmon et al., 2016; Millar et al., 2018; Roo et al., 2017). Cependant, la manipulation directe du réfèrent physique entraîne une **occultation** et une déformation des données numériques projetées. En effet la visualisation étant généralement projetée par le haut, quand l'utilisateur entre en contact avec sa main sur le réfèrent physique la projection apparaît sur sa main et non sur le réfèrent physique (cf. Figure 2.9).

De plus, la sélection au contact requiert de pouvoir manipuler l'entièreté du réfèrent physique ou des données numériques. L'**atteignabilité** et la **disponibilité d'accès** sont donc nécessaires pour interagir au contact du réfèrent physique. Par exemple, les projets de visualisations de données situées projetées sur des maquettes de villes ou de quartiers (cf. Figure 2.11) proposent

des maquettes de taille conséquente. L'utilisateur n'a que difficilement accès au centre du référent physique. Le référent physique ou les données numériques doivent donc être placés dans un environnement réduit où l'utilisateur peut facilement les sélectionner.

Ces interactions souffrent donc de trois problèmes majeurs :

- **L'occultation** : interagir au contact du référent physique entraîne une superposition de la main ou de l'outil utilisé pour interagir et des données projetées sur la surface du référent.
- **L'atteignabilité** : certains référents physiques sont parfois inatteignables car trop grands ou trop éloignés de l'utilisateur. L'utilisateur n'a donc aucune possibilité d'aller interagir avec une zone distante de lui.
- **La disponibilité d'accès** : dans certains contextes, le référent physique nécessite d'être protégé car trop fragile (p. ex. artefact d'un musée). L'utilisateur ne peut donc pas directement toucher ou placer un élément sur le référent physique.

Pour pallier ces problèmes, une solution utilisée dans de nombreux environnements immersifs (réalité virtuelle, réalité augmentée avec un casque, etc.) est d'interagir à distance avec le référent physique ou les données numériques. L'utilisateur n'a plus à entrer physiquement au contact du référent physique ou des données et donc peut se trouver à distance évitant aussi l'occultation.

2.2 Sélection à distance

Sélectionner à distance requiert de transférer un mouvement de l'utilisateur en une action comprise par le système. Cette action peut avoir lieu **en l'air** ou au **contact d'une surface**.

Par ailleurs, lors des mouvements exécutés en l'air ou appliqués sur la surface, les translations et rotations constituant ces mouvements peuvent être basées sur **la présence d'une relation spatiale** avec le référent ou la représentation des données situées. **En l'air** ces mouvements, généralement 3D, impliquent une ou des parties du corps de l'utilisateur. **Sur une surface** il peut s'agir des mouvements de la surface ou des mouvements d'une partie du corps de l'utilisateur sur la surface. Dans les deux cas et en présence d'une référence spatiale, les mouvements sont considérés comme étant réalisés dans le référentiel du référent ou des données situées.

Dans le cas contraire, en **l'absence de relation spatiale**, ces mouvements sont interprétés indépendamment du référent ou des données situées : un référentiel arbitraire est utilisé tel qu'un référentiel égocentrique, le cadre d'une table ou d'un objet porté par l'utilisateur, etc.

Ces deux propriétés peuvent être résumées par le tableau ci-dessous et des exemples de la littérature illustrent les quatre catégories résultantes dans les 4 sections suivantes.

Tableau 1 – Propriétés des sélections à distance

	En l'air	Support de contact
Absence de relation spatiale	Section 2.2.1	Section 2.2.3
Présence de relation spatiale	Section 2.2.2	Section 2.2.4

2.2.1 Sélection en l'air sans relation spatiale

La sélection en l'air sans relation spatiale demande à l'utilisateur de réaliser des mouvements qui sont interprétés par le système indépendamment de la position et de l'orientation du dispositif d'interaction par rapport au référent physique ou aux données situées.

Ces mouvements se déroulent alors dans un espace limité dédié uniquement à l'interaction en entrée. À notre connaissance peu de travaux se sont intéressés à la sélection en l'air sans relation spatiale. Premièrement, l'utilisation de gestes en l'air permet d'interagir avec un langage plus naturel pour l'utilisateur (on parle d'interaction utilisateur naturelle). Ensuite, le pointage avec son corps (généralement son doigt) est une tâche fondamentale qui apparaît dès la petite enfance (Butterworth, 2003) pour désigner directement un objet ou une zone. Cette désignation se fait en direction de l'objet que l'enfant souhaite indiquer et donc dans le même référentiel que celui-ci. Utiliser les mouvements d'une partie de son corps dans un autre référentiel que celui du référent physique ou des données semblerait donc contre-intuitif.

Néanmoins, certains travaux ont étudié l'utilisation de mouvements en l'air sans relation spatiale. Brasier et collab. (Brasier et al., 2020) proposent l'utilisation d'un pavé tactile virtuel avec un *mapping* absolu dans l'HoloLens selon deux propriétés : le placement du pavé virtuel (hanche, cuisse, poignet) et l'orientation (vertical ou horizontal). L'utilisateur contrôle un curseur en pointant une zone du pavé virtuel dans un environnement immersif. Ces différents pavés virtuels ont été comparés à des techniques de sélection avec relation spatiale proposées par l'HoloLens (cf. Figure 2.13). Il semblerait que les utilisateurs soient moins rapides avec les techniques de sélection sans relation spatiale. Cependant, l'utilisation de techniques sans

relation spatiale proches du corps permet de réduire la taille des mouvements à effectuer et donc limiter la fatigue de l'utilisateur.

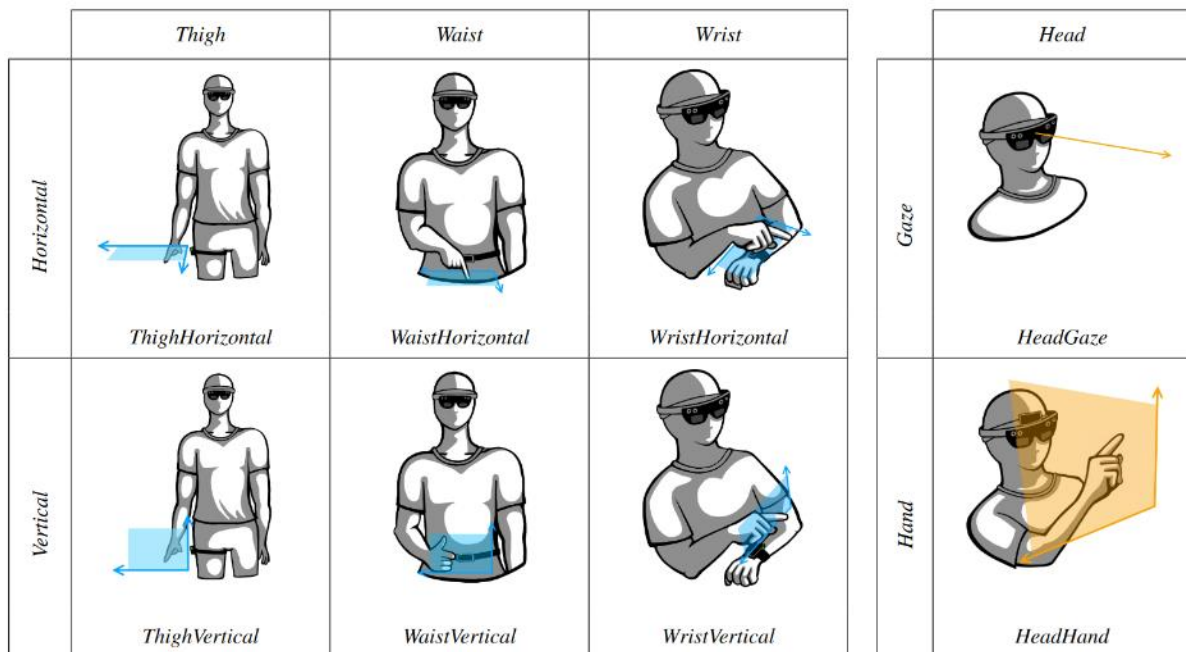


Figure 2.13 – L'espace de conception de Arpads (Brasier et al., 2020)

Dans ce travail, l'orientation et la position du pavé tactile numérique sont contrôlées. Andujar et Argelaguet (Andujar & Argelaguet, 2007) proposent de laisser l'utilisateur choisir la position et l'orientation du pavé virtuel. L'utilisateur sélectionne, à l'aide d'un lancer de rayon, un point du pavé virtuel pour contrôler le curseur selon un contrôle relatif. La technique de sélection en l'air sans relation spatiale est comparée à une technique en l'air avec relation spatiale. Les résultats montrent que réaliser certaines tâches avec la technique sans relation spatiale prend plus de temps que ces mêmes tâches réalisées avec une technique en l'air avec relation spatiale malgré le confort apporté de la technique sans relation spatiale (cf. Figure 2.12).

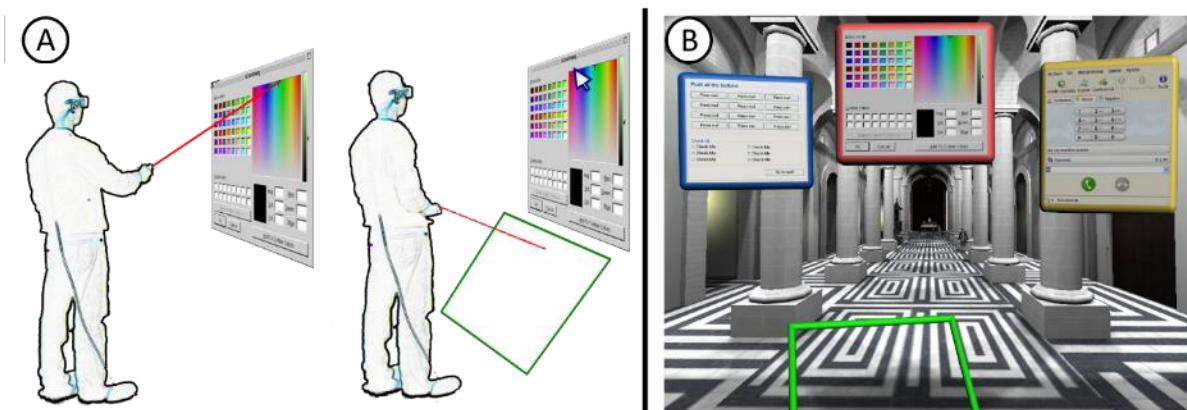


Figure 2.12 – VirtualPads (Andujar & Argelaguet, 2007) avec (A) le concept de technique de sélection directe et indirecte et (B) le prototype avec les fenêtres et le pavé virtuel en vert

Les techniques de sélection en l'air sans relation spatiale semblent donc être moins performantes que les techniques de sélection avec relation spatiale considérées dans ces travaux. Cependant, les techniques en l'air sans relation spatiale permettent de réduire les mouvements de l'utilisateur et par conséquent, réduire la fatigue de l'utilisateur.

2.2.2 Sélection en l'air avec relation spatiale

La sélection en l'air avec relation spatiale transforme le mouvement en fonction de la position et l'orientation du dispositif d'interaction par rapport au référent physique ou aux données. Le dispositif d'interaction est ensuite projeté sur le référent physique ou les données situées. Cette projection se fait alors généralement par des techniques de lancer de rayon (*raycasting*).

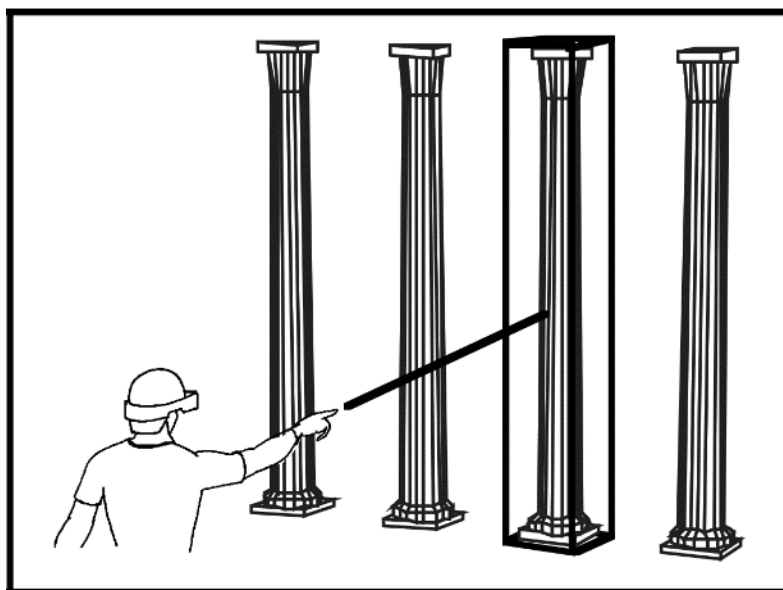


Figure 2.14 – La technique de sélection par lancer de rayon proposée par Mine (Mine, 1995)

Mine (Mine, 1995) définit le lancer de rayon comme une projection (à l'instar d'un faisceau laser) sortant de la main ou du doigt de l'utilisateur. L'intersection entre les objets numériques ou physiques et le rayon définit l'objet sélectionné (cf. Figure 2.14). Le lancer de rayon peut alors se définir selon plusieurs propriétés (Argelaguet & Andujar, 2013) : l'origine du rayon, la direction du rayon, la forme du rayon ou le type de dispositif contrôlant le rayon, etc.

Premièrement, lancer un rayon nécessite un point de départ (c'est-à-dire une origine). Hincapié-Ramos et collab. (Hincapié-Ramos et al., 2015) étudient différentes origines du lancer de rayon pour la sélection à l'aide d'un dispositif dédié dans un environnement de réalité augmentée porté sur la tête. Les résultats de l'étude montrent que placer l'origine du lancer de rayon sur le dispositif n'est pas la solution optimale (en matière de temps de complétion et en nombre d'erreurs). Néanmoins, la majorité des travaux de recherche sur la sélection à distance basée

sur un lancer de rayon placent l'origine du rayon au niveau du dispositif (main, doigt ou contrôleur).

Définir l'origine d'un lancer de rayon n'est pas une condition suffisante pour le représenter. En effet, le lancer de rayon a une origine et une direction. Cette direction peut être définie soit par l'orientation du dispositif permettant le lancer de rayon, soit par la présence d'un autre point dans l'espace. Dans ce dernier cas, le rayon correspond à la demi-droite partant de l'origine du rayon et croisant le deuxième point. Teather et Stuerzlinger (Teather & Stuerzlinger, 2013) proposent différentes conceptions de lancer de rayon pour la sélection 3D dans un environnement virtuel sur ordinateur. Certains lancers de rayon sont alors définis selon un point départ (les yeux ou un dispositif) et soit une orientation (l'orientation de la tête ou du dispositif) (cf. Figure 2.15 A C D) soit une demi-droite croisant un point contrôlé par l'utilisateur (cf. Figure 2.15 B). Ces différentes techniques (ou des dérivés) ont fait l'objet d'autres travaux présents dans la littérature (Teather et al., 2014; Teather & Stuerzlinger, 2011).

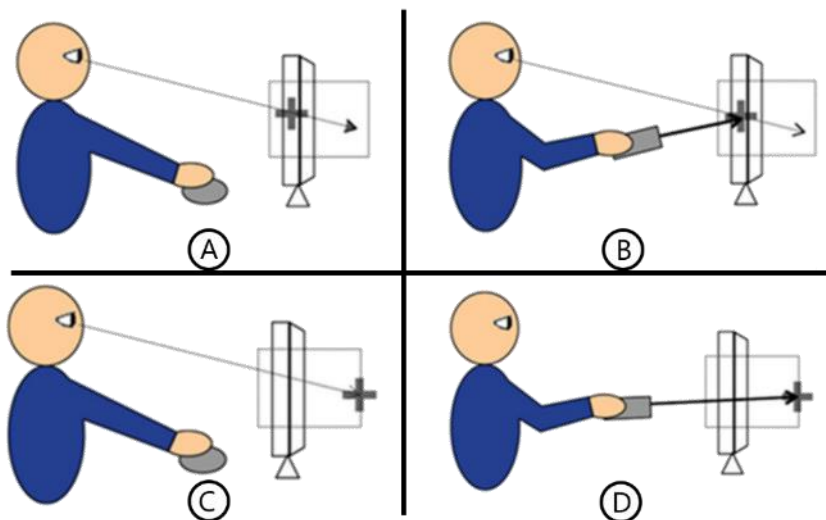


Figure 2.15 – Les quatre techniques de sélection de Teather et Stuerzlinger (Teather & Stuerzlinger, 2013)

Un lancer de rayon peut aussi être caractérisé par sa taille ou l'angle d'ouverture du rayon. En effet, de nombreuses techniques proposent de lancer non pas un rayon, mais un cône (Liang & Green, 1994; Olwal et al., 2003; G. Schmidt et al., 2006; Steed, 2006; Steinicke et al., 2006). Utiliser un cône plutôt qu'un rayon rectiligne permet d'élargir la zone de contact entre le rayon et les objets à sélectionner. Hincapié-Ramos et collab. (Hincapié-Ramos et al., 2015, 2016) proposent d'utiliser l'inclinaison de leur dispositif pour contrôler l'ouverture du cône entre 0° (un lancer de rayon en ligne droite) et 6° . Forsberg et collab. (Forsberg et al., 1996) utilisent la

distance entre l'origine du cône et un point pour définir l'ouverture de celui-ci ; plus le point est proche (resp. loin) de l'origine plus l'ouverture sera réduite (resp. élargie).

Enfin, le lancer de rayon peut être contrôlé à partir de différents dispositifs. Pham et Stuerzlinger (Pham & Stuerzlinger, 2019) comparent 2 types de dispositifs permettant de contrôler un rayon : un stylet conçu lors d'une pré-expérimentation et le contrôleur du HTC Vive (cf. Figure 2.16). Ces deux dispositifs ont été étudiés pour une tâche de pointage dans un contexte de réalité virtuelle et un contexte de réalité augmentée. Les résultats de cette étude montrent que le stylet est plus performant (temps et nombre d'erreurs) que le contrôleur du HTC Vive.



Figure 2.16 – Les deux dispositifs étudiés par Pham et Stuerzlinger (Pham & Stuerzlinger, 2019)

Concevoir une technique de sélection en l'air avec relation spatiale nécessite donc de prendre en considération les différentes caractéristiques du lancer de rayon. Cependant, les lancers de rayon à l'aide de dispositifs tenus en main sont susceptibles de causer de la fatigue (cet effet est connu sous le nom de « *gorilla-arm effect* » (Hincapié-Ramos, Guo, Moghadasian, et al., 2014; Kharlamov et al., 2016)).

2.2.3 Sélection sans relation spatiale avec support de contact

La sélection sans relation spatiale et avec un support de contact demande à l'utilisateur d'effectuer sur une surface, un mouvement avec un dispositif sans que la position et l'orientation du geste ou de la surface soient interprétées en regard du référent ou des données situées. Par exemple, sélectionner un élément avec une souris sur un ordinateur est une interaction sans relation spatiale et avec un support de contact (p. ex. la table), car la souris doit être posée sur une surface et la position et l'orientation de la souris par rapport à l'écran n'influencent pas l'interaction. Pour la sélection dans un environnement immersif, les techniques sans relation spatiale et avec support de contact se rapprochent des solutions en l'air sans relation spatiale hormis le fait que l'utilisateur ne doit pas effectuer des mouvements en l'air, ce qui peut réduire la fatigue inhérente à ces sélections. Ces supports de contact peuvent varier selon la taille, la forme ou la position du support de contact.

À l'instar des pavés tactiles sur ordinateur portable, de nombreux travaux de recherche explorent la sélection sans relation spatiale en utilisant une surface 2D tactile (p. ex. un smartphone). Chen et collab. (Chen et al., 2020) proposent d'utiliser soit un smartphone soit une smartwatch comme support de contact dans une application de réalité mixte avec un

HoloLens. L'utilisateur peut alors déplacer un curseur en effectuant des mouvements sur le support de contact dans le champ de vision du casque pour sélectionner un élément. Le curseur peut alors se déplacer de manière relative ou absolue. Les auteurs comparent, au travers d'études utilisateur, différentes techniques de sélection sans relation spatiale à une technique de sélection avec relation spatiale (un lancer de rayon basé sur la position et l'orientation de la tête). Ces études montrent que les sélections sans relation spatiale avec support de contact permettent aux utilisateurs de faire moins d'erreurs. Cependant, les utilisateurs semblent moins rapides avec les techniques sans relation spatiale et avec support de contact. Dans cette étude, le smartphone était tenu en main et la smartwatch portée sur le poignet.

Dans FaceTouch, Gugenheimer et collab. (Gugenheimer et al., 2016) étudient différentes positions d'un support de contact (c'est-à-dire un pavé tactile) pour la sélection de cibles en réalité virtuelle : sur la face du casque, sur le côté du casque ou tenu en main (cf. Figure 2.17). L'emplacement sur la face du casque de réalité virtuelle semble la plus précise et permet de sélectionner des éléments plus rapidement que les autres emplacements (sur le côté ou tenu en main). De plus, pouvoir positionner le support de contact sur le casque permet à l'utilisateur d'avoir une main libre pour réaliser d'autres actions.



Figure 2.17 – Les trois emplacements du pavé tactile évalués dans (Gugenheimer et al., 2016)

D'autres travaux étudient des supports de contact originaux pour l'interaction. Par exemple, Saidi et collab. (Saidi et al., 2019) proposent aux utilisateurs d'utiliser leur corps comme support d'interaction avec une souris hémisphérique. Les utilisateurs peuvent alors tourner et déplacer la souris sur plusieurs parties de leur avant-bras (cf. Figure 2.18).

Sélectionner avec un dispositif sans relation spatiale et avec un support de contact permet à l'utilisateur d'avoir un retour kinesthésique contrairement aux interactions en l'air. Cependant, la présence d'un support limite l'interaction à la taille de celui-ci. Les applications immersives proposant un environnement potentiellement infini, le rapport entre la taille du support de

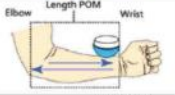
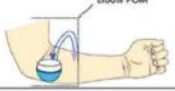
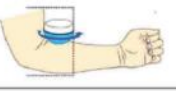


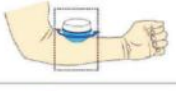




PLACE OF MOTION	TANGIBLE ACTION (TA)		
	TRANSLATION	ROTATION	ROLL
Length POM		N/A	N/A
Elbow POM			
Middle POM			
Wrist POM			

Figure 2.18 – L'espace de conception des interactions de (Saidi et al., 2019)

contact et de l'environnement peut requérir soit de nombreuses actions de *clutching* dans le cas d'une sélection relative soit une grande précision dans le cas d'une sélection absolue.

2.2.4 Sélection avec relation spatiale et avec support de contact

La sélection avec relation spatiale et avec un support de contact requiert d'agir sur le support ou de déplacer et d'orienter le support de contact dans le référentiel du référent physique ou des données numériques.

Ces sélections sont courantes dans les applications de réalité augmentée avec un smartphone ou une tablette (c'est-à-dire où le référent physique est rendu en temps réel grâce à l'écran et les données sont affichées par-dessus). ElSayed et collab. (ElSayed et al., 2015), proposent un système de visualisation de données situées à l'aide d'une tablette dans un magasin. Les utilisateurs peuvent interagir au contact de la tablette pour filtrer, trier, noter les différents produits affichés grâce au rendu vidéo fourni par la caméra de la tablette (cf. Figure 2.19).

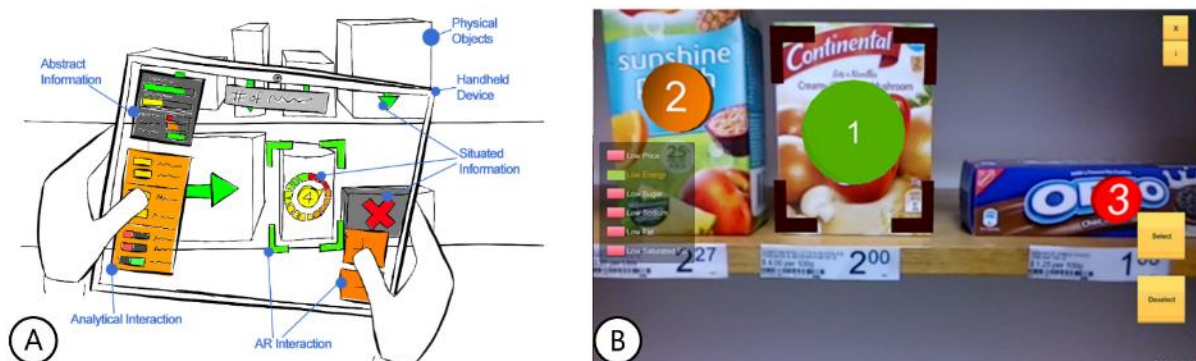


Figure 2.19 – Interaction directe avec un support de contact dans un magasin avec (A) le prototype du système et (B) le système de classement des produits en fonction d'une donnée

Hincapié-Ramos et collab. (Hincapié-Ramos, Roscher, et al., 2014) proposent d'utiliser un écran transparent à poser sur le référent physique pour interagir directement avec lui.

L'utilisateur manipule l'écran sur, par exemple, une page de texte pour interagir avec le document et son contenu. D'autres approches similaires utilisent des vitrines transparentes pour des documents textuels (Maeda et al., 2011; Sharma et al., 2013). L'utilisation d'un support de contact pour sélectionner des données situées ou des parties de référents physiques requiert généralement un dispositif interactif tenu en main. Le rendu vidéo permettant l'interaction est alors fortement lié à la position et l'orientation du dispositif en fonction du référent physique. Cependant, le rendu vidéo n'existe qu'au travers de la taille de l'écran dispositif interactif. De plus, ce dispositif tenu en main (avec une ou les deux mains) ne facilite pas les interactions gestuelles.

2.3 Synthèse

La sélection d'un élément dans un dispositif de visualisation de données situées peut prendre place soit dans l'espace numérique (les données) soit dans l'espace physique (le référent). Quand ces deux espaces sont parfaitement superposés (réalité augmentée par vidéoprojection) l'interaction se fait généralement au contact. La sélection au contact permet donc d'interagir à la fois avec les données et le référent physique. Cependant, cette sélection soulève trois problèmes majeurs : l'occultation, l'atteignabilité et la disponibilité d'accès (cf. section 2.1). Pour pallier ces problèmes, il est nécessaire d'interagir à distance du référent physique ou des données numériques. Cependant, à notre connaissance, peu de travaux s'intéressent à la sélection à distance d'objets physiques.

Deux propriétés majeures permettent de définir l'interaction à distance : la présence d'une relation spatiale entre le mouvement et le référent physique ou les données numériques et la présence d'un support de contact pour interagir.

La sélection sans relation spatiale à distance permet de définir un espace d'interaction décorrélé des espaces numérique et physique. Cette absence de relation permet de définir un espace à part et personnalisé selon la position, l'orientation ou la taille qui correspond le mieux à l'application développée. Cependant, ces sélections nécessitent de transformer un déplacement dans l'espace d'interaction réduit vers le référent physique ou les données numériques. Dans les études comparatives, il semble que les techniques de sélection sans relation spatiale sont moins rapides en comparaison des sélections **avec relation spatiale**.

La sélection avec relation spatiale à distance interprète les mouvements réalisés par l'utilisateur en fonction du référentiel du référent physique ou des données numériques. Le référent physique pouvant avoir une taille conséquente (p. ex. un bâtiment) et l'environnement

numérique affichant les données pouvant être potentiellement infini, d'amples mouvements soit d'une partie du corps soit de l'outil d'interaction sont requis. Ces mouvements peuvent entraîner de la fatigue chez l'utilisateur.

La sélection avec un support de contact à distance permet à l'utilisateur d'avoir un retour kinesthésique de ses actions. Cependant, l'interaction est généralement limitée à la taille ou la forme de ce support de contact. Dans un environnement numérique potentiellement infini ou en présence d'un référent physique trop imposant, la taille du support de contact peut entraîner des actions de « *clutching* » et donc dégrader la performance de l'utilisateur.

La sélection en l'air (sans support de contact) à distance demande à l'utilisateur d'effectuer des mouvements avec une partie de son corps. La diversité des mouvements possibles permet de développer de nombreuses techniques de sélection différentes. Néanmoins, l'utilisation d'amples mouvements peut causer de la fatigue chez l'utilisateur et donc réduire l'utilisabilité de l'application sur le long terme.

La conception des systèmes permettant d'interagir avec un référent physique et des données numériques doit prendre en compte de nombreux paramètres. En effet, outre les multiples solutions mises en évidence dans ce chapitre, un tel contexte d'interaction peut être scindé en deux grandes tâches : sélection sur le référent physique et manipulation des données numériques afférentes à la sélection physique. Or, sélectionner tout ou partie d'un référent physique est une tâche encore peu étudiée, alors qu'elle est primordiale pour, par exemple, afficher les données numériques relatives au référent physique sélectionné. De plus, la manipulation des données numériques affichées, après sélection sur le référent physique, offre encore une multitude d'options telles que la position par rapport au référent, les liens mis en évidence, le support d'affichage, etc. La conception et l'évaluation de tels systèmes est donc une tâche complexe, car ouverte à de nombreuses options d'interactions différentes. En réponse à ce contexte, des approches prédictives de l'efficacité d'une interface ont été développées et font l'objet de mises à jour récentes. Ces approches ont l'avantage de pouvoir faciliter la conception, la comparaison et l'évaluation de systèmes interactifs. Nous synthétisons ces approches dans la section suivante.

3 Les modèles prédictifs

L'évaluation des interfaces repose traditionnellement sur des études utilisateur, qui sont parfois mises à mal par la complexité des tâches, la multiplicité des modalités d'interaction ou les

contextes d'utilisation particuliers. En revanche, l'évaluation basée sur des modèles (Kieras, 2009) permet aux concepteurs d'évaluer un système ou une technique d'interaction avant sa mise en œuvre. Dans cette section, nous présentons brièvement et en premier lieu différents types de modélisation pour l'évaluation de systèmes interactifs. Nous détaillons principalement le Keystroke-Level Model (KLM) car il est simple d'utilisation et a fait ses preuves dans de nombreux domaines. Nous détaillons ensuite le modèle KLM original, son histoire et son fonctionnement. Enfin, nous détaillons les différentes étapes nécessaires à la création d'une extension du modèle KLM.

3.1 Les différents types de modèles pour évaluer des systèmes interactifs

Kieras (Kieras, 2009) définit trois évaluations basées sur des modèles différents : les modèles de tâches, les modèles d'architecture cognitive, et la famille des modèles dérivés de GOMS (pour Goals, Operators, Methods, Selection). Les modèles de tâches représentent la tâche sous forme de réseau de tâches. Chaque tâche peut démarrer quand toutes les tâches nécessaires à son fonctionnement sont terminées. Il est important de noter qu'une tâche peut être une tâche réalisée par l'utilisateur grâce à une action physique de sa part, par l'utilisateur grâce à une action cognitive ou encore une tâche réalisée par la machine (Chubb, 1981; Laughery, 1989). Les modèles d'architecture cognitive modélisent des composants perceptuels, cognitifs ou moteurs où chacun de ces composants est composé de propriétés venant des recherches scientifiques en psychologie (Byrne, 2002). Les composants et les interactions qu'ils peuvent avoir entre eux sont simulés à l'aide d'un logiciel. Enfin, les modèles dérivés de GOMS prennent racine dans les théories de résolution de problème et d'acquisition de compétences des êtres humains. Les modèles dérivés de GOMS décrivent les connaissances nécessaires et les quatre composantes cognitives de la performance qualifiée dans les tâches : les objectifs, les opérateurs, les méthodes et les règles de sélection (Olson & Olson, 1990). Les modèles de la famille GOMS semblent avoir fait leurs preuves même s'ils ne sont pas utilisés aussi largement que leur simplicité et leur performance le justifieraient (Kieras, 2009). Dans cette section, nous passons en revue les travaux antérieurs sur un modèle de prédiction issu de la famille des modèles GOMS : le Keystroke-Level Model ou KLM.

3.2 Le modèle KLM de la famille des modèles GOMS

L'acronyme GOMS (Kieras, 2009) peut s'expliquer comme ceci : l'utilisateur veut atteindre un objectif (Goal) avec le système ; les opérateurs (Operators) sont les actions de base qui peuvent être exécutées sur le système par l'utilisateur ; les méthodes (Methods) sont des suites

d'opérateurs qui, appliqués ensemble, peuvent réaliser un objectif ; et les sélections (Selection) définissent quelle méthode doit être utilisée pour atteindre un objectif (s'il y a plusieurs méthodes pour atteindre un même objectif). Plusieurs modèles s'inspirant de GOMS ont vu le jour.

Le Keystroke-Level Model (KLM) (Card et al., 1980) fait partie de la famille GOMS. KLM est conçu pour prédire le temps de réalisation et la séquence des opérateurs pour réaliser une tâche. Par rapport au modèle initial GOMS, KLM n'utilise que les concepts d'opérateurs (O) et de méthodes (M), ignorant ainsi les objectifs et la sélection. Avec KLM, la modélisation d'une tâche nécessite la décomposition de la tâche en un ensemble d'opérateurs tels que l'appui sur une touche, le pointage, les actions cognitives, le guidage, etc.

L'hypothèse émise par KLM est que l'utilisateur effectue une suite d'actions pour réaliser une tâche plus importante. Une tâche se décompose alors en deux phases :

1. L'acquisition de la tâche correspond à la phase où l'utilisateur conceptualise et développe une représentation mentale de la tâche à réaliser.
2. L'exécution de la tâche correspond aux processus nécessaires pour réaliser la séquence d'action pour réaliser la tâche.

Le temps de réalisation complet de la tâche peut donc être modélisé comme la somme du temps d'acquisition de la tâche et le temps d'exécution de la tâche :

$$T_{t\grave{a}che} = T_{acquisition} + T_{ex\acute{e}cution} \quad (1)$$

Le modèle KLM ne modélise que le temps d'exécution de la tâche car c'est le seul temps sur lequel le concepteur de l'interface peut avoir du contrôle. Le modèle KLM était à l'origine conçu pour les systèmes interpréteurs de commande. Card et collab. (Card et al., 1980) proposent six opérateurs pour interagir avec ces systèmes dont quatre correspondent à une action physique : Appui sur une touche (Keystroke K), Pointage (Pointing P), Homing (H), Dessin (Drawing D). Les deux opérateurs suivants représentent le temps de préparation mentale pour réaliser une action (Mental M) et le temps de réponse d'un système (Response Time R(t)). Chaque opérateur est associé à une unité de temps (c'est-à-dire le temps que l'utilisateur met pour exécuter l'action modélisée par l'opérateur). Les opérateurs de KLM et les unités de temps associées sont présentés en Tableau 2.

Modéliser une tâche demande donc au concepteur de l'interface de comprendre et d'anticiper chacune des actions physiques que l'utilisateur va devoir effectuer pour réaliser sa tâche.

Tableau 2 – Les opérateurs originaux de KLM

Type	Opérateur	Description	Unité de temps (s)
Physique	Keystroke (K)	Appui sur une touche :	
		Meilleur dactylographe (135 mots par minute)	0,08
		Bon dactylographe (90 mots par minute)	0,12
		Dactylographe moyen (55 mots par minute)	0,20
		Dactylographe débutant (40 mots par minute)	0,28
		Écrire des lettres au hasard	0,5
	Écrire des symboles complexes	0,75	
	Pire dactylographe	1,20	
	Pointing (P)	Pointage avec une souris	1,10
	Homing (H)	Déplacer les mains sur le clavier ou tout autre dispositif	0,40
	Drawing (D)	Dessiner n lignes droites avec une longueur totale l	$0,90_{nD} + 0,16_{lD}$
Mental	Mental Act (M)	Préparation mentale à une action	1,35
System	R(t)	Temps de réponse d'un système	t

Contrairement aux actions physiques, les opérateurs de type Mental ne sont pas directement observés par le concepteur. L'opérateur mental M doit donc être placé à l'aide d'heuristiques basées sur les connaissances que l'on a des utilisateurs et de leurs compétences. Le modèle KLM propose 5 heuristiques pour placer les opérateurs mentaux entre les opérateurs physiques :

- **Règle 0** : insérer un opérateur Mental (**M**) devant chaque opérateur Keystroke (**K**) qui ne fait pas partie d'une « chaîne d'arguments ». Insérer un opérateur Mental (**M**) devant tous les opérateurs Pointing (**P**) qui permettent de sélectionner des commandes.
- **Règle 1** : si un opérateur Mental (**M**) peut être entièrement anticipé par l'utilisateur dans l'opérateur qui le précède, alors supprimer l'opérateur Mental (**M**). Par exemple, pointer une zone avec la souris et appuyer sur une touche du clavier peut être entièrement anticipé par l'utilisateur. La séquence d'actions Pointing (**P**), Mental (**M**) et Keystroke (**K**) devient une séquence Pointing (**P**) et Keystroke (**K**) (**PMK** → **PK**).
- **Règle 2** : si une chaîne de Mental (**M**) et Keystroke (**K**) appartient à une « unité cognitive » (p. ex. le nom d'une commande dans un éditeur) alors supprimer tous les opérateurs Mental (**M**) sauf le premier.

- **Règle 3** : si la touche du clavier (**K**) pour finaliser la saisie d'un argument d'une méthode et pour déclencher la commande est identique, alors supprimer l'opérateur Mental (**M**) devant le Keystroke (**K**).
- **Règle 4** : si la chaîne de caractères est une commande connue du système (c'est-à-dire interprétée par le système comme des mots-clés) alors supprimer l'opérateur Mental (**M**) devant le dernier Keystroke (**K**) ; mais si la chaîne de caractère n'est pas un mot-clé alors garder l'opérateur Mental (**M**).

Le modèle KLM est composé d'opérateurs, d'unités de temps et d'heuristiques permettant de modéliser la tâche à réaliser par l'utilisateur. Le modèle KLM est limité aux environnements de bureau, mais l'émergence de nouveaux dispositifs ou les besoins des utilisateurs ont conduit les chercheurs à concevoir des extensions du modèle KLM. Les concepteurs d'extensions de KLM doivent alors adapter les opérateurs, les unités de temps, les heuristiques et leurs méthodes d'évaluation aux nouveaux contextes d'utilisation. Nous présentons les principales extensions et leur mode de fonctionnement dans la section suivante.

3.3 Extensions du modèle KLM

Étendre un modèle KLM permet de prédire le temps de réalisation d'une tâche soit avec un système spécifique différent d'un interpréteur de commande soit ayant lieu dans un environnement différent. La conception d'un nouveau modèle KLM peut prendre plusieurs formes. En effet, chaque partie du modèle peut être partiellement ou entièrement modifiée. Dans cette section, nous détaillerons les principaux dispositifs interactifs qui ont fait l'objet d'une extension du modèle KLM. Ensuite, nous discuterons les différentes parties du modèle qui sont modifiées dans la littérature : les opérateurs, les unités de temps et les heuristiques. Enfin, nous expliquerons les différentes évaluations et méthodes statistiques utilisées pour valider l'extension du modèle.

3.3.1 *Les extensions de KLM pour de nouveaux dispositifs interactifs*

L'une des raisons d'étendre le modèle KLM est d'évaluer de nouvelles technologies ou applications qui n'étaient pas prises en compte dans le modèle KLM original (Al-Megren et al., 2018). Dans cette section nous présentons plusieurs exemples d'extensions KLM à de nouveaux systèmes interactifs : les systèmes embarqués dans un véhicule et les téléphones mobiles ou smartphones.

Les systèmes embarqués dans un véhicule

Les systèmes interactifs ont été intégrés par des concepteurs dans de nombreux outils de notre vie quotidienne. Par exemple, les voitures sont équipées d'interfaces interactives permettant de contrôler la musique ou le système de navigation. Les concepteurs de systèmes interactifs embarqués doivent réduire le temps de réalisation des différentes tâches pour ne pas surcharger l'utilisateur qui est en train de conduire. De nombreux modèles ont été conçus pour prédire le temps de réalisation de différentes tâches (utilisation du système de navigation GPS ou de navigation dans les menus composant l'interface (S. C. Lee et al., 2019; Manes et al., 1997; Manes & Green, 1997; Nowakowski et al., 2000; Schneegaß et al., 2011)). Par exemple, Green (P. Green, 1999) propose une extension de KLM pour évaluer des tâches de navigation (GPS) selon la règle des 15 secondes (c'est-à-dire que la tâche de navigation ne doit pas prendre plus de 15 secondes pour ne pas empiéter sur l'activité principale de conduite).

Les téléphones mobiles et smartphones

L'avènement des smartphones et des téléphones mobiles a conduit les chercheurs à proposer de multiples extensions pour l'interaction avec un téléphone mobile. Par exemple, Holleis et collab. (Holleis et al., 2007), conçoivent un modèle KLM pour l'interaction avec un téléphone mobile (cf. Figure 2.20). Ce modèle peut décrire à la fois des tâches classiques (p. ex. écrire un message) ou des tâches avancées liées au contexte d'utilisation du téléphone (p. ex. prendre en photo des marqueurs, à l'instar des QR codes, placés dans l'environnement). Plus récemment, d'autres travaux ont exploré les extensions de KLM pour des smartphones avec un écran tactile (El Batran & Dunlop, 2014; A. Lee et al., 2015; Rice & Lartigue, 2014).



Figure 2.20 – Illustration du besoin spécifique de création d'un modèle KLM pour l'interaction avec un téléphone mobile avec (A) une utilisation dans un environnement public et (B) la spécificité du téléphone mobile avec trois zones d'interaction (Holleis et al., 2007)

Prédire le temps de réalisation d'une tâche avec un nouveau système interactif requiert la conception d'un nouveau modèle. Ce nouveau modèle peut alors définir de nouveaux opérateurs et leurs unités de temps associées. Ces nouveaux opérateurs peuvent soit modéliser une action physique (p. ex. l'appui sur une touche d'un téléphone) ou une action cognitive (p. ex. le changement d'attention entre un dispositif et son environnement). La modélisation

3.3.3 Définition des unités de temps

L'unité de temps associée à un opérateur modélise le temps que l'utilisateur mettra pour réaliser l'action modélisée par l'opérateur. Cette unité de temps peut être unique ou dépendre d'un ou plusieurs paramètres (p. ex. l'opérateur Keystroke du modèle KLM original dépend du niveau de compétence de l'utilisateur). Il est alors possible d'associer plusieurs valeurs à cet opérateur pour modéliser différentes valeurs de ce paramètre. Une autre approche possible est d'utiliser une équation, à l'instar de la loi de Fitts (Fitts, 1954) pour modéliser une unité de temps associée à un opérateur. Dans cette section, nous détaillerons ces trois approches en les illustrant par des exemples.

Unité de temps unique

L'association d'une seule unité de temps à un opérateur semble l'approche la plus commune dans le modèle KLM original. En effet, cette approche est utilisée pour trois de ses six opérateurs : Pointing, Homing, Mental Act. Dans les extensions du modèle KLM pour de nouveaux dispositifs, cette approche est aussi répandue. Par exemple, Erazo et Pino (Erazo & Pino, 2015) associent une seule unité de temps pour les opérateurs Gripping et Releasing (ouverture et fermeture de la main dans une interface gestuelle avec la main). Holleis et collab. (Holleis et al., 2007) utilisent cette approche pour certains de leurs opérateurs pour leur extension de KLM au téléphone mobile. L'opérateur Macro Attention Shift correspond au changement d'attention entre le téléphone et l'environnement (p. ex. un bruit dans l'environnement interrompant l'utilisateur) et est défini avec une seule unité de temps. Cette approche permet aux concepteurs d'interface de facilement construire une modélisation KLM pour prédire le temps de réalisation d'une tâche. Cependant, elle peut manquer de précision si l'action réalisée est dépendante d'un autre facteur.

Plusieurs unités de temps associées à un opérateur

Certains opérateurs se définissent en fonction d'un paramètre permettant d'affiner la modélisation. Dans le KLM original, Card et collab. (Card et al., 1980) définissent plusieurs unités de temps pour l'opérateur Keystroke selon soit le niveau de compétence de l'utilisateur soit le type de touche appuyée sur le clavier. Cette approche se retrouve dans plusieurs extensions récentes. Erazo et Pino (Erazo & Pino, 2015) définissent trois unités de temps pour leurs opérateurs Swiping et Swipe preparation (modélisant respectivement le mouvement de la main dans une direction et le geste de préparation à ce mouvement) : pour les mouvements horizontaux, les mouvements verticaux et une moyenne des deux unités de temps précédentes. Schneegaß et collab. (Schneegaß et al., 2011) modélisent l'opérateur Turn (l'action de tourner

un bouton rotatif) selon six valeurs de rotation : -180° , -90° , -45° , 45° , 90° , 180° . Le paramètre de rotation de cette extension KLM est décomposé en six valeurs discrètes dans le but de simplifier l'utilisation du modèle. Cependant, ce paramètre peut aussi représenter une valeur continue. Dans ce cas, associer une infinité de valeurs à un opérateur nécessite de mettre en place une équation pour, à partir de n'importe quelle valeur que prendra le paramètre, définir une unité de temps utilisable par le modèle.

Définition d'unités de temps associées à un opérateur selon une équation

Modéliser le temps de réalisation d'une action selon une équation paramétrique est une méthode très utilisée dans le domaine de l'Interaction Humain-Machine. Par exemple, les tâches de pointage ont été modélisées au travers de la loi de Fitts (Fitts, 1954) et de ses dérivés (Ens et al., 2016; Jagacinski & Monk, 1985; Jin et al., 2020; Kopper et al., 2010). Ces équations permettent, en indiquant les valeurs des paramètres de l'équation, de retourner une unité de temps modélisant le temps de réalisation de la tâche. Dans le modèle KLM original, l'opérateur Drawing modélise le temps de réalisation de dessin selon 2 paramètres continus : le nombre de segments et la longueur des segments composant le dessin. El Batran et Dunlop (El Batran & Dunlop, 2014), modélisent trois opérateurs pour l'interaction avec l'écran tactile d'un smartphone : Swipe (le mouvement d'un doigt dans une direction), Tap (l'appui sur l'écran du téléphone en fonction de la distance entre la position du doigt et l'objet ciblé) et Zoom (le rapprochement ou l'éloignement de deux doigts sur l'écran tactile du smartphone). À l'instar de la loi de Fitts, les unités de temps associées à ces opérateurs sont modélisées à l'aide d'équations selon l'indice de difficulté de la tâche. Cette approche permet d'avoir une unité de temps très précise en fonction des différents paramètres de l'équation. Cependant, cela peut être difficile à utiliser pour des concepteurs non expérimentés.

Définition d'unités de temps selon plusieurs approches

Certaines extensions du modèle KLM proposent les trois méthodes pour définir l'unité de temps associée à chaque opérateur. Lee et collab. (A. Lee et al., 2015), modélisent l'interaction avec un smartphone dans des applications de jeux mobiles. Leur modèle permet de modéliser quatre

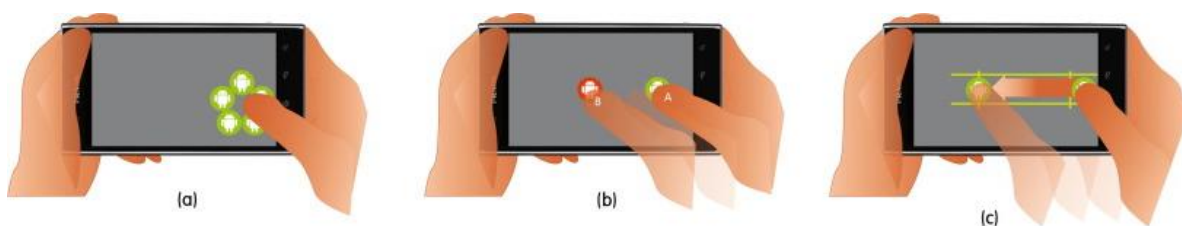


Figure 2.22 – Les opérateurs (a) Tapping, (b) Pointing et (c) Dragging de l'extension KLM proposée par (A. Lee et al., 2015)

opérateurs : Tapping, Pointing, Dragging et Flicking. Les trois premiers opérateurs mentionnés (cf. Figure 2.22) proposent à la fois une unité de temps moyenne, plusieurs unités de temps selon plusieurs valeurs discrètes des paramètres et une équation. Cette modélisation permet aux concepteurs de choisir selon leurs besoins.

La création d'un modèle KLM nécessite donc de créer des opérateurs et de définir leurs unités de temps associées. Pour modéliser une tâche à l'aide d'un modèle KLM, la tâche doit être décomposée en une suite d'opérateurs physiques (c'est-à-dire correspondant à des mouvements moteurs effectués par les utilisateurs). Cependant, les opérateurs cognitifs ou mentaux ne peuvent pas être observés et dépendent donc de l'utilisateur et de la tâche qui doit être réalisée. Pour placer les opérateurs cognitifs, il convient de suivre des heuristiques indiquant l'endroit où les opérateurs mentaux doivent se retrouver dans la modélisation de la tâche.

3.3.4 Les heuristiques

Le placement des opérateurs mentaux dans les modélisations KLM dépend des compétences et des connaissances des utilisateurs (Card et al., 1980). Ces actions mentales sont difficilement observables. Le placement des opérateurs mentaux est alors à l'appréciation des personnes en charge de la conception. Kieras (Kieras, 2001) enrichit les heuristiques de KLM de suggestions concernant les opérateurs mentaux. Par exemple, Kieras suggère que le nombre d'opérateurs mentaux est plus important que leur placement au sein de la modélisation et qu'il faut garder une consistance entre chaque placement ou suppression d'opérateurs mentaux.

La création de nouveaux opérateurs implique la création de nouvelles heuristiques ou l'adaptation des heuristiques originelles. Holleis et collab. (Holleis et al., 2007), dans leur modèle KLM pour les interactions avec un téléphone, ajoutent à la règle 0 (celle qui ajoute les opérateurs **M** avant les opérateurs **K** et **P**) trois de leurs nouveaux opérateurs : Homing, Macro Attention Shift et Gesture. En plus, une règle, provenant des suggestions de Kieras, est ajoutée aux heuristiques originales : « Règle 5 : si vous n'êtes pas sûr, priorisez le nombre d'opérateurs **M** plutôt que leur placement ». Des changements identiques peuvent être retrouvés dans le modèle proposé par Erazo et Pino (Erazo & Pino, 2015) en ajoutant une condition à une des règles selon le niveau de compétence des utilisateurs.

Ces règles permettent aux concepteurs d'anticiper et d'être consistants pour modéliser des actions qui ne peuvent être visuellement observées. L'ajout d'opérateurs mentaux est primordial pour concevoir une modélisation KLM adéquate pour prédire les temps de réalisation d'une

tâche. Une fois le modèle créé, il est nécessaire d'évaluer sa capacité à prévoir le temps de réalisation d'une tâche.

3.3.5 *Les évaluations de modèles KLM*

L'évaluation d'un modèle KLM consiste à comparer le temps prédit par le modèle et le temps réalisé par des utilisateurs. Il est alors nécessaire de modéliser une ou plusieurs tâches avec le modèle. Card et collab. (Card et al., 1980), évaluent le modèle KLM en combinant 10 systèmes et 14 tâches. Chaque combinaison de systèmes et de tâches permet de comparer le temps observé et le temps prédit. Cette méthode d'évaluation du modèle est prédominante dans les différentes extensions de KLM (Erazo & Pino, 2018; Holleis et al., 2011, 2007; A. Lee et al., 2015; S. C. Lee et al., 2019; Myung, 2004; Schneegaß et al., 2011).

Les résultats de ces études sont ensuite statistiquement étudiés selon différentes méthodes : analyse de corrélation, pourcentage de différence, ratio, RMSE (Racine de l'erreur quadratique moyenne), etc. (Al-Megren et al., 2018). Afin d'analyser les résultats de ces différentes méthodes, différents critères permettent d'indiquer si le modèle est valide. Par exemple, Olson et Olson (Olson & Olson, 1990) proposent un critère d'erreur de 20% pour les pourcentages de différence et les ratios et Card et collab. (Card et al., 1980) proposent une échelle entre 20 et 30% pour le RMSE.

Parmi les différents types de modèles conçus pour évaluer des systèmes interactifs, le modèle KLM semble être l'un des plus intéressants. En effet, il est à la fois facile d'utilisation pour les concepteurs et permet de définir différents niveaux de précision pour décrire les tâches réalisées par l'utilisateur. Le modèle KLM a fait ses preuves dans de nombreux domaines (smartphone, transport, interfaces gestuelles avec une Kinect). Dans cette section, nous avons pu voir que la grande variété d'opérateurs (physiques ou mentaux) et d'unités de temps qui peuvent leurs être associées confère une grande profondeur au modèle. L'utilisation d'heuristiques basées sur des travaux sur le comportement humain avec des systèmes interactifs permet aux concepteurs de comprendre quand les utilisateurs devront réfléchir à leurs prochaines actions. Enfin, la validation de ces modèles au travers d'une comparaison entre le temps prédit et le temps observé est primordiale pour indiquer l'efficacité du modèle aux personnes souhaitant l'utiliser par la suite.

3.4 Résumé

Cet état de l'art a permis de mettre en avant plusieurs propriétés intéressantes de la conception et l'évaluation d'un système de visualisation de données situées. La technologie utilisée pour

réunir un référent physique et des données numériques entraîne différentes problématiques. La réalité augmentée mobile oblige à voir les données numériques au travers d'une petite fenêtre (l'écran du dispositif). D'autre part, la réalité augmentée par vidéoprojection ne s'adapte que très peu à des environnements dégradés.

Pour interagir avec ces systèmes de visualisations de données situées, plusieurs solutions existent. L'interaction au contact est prédominante dans les systèmes de visualisation par vidéoprojection mais elle souffre de trois problèmes majeurs : l'occultation, l'atteignabilité et la disponibilité d'accès. L'interaction à distance est alors une solution adéquate pour pallier ces problèmes. Cependant, plusieurs facteurs permettent de définir plus précisément l'interaction à distance : la présence d'une relation spatiale entre le mouvement et le référent physique ou les données numériques.

Concevoir et évaluer un système de visualisation de données situées interactif est alors complexe selon les méthodes traditionnelles d'études utilisateur. Une autre approche possible est d'évaluer les composantes d'un système à l'aide d'un modèle de prédiction. Le modèle KLM de la famille GOMS permet d'évaluer le temps de réalisation d'une tâche en amont de son développement. Cependant, ce modèle n'est pas conçu pour les systèmes de visualisations de données situées. En s'inspirant du modèle KLM, plusieurs extensions ont vu le jour pour s'adapter à de nouveaux dispositifs ou contextes d'utilisation. À notre connaissance, aucun d'eux n'est conçu pour modéliser l'interaction avec un casque de réalité augmentée. Dans les chapitres suivants, nous présentons : 1) une technique d'interaction pour interagir avec un référent physique à distance pour un système de visualisation de données par vidéoprojection (Chapitre TouchGlass) et 2) un modèle de prédiction basé sur KLM pour l'interaction avec des données situées dans l'espace numérique (Chapitre KLM).

Chapitre 3 Contexte et Outils

Contenu du Chapitre	
1 Contexte : neOCampus.....	43
1.1 Les types de données	44
1.2 Les types de référents	45
2 Outil Logiciel	46
2.1 Scénarios d’usages et exigences des maquettes interactives	47
2.2 Exigences pour la construction de maquette interactive.....	49
2.3 Les maquettes physiques	50
2.4 Interaction et visualisation.....	52
2.5 Processus de création modulaire.....	53
2.7 Synthèse	60
3 Outil Conceptuel.....	61
3.1 L’espace de conception	62
3.2 Illustration du pouvoir descriptif de notre espace.....	66
3.3 Illustration du pouvoir comparatif de notre espace	71
4 Synthèse.....	78

Dans ce chapitre, nous détaillerons le contexte dans lequel s'inscrivent les travaux de cette thèse : le projet neOCampus. Ce projet de « campus intelligent » a été imaginé au départ à l’université Toulouse III Paul Sabatier puis s’est étendu à d’autres universités locales. Ce projet regroupe de multiples acteurs venant du monde industriel ou universitaire. Pour concevoir et développer des solutions adaptées à plusieurs acteurs du projet, nous avons dû mettre en place deux outils permettant de faire émerger des contributions pertinentes pour ce projet :

- Le premier est un outil logiciel pour aider les concepteurs de visualisations de données situées à concevoir un système complet. En partant d'une modélisation physique d'un bâtiment ou du campus en Lego jusqu'à la projection interactive de données sur cette maquette, cet outil permet d'accompagner le concepteur tout le long du processus de création d'une visualisation de données situées.
- Le second est un outil conceptuel permettant de décrire les systèmes de visualisation de données situées selon plusieurs dimensions pertinentes. Cet outil permet de décrire et de comparer différentes approches dans la conception d'un système de visualisation de données situées proches d'un référent physique.

1 Contexte : neOCampus

Cette thèse s'inscrit dans le cadre du projet neOCampus, une initiative de l'Université de Toulouse, lancée en juin 2013 dans le but de créer, à l'instar d'une ville intelligente, un campus intelligent. Les objectifs du projet sont triples :

1. Améliorer le confort quotidien des étudiants et du personnel de l'université
2. Diminuer l'empreinte écologique de ses bâtiments
3. Réduire ses coûts d'exploitation (fluides, eau, électricité, etc.).

Pour atteindre ses objectifs, le projet neOCampus s'appuie sur l'utilisation du grand nombre d'appareils connectés détenus par les utilisateurs (smartphones, smartwatches, tablettes) et un réseau de capteurs connectés, pour mieux collecter et exploiter les données. À l'instar d'une petite ville avec ses 407 000m² de surface bâtie, 70 structures de recherche, plusieurs solutions de mobilité et plus de 39000 employés et étudiants, l'amélioration de la qualité de vie à l'intérieur du campus de l'Université de Toulouse est un défi majeur. Pour le relever, le projet neOCampus s'appuie sur le grand nombre de dispositifs (appareils et capteurs) présents sur le campus. En effet, les dispositifs et capteurs collectent une grande quantité de données hétérogènes en temps réel. Cette grande quantité de données nécessite la création de nouveaux outils pour les explorer, les gérer et les analyser.

Les défis scientifiques sont nombreux (stockage, analyse, réseau, détection d'anomalies, etc.) et requièrent un large éventail de compétences pour être relevés. Pour atteindre ses objectifs et relever ces défis, le projet neOCampus a favorisé une approche multidisciplinaire avec 11 laboratoires participant au projet (une liste détaillée des laboratoires participants peut être

trouvée ici⁷). Chaque laboratoire apporte sa propre expertise scientifique et l'université est utilisée comme une plateforme expérimentale pour réaliser des tests à grande échelle et dans des conditions réelles.

Le groupe de recherche Elipse (Étude de L'Interaction Personne-Système) de l'Institut de Recherche de Toulouse (IRIT) fait partie intégrante de ce projet. Ce groupe de recherche spécialisé en Interaction Humain-Machine apporte ses compétences de conception et d'évaluation de systèmes innovants et interactifs pour visualiser de grandes quantités de données hétérogènes. Le groupe de recherche Elipse s'inscrit donc parfaitement dans le projet neOCampus et focalise son apport sur l'exploration et la visualisation de ces données hétérogènes fournies par tous les capteurs présents sur le campus de l'université.

Cette thèse menée au sein de ce groupe et de ce projet a donc pour but de concevoir, développer et évaluer des solutions analytiques ou logicielles pour faciliter l'accès, la manipulation et la compréhension de ce large éventail de données en s'inspirant de nouveaux paradigmes comme la visualisation de données situées. Dans ce contexte, neOCampus offre de nombreuses données et de nombreux référents différents. Plusieurs applications de ces référents et de ces données sont présentées en Annexe.

1.1 Les types de données

Pour réaliser les objectifs d'un campus plus agréable et limitant son empreinte écologique, des capteurs récoltant des données en temps réel ont été installés dans des endroits stratégiques de l'université. Afin de limiter la consommation énergétique des bâtiments, des capteurs ont été installés sur des bornes à l'extérieur des bâtiments ou sur des compteurs à l'intérieur d'un étage. Un capteur permet de collecter un type de données précis. Par exemple, des capteurs de consommation d'électricité, de gaz ou d'eau permettent de collecter la consommation d'un bâtiment en temps réel ou plus finement d'un étage. Ces données peuvent être utilisées par l'économe des flux de l'université qui va pouvoir les analyser afin de détecter des anomalies (p. ex. une consommation d'électricité très importante dans un bâtiment en période de vacances scolaires) ou analyser des tendances.

Pour améliorer la qualité de vie des étudiants et des enseignants, des capteurs ont été placés à l'intérieur des salles. Ces capteurs récoltent de nombreuses données : qualité de l'air, température, luminosité, humidité, nombre de personnes dans une pièce, etc. Ces capteurs se

⁷ <https://www.irit.fr/neocampus/fr/>

trouvent à l'intérieur des salles de classe ou de réunion et peuvent être utilisés pour analyser l'activité d'une salle ou déclencher des événements comme l'abaissement des rideaux électriques si la luminosité est trop forte. La température peut également être analysée pour déclencher ou arrêter l'activité du chauffage central d'une partie d'un bâtiment.

Toutes ces données sont relatives à différents environnements ou objets collectant les données. Situer ces données proches de leur contexte d'acquisition pour les représenter afin de mieux les comprendre et les analyser (cf. chapitre État de l'art section 1) nécessite donc de bien analyser les différents référents physiques associés à ces données.

1.2 Les types de référents

Transformer une représentation de données « classiques » en une représentation de données situées nécessite de les représenter à proximité de leur référent physique (de l'endroit, de l'objet ou du contexte) d'où elles proviennent. Pour représenter d'une manière située des données relatives à un campus, différentes échelles ou granularités sont à explorer :

1. Le **capteur** est le plus petit composant auquel une donnée peut être liée. C'est lui qui récolte les données. Il peut être placé à l'intérieur d'une salle, dans une zone spécifique d'un étage ou d'un bâtiment ou à l'extérieur d'un bâtiment (à l'instar d'un compteur électrique ou de gaz).
2. La **salle** est une composante essentielle d'un campus universitaire. Elle est généralement composée de matériel et, dans le cas de notre campus intelligent, de plusieurs capteurs placés à différents endroits stratégiques. Par exemple, dans une salle de classe, les capteurs de luminosité permettant le contrôle automatique des volets vont être placés au niveau des étudiants et les capteurs permettant de compter le nombre d'étudiants présents dans la pièce seront placés proches des portes d'entrée.
3. Un **groupe de salles** est une entité qui peut être physique (un étage, un bâtiment) ou conceptuelle (p. ex. les bureaux d'une équipe de recherche). On y trouve des capteurs permettant d'avoir des données de plus grandes granularités comme la consommation énergétique ou le taux d'occupation d'un bâtiment.
4. Le **campus** est l'entité la plus large de notre spectre de référents physiques. On y trouve des données d'encore plus gros grains. On peut y représenter les flux de circulation, des données météorologiques ou encore l'emplacement des bâtiments.

Représenter des données d'une manière située avec cette pluralité de référents physiques demande d'adapter les techniques d'interaction et les technologies de représentation de

données. Par exemple, si l'on veut visualiser les données de tous les capteurs d'un bâtiment, il semble complexe de les afficher sur les capteurs et que l'utilisateur se déplace dans tout un bâtiment pour les récupérer puis les analyser. Une solution possible, et présentée dans le chapitre État de l'art (cf. section 1.2), est l'utilisation d'une maquette réduite d'un bâtiment.

2 Outil Logiciel

Les maquettes réduites de bâtiments sont un outil puissant pour visualiser les données dans leur contexte d'acquisition quand celui-ci est trop volumineux, mais aussi pour communiquer des informations à des non-experts (Jansen et al., 2013). Cependant, concevoir un système interactif de visualisation de données situées nécessite :

1. Une maquette physique du bâtiment,
2. Une maquette numérique et
3. Un lien logiciel entre physique et numérique.

Architectes et promoteurs produisent de telles maquettes dans leur tâche classique mais sculpter des maquettes physiques ou concevoir des maquettes numériques n'est pas à la portée de nombreux utilisateurs. Notre approche se base alors sur une construction en briques Lego⁸. Les maquettes en briques Lego ont l'avantage d'être facilement concevables et il existe divers logiciels^{9,10} permettant de créer des modèles numériques en manipulant des briques Lego numériques.

Notre système utilise la réalité augmentée par vidéoprojection pour rendre les données numériques proches du monde physique sans avoir besoin d'un dispositif de visualisation additionnel (p. ex. une tablette ou un smartphone). Pour faire le lien entre le monde physique et le monde numérique nous avons implémenté un module Unity permettant de gérer la position des vidéoprojecteurs qui viendront illuminer le modèle physique, une liste de *tags* permettant de grouper des briques Lego pour construire des salles, des étages ou des bâtiments.

Cette section se décompose alors en une explicitation des scénarios qui ont motivé la construction de cet outil et de ses exigences. L'approche basée Lego, en plus de la facilité de conception, offre une modularité qui peut être utilisée pour la conception d'un système de visualisation de données sur une maquette d'un bâtiment ou d'un campus. En effet, plusieurs

⁸ <https://www.lego.com/fr-fr>

⁹ <https://www.bricklink.com/v3/studio/download.page>

¹⁰ <https://www.lego.com/en-us/ldd>

changements peuvent être appliqués sur la maquette d'une salle ou d'un bâtiment. Cette modularité s'exprime aussi dans la diversité des données ou des différentes maquettes physiques à concevoir. Cette section inclut donc une description de l'interaction et de la visualisation modulaire. Ensuite nous explicitons notre outil et le cheminement que l'utilisateur doit suivre pour créer un système de visualisation interactive sur une maquette. Enfin, nous montrons comment cet outil a pu être mis en place pour construire trois applications différentes.

2.1 Scénarios d'usages et exigences des maquettes interactives

Pour illustrer concrètement l'intérêt d'un outil facilitant la conception d'une visualisation de données situées sur une maquette physique nous proposons différents scénarios issus de différents entretiens. Dans le contexte d'un campus intelligent, nous avons interrogé deux architectes, un BIM manager (Building Information Modeling¹¹), une économiste des flux, la responsable du projet neOCampus et des usagers du campus (p. ex. étudiants, salariés, professeurs). Ces entretiens ont permis de faire émerger des scénarios d'usage. Nous les présentons ci-après.

2.1.1 Scénario 1 : Visualisation de la consommation énergétique

Contrôle instantané : L'économiste des flux doit avoir accès en permanence à des données actuelles ou antérieures pour comparer des valeurs, détecter des anomalies ou adapter les besoins en énergie. Elle répète ces tâches plusieurs fois par jour pour comparer les données de salles, étages, bâtiments ou groupe de bâtiments entre elles, mais aussi pour évaluer leur évolution avec des données historiques. Quotidiennement, l'économiste des flux doit alors faire des allers-retours entre des tableurs pour analyser les données et une carte du campus pour vérifier d'où proviennent ces données (d'un bâtiment, d'un capteur ou d'un compteur électrique).

La visualisation de données situées est alors une solution intéressante car elle permet de spatialement mettre en relation les données et d'où elles proviennent (les référents physiques). Cependant, afficher puis aller visualiser et analyser ces données sur les bâtiments, compteurs ou capteurs réels est impossible. L'utilisation d'une maquette physique peut donc être un facilitateur permettant une vue d'ensemble rapide des données supportant la tâche de comparaison, et offrant des points de repère physiques. De plus, les campus étant évolutifs (destruction, construction ou rénovation de bâtiment) cette maquette doit être modulaire.

¹¹ Le BIM est un processus permettant au travers d'une maquette numérique de caractériser de manière presque illimitée chaque objet d'un écrou à un mur composant un bâtiment

Sensibilisation du public et amélioration de la qualité de vie : Réduire la consommation énergétique d'un campus dépend en partie de la sensibilisation de ses usagers. Il est primordial de faire prendre conscience au personnel et aux étudiants de leur impact écologique sur le campus. Ces données sont généralement disponibles sur des sites web qui peuvent avoir du mal à susciter de l'engagement chez les utilisateurs (Piccolo et al., 2014). Or, la visualisation de données situées peut-être un moyen d'engager l'utilisateur pour des données énergétiques (Jönsson et al., 2010). Dans ce cas, l'utilisation d'une maquette réduite du campus placée dans un des bâtiments peut permettre à tout le monde de prendre connaissance de la consommation énergétique d'un ou de plusieurs bâtiments. L'étudiant voulant s'inscrire à l'université ou la chercheuse entrant dans son laboratoire verrait la maquette dans le hall du bâtiment et viendrait découvrir la consommation énergétique de la semaine passée du campus ou de son équipe de recherche. De plus, d'autres types de données peuvent aisément venir en complément des données énergétiques (p. ex. le nombre de personnes attendant au restaurant universitaire, les places libres à la bibliothèque, etc.). Des étudiants voulant réserver une salle de travail à la bibliothèque universitaire pourraient alors entrer dans la bibliothèque universitaire, s'approcher de la maquette, sélectionner une ou plusieurs salles de travail et voir si celles-ci sont disponibles.

Pour engager l'utilisateur et lui fournir de nouveaux services, l'utilisation d'une maquette interactive est une idée intéressante. Cette maquette doit cependant être interactive pour permettre l'exploration des utilisateurs et renforcer leur engagement dans le système. De plus, les usagers interrogés ont également souligné le fait que l'utilisation d'une maquette physique interactive accentuerait l'intérêt des passants pour les données affichées par rapport à un affichage classique de type compteur.

2.1.2 Scénario 2 : Maquette interactive pour l'affichage de données architecturales

Maquette publique : Le BIM manager interrogé a fait référence à un projet récent construit avec un processus BIM et de nouveaux matériaux. L'un des objectifs de ce projet est de proposer une approche originale basée sur une maquette permettant de fournir un outil de communication donnant accès aux données du bâtiment à tous les visiteurs. Ces données peuvent être la composition des nouveaux matériaux, leur évolution dans le temps, mais aussi l'historique du bâtiment en fonction des années. Les visiteurs devraient pouvoir naviguer dans l'ensemble des données en étant proches de la maquette du bâtiment. Une maquette physique interactive pourrait être utilisée comme support d'information pour les visiteurs, mais pourrait aussi faciliter des tâches de maintenance, anticiper certains coûts, etc.

La maquette interactive doit permettre aux utilisateurs d'interagir pour découvrir l'ensemble du bâtiment et des données associées. De plus, cette visualisation doit être modulaire pour pouvoir proposer différents types de données.

Gestion du patrimoine : Tous les bâtiments doivent pouvoir anticiper l'usure et prévoir les coûts de maintenance et futurs travaux à effectuer. Les gestionnaires immobiliers pourraient ainsi profiter d'un modèle physique du bâtiment pour mettre en évidence le lieu des dépenses à planifier afin de mieux communiquer aux propriétaires ces coûts et aider à anticiper le meilleur timing pour engager les travaux les plus coûteux. Par exemple, un gestionnaire immobilier pourrait vouloir anticiper l'évolution de l'éclairage de tout un bâtiment. Il utiliserait alors une maquette interactive où il pourrait sélectionner les dispositifs d'éclairage et visualiser les informations relatives à chaque dispositif (p. ex. date de dernière réparation, coût énergétique, etc.). Un tel support pourrait également être prolongé jusqu'à la fin de vie du bâtiment : le démontage nécessite de connaître l'emplacement des conduits principaux, la taille des murs, etc. Cette maquette doit pouvoir évoluer dans le temps et donc être modulaire.

2.2 Exigences pour la construction de maquette interactive

Ces scénarios nous ont permis de mettre en évidence des besoins pour la mise en place d'un système interactif basé sur un modèle physique 3D supportant la visualisation et l'interaction avec des données situées : modularité et interactivité.

2.2.1 *Modularité*

L'objet et les données qui lui sont associées évoluent au fil du temps. Il est donc important d'avoir une approche modulaire pour représenter ces données et ainsi supporter un processus itératif. Nous définissons la modularité comme un élément clé de notre approche. Elle se trouve à tous les niveaux de notre processus allant de la construction de la maquette en Lego, jusqu'à l'interaction avec les données en passant par la liaison entre le monde physique et le monde numérique.

2.2.2 *Interactivité*

L'interaction 3D peut être difficile pour les utilisateurs novices mais l'utilisation d'un modèle physique constitue un support prometteur, offrant des points de repère, un aperçu global et une référence physique. Ainsi, l'utilisation de simples gestes en l'air ou de dispositifs basés sur l'environnement de la maquette comme surface tactile peut permettre d'interagir avec le modèle physique. Ceci est d'autant plus important que nos scénarios d'utilisation se déroulent dans un contexte public avec des environnements souvent contraignants.

2.3 Les maquettes physiques

Dans un contexte de campus intelligent, l'interaction et la représentation des données situées doivent s'adapter au type de référent physique. En effet, ces maquettes peuvent être difficiles d'accès car trop imposantes ou protégées (cf. chapitre État de l'art section 2.1).

Dans cette section nous présentons plusieurs types de maquettes 3D. Ensuite, nous extrayons une liste de propriétés de briques ou d'un groupement de briques et montrons comment ces propriétés peuvent contribuer à établir un lien entre une brique et des éléments d'une visualisation ancrée sur la maquette. Enfin, nous présentons différentes solutions possibles pour sélectionner une brique ou un groupe de briques.

2.3.1 Les différents types de maquettes

Sur la base des entrevues effectuées avec les architectes, 3 types de maquettes ont été identifiés. Une **maquette 2,5D** permet à l'utilisateur d'avoir une vue aérienne (cf. Figure 3.1). Cette maquette peut contenir un bâtiment (p. ex. un grand centre pour une convention) ou plusieurs (p. ex. un campus, un quartier) et dans ce dernier cas, chaque brique représenterait un bâtiment. Cette maquette permet de mettre l'accent autant sur les bâtiments (les briques) que sur l'environnement de ces briques (p. ex. les routes pour montrer la circulation dans un quartier). Comme dans les visualisations de données situées présentées dans l'état de l'art, ces maquettes peuvent être facilement augmentées par un vidéoprojecteur. Une **maquette 3D fermée** représente la vue extérieure d'un bâtiment, les briques permettent de modéliser les façades, les fenêtres, etc. Cette maquette permet de comprendre la géométrie du bâtiment et l'impact que peut avoir la structure (taille, forme) du bâtiment sur des données comme la consommation énergétique. Dans ces deux premières maquettes, il est cependant difficile de voir l'intérieur d'une salle. Des technologies basées sur de la réalité augmentée (*X-Ray vision*) permettent de visualiser au travers d'une structure occultante (Sandor et al., 2010). Cependant, cette visualisation dégrade la représentation de l'objet physique et potentiellement des données affichées par-dessus. Le dernier type de maquette est une maquette **3D ouverte** (cf. Figure 3.1) combinant la maquette **2,5D** et la maquette **3D fermée**. Cette maquette propose de représenter chaque étage du bâtiment séparé des autres (cf. Figure 3.1). Afin de pouvoir modéliser

l'intérieur du bâtiment, le toit est retiré de la maquette. Les briques peuvent donc représenter le matériel d'une salle, les murs, le sol, etc.

Ces trois types de maquettes permettent donc de représenter différents référents physiques et leurs données associées. Nous utiliserons et illustrerons ces 3 types de maquettes en fin de

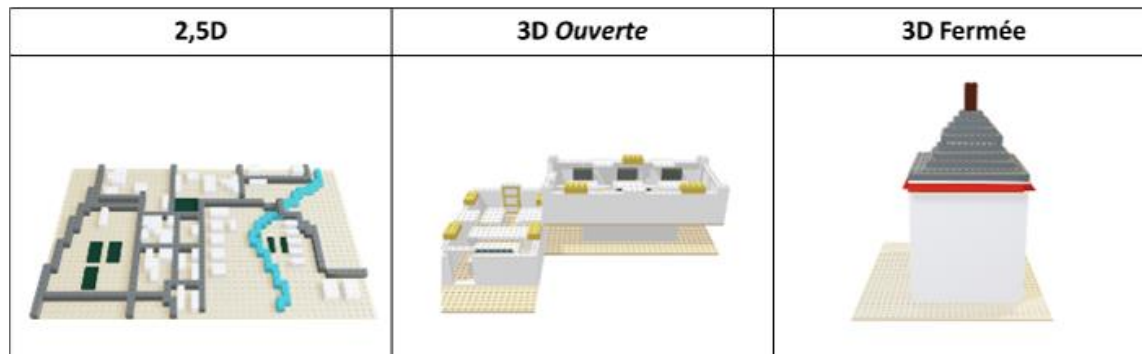


Figure 3.1 – Les trois types de maquettes d'architecture étudiés

section grâce à des scénarios et des mises en application.

2.3.2 Propriétés des briques

Chaque brique Lego peut être décrite selon ses propriétés intrinsèques. Ces propriétés seront utilisées pour l'interaction et la visualisation de données.

Taille : utiliser des briques Lego permet de décrire les pièces et la maquette selon une unité de mesure standardisée. Chaque pièce est donc décrite selon 3 dimensions : largeur, longueur et hauteur. La taille d'une brique Lego d'unité 1x1x1 mesure 0,8×0,8×0,96cm. Cependant, il existe des pièces plus petites (appelées briques « plates ») mesurant 1/3 de la hauteur d'une pièce standard.

Forme : les briques Lego les plus communes sont rectangulaires. Cependant certaines briques sont différentes et peuvent donc représenter des murs, des surfaces planes, des colonnes, etc.

Données associées : cette propriété provient de notre association entre le modèle numérique et la maquette physique. Elle permet de définir si une pièce est « porteuse » d'une donnée ou si elle est uniquement utilisée comme une pièce de décor de la maquette.

Les briques peuvent être regroupées afin de décrire des ensembles plus complexes comme une salle, un étage, ou l'ensemble des bâtiments possédant une même caractéristique, etc. Ce groupe offre donc des propriétés additionnelles utilisables pour l'interaction et pour la visualisation. Sélectionner les propriétés de groupes de briques permettra d'ancrer des données relatives à la

partie du bâtiment représentée par ce groupe de brique (une salle, un étage, etc.). Certaines propriétés peuvent donc servir à décrire un groupe de briques.

Nombre : un groupe de briques Lego peut être décrit par le nombre de pièces qui le compose. Plus le nombre de pièces composant le groupe est important, plus la surface d'affichage devient potentiellement grande et donc l'affichage de données complexes (graphes) sur les briques est possible.

Topologie : La topologie d'un groupe de briques définit la distribution des briques au sein d'un groupe. Les briques Lego pouvant s'emboîter entre elles, la plupart des groupes sont composés de briques adjacentes pour former un étage d'un bâtiment, une salle ou une façade. Cependant il est possible que le groupe soit éclaté au sein de la maquette (p. ex. les différents restaurants universitaires d'un même campus).

2.4 Interaction et visualisation

Nous proposons différentes solutions pour sélectionner une brique ou un groupe de briques pour chaque type de maquette. Nous nous sommes concentrés uniquement sur la tâche de sélection, celle-ci étant fondamentale car elle précède toutes les autres (Looser et al., 2007). La sélection peut prendre plusieurs formes (sans ou avec support de contact, à distance/au contact cf. chapitre État de l'art section 2). Cependant, la sélection au contact dans les systèmes employant la vidéoprojection souffre de problèmes d'atteignabilité, d'occultation et de disponibilité d'accès (cf. chapitre État de l'art section 2.1). Une autre solution possible est d'interagir à distance avec la maquette à l'aide d'un lancer de rayon. Nous détaillons plusieurs solutions proposées par notre outil pour paramétrer et contrôler les lancers de rayon. La ou les briques sélectionnées permettent d'accéder à des données numériques relatives à celles-ci. Nous détaillons ensuite les différentes visualisations proposées par notre outil.

2.4.1 Interaction avec la maquette à l'aide d'un lancer de rayon

La technique de sélection par lancer de rayon doit pouvoir sélectionner une ou plusieurs briques Lego pour pouvoir afficher des informations relatives soit à la brique soit au groupe de briques sélectionnées. Dans le premier cas, nous proposons au concepteur de pouvoir définir un lancer de rayon où la brique sélectionnée est la première brique se trouvant à l'intersection avec le rayon (à l'instar de Mine (Mine, 1995)). Malheureusement, cette technique ne permet pas de sélectionner des cibles occultées par d'autres. Pour ce faire, notre outil offre la possibilité de contrôler un deuxième rayon (à la manière de iSith (Wyss et al., 2006) cf. Figure 3.2).

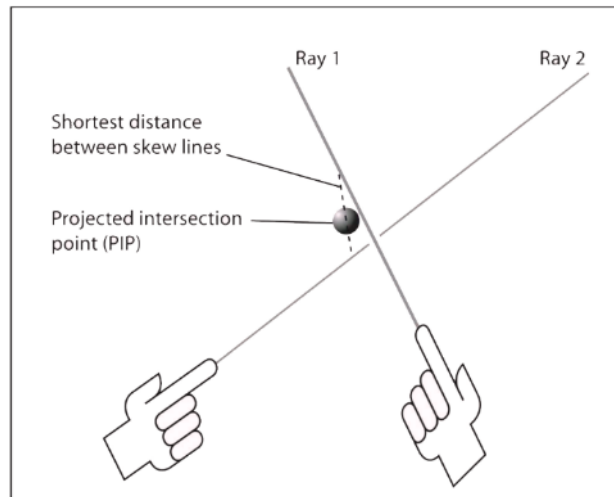


Figure 3.2 – Les deux lancers de rayon définissant un curseur dans le prototype iSith (Wyss et al., 2006)

L'intersection de ces rayons définit alors un curseur. Si ce curseur est en contact avec une brique Lego, celle-ci est sélectionnée.

La sélection de plusieurs briques Lego peut être réalisée à l'aide de plusieurs sélections d'une seule brique. Cependant, cela peut être sous-optimal si de nombreuses briques doivent être sélectionnées. Pour sélectionner des briques adjacentes, notre outil offre la possibilité d'élargir la taille du rayon. Cette sélection peut s'apparenter à la sélection par lancer de cône qui, au lieu de lancer un rayon en ligne droite, utilise un cône comme outil de sélection (Hincapié-Ramos et al., 2015). Ce ou ces rayons élargis permettent d'augmenter la zone de contact et donc de pouvoir sélectionner plusieurs briques adjacentes.

2.4.2 Visualisation située, ancrée et modulaire

La visualisation de données situées proche d'un référent physique consiste à afficher des données soit proches soit directement sur leur source ou une représentation physique de celle-ci. Selon le type de données, elles peuvent être affichées sur la maquette elle-même ou juste à côté. Pour des données simples (p. ex. la température instantanée d'une salle résumée en un chiffre) notre environnement affiche la valeur directement sur la brique. Les données complexes (p. ex. historique d'une donnée, graphe, etc.) sont projetées sous forme de graphe juste à côté de la maquette.

2.5 Processus de création modulaire

Nous proposons donc une implémentation d'un outil logiciel permettant d'associer rapidement et simplement des données numériques aux éléments (briques ou groupe de briques) d'une

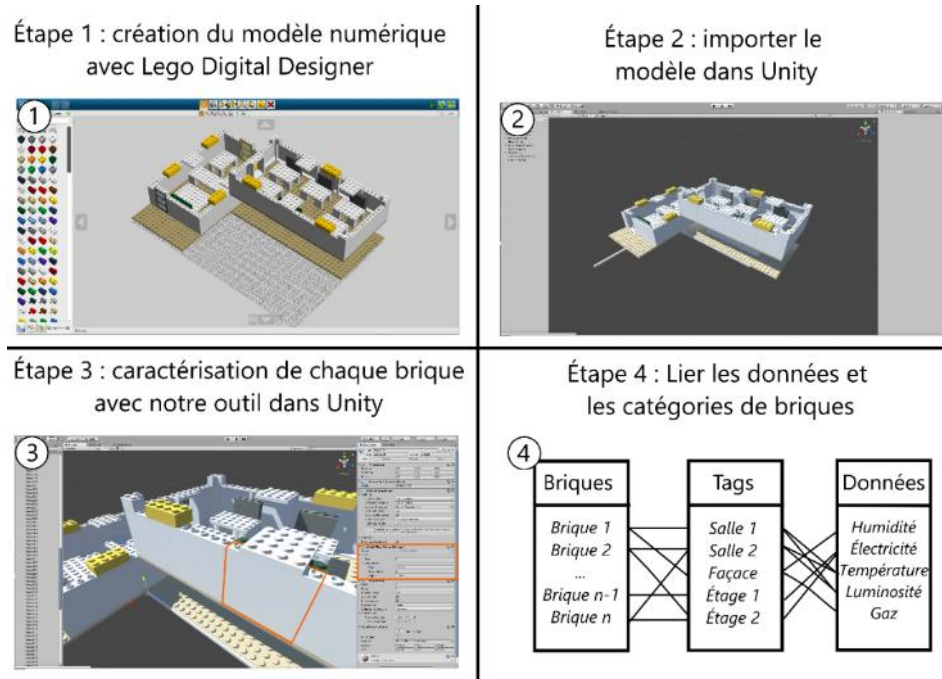


Figure 3.3 – Ensemble de processus de création à l'aide de notre outil logiciel

maquette physique en vue de son utilisation pour la visualisation de données situées. Ce processus comporte 4 étapes résumées dans la Figure 3.3 et a été implémenté dans Unity¹².

2.5.1 Construction physique et numérique de la maquette

La première étape du processus consiste à construire la maquette physique en utilisant des briques Lego et construire sa version numérique à l'aide d'un logiciel tiers (cf. Figure 3.4). Cette solution à bas cout est déjà utilisée par des architectes (Hoggenmueller & Wiethoff, 2016). Dans notre implémentation, nous construisons la maquette numérique en utilisant un logiciel existant, Lego Digital Designer¹⁰. D'autres approches ont été étudiées en utilisant des algorithmes transformant un modèle 3D en modèle Lego (Luo et al., 2015), ce qui peut

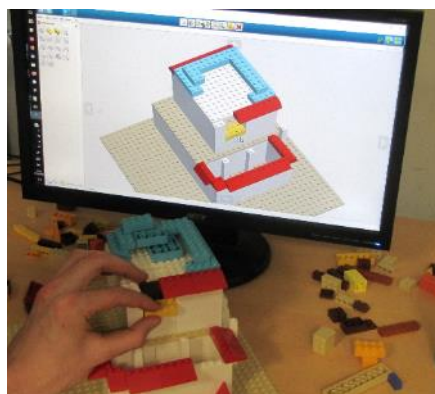


Figure 3.4 – Construction d'une maquette numérique et d'une maquette physique

¹² <https://unity.com/fr>

permettre de générer un modèle Lego à partir d'une maquette déjà conçue avec d'autres environnements de modélisation 3D.

2.5.2 Créer un modèle 3D dans Unity

Cette deuxième étape consiste à créer un environnement interactif 3D basé sur la maquette Lego préalablement modélisée. Pour ce faire, la version numérique de la maquette Lego réalisée à l'étape 1, doit être convertie en un fichier au format wavefront (c'est-à-dire un fichier au format .obj) et importé dans Unity. Chaque brique définie dans Lego Digital Designer devient alors un objet Unity indépendant (cf. Figure 3.5). Ce nouvel objet peut être déplacé, sélectionné et modifié indépendamment du reste de la maquette.

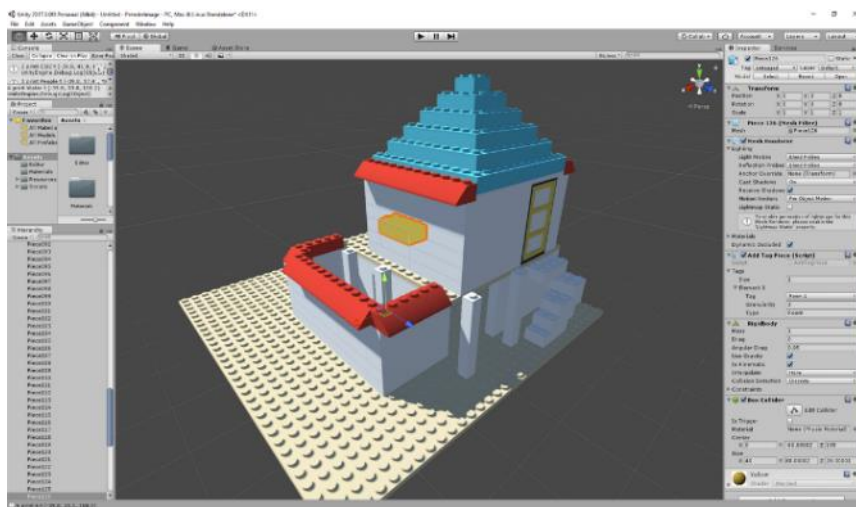


Figure 3.5 – Importation de la maquette numérique dans Unity

2.5.3 Caractériser architecturalement chaque brique

La troisième étape améliore la maquette numérique avec des informations sémantiques caractérisant chaque objet Unity d'un point de vue architectural. L'utilisateur sélectionne chaque brique, soit sur la vue 3D soit dans la liste hiérarchique d'objets proposée par Unity, et spécifie si l'objet est un dispositif, un mur, une fenêtre, etc. Il est aussi possible de définir et caractériser un groupe de briques de la maquette (cf. Figure 3.6). Ce groupe peut alors être caractérisé comme étant une salle, un étage, une façade, etc. Nous obtenons ainsi un fichier XML permettant de lier la définition de la structure architecturale du bâtiment à sa représentation numérique dans Unity.

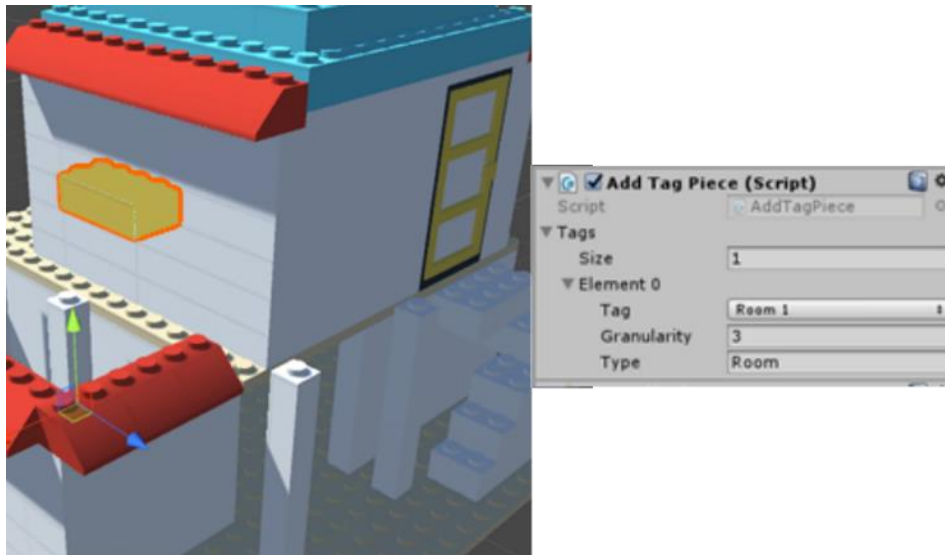


Figure 3.6 – Illustration de notre script permettant de caractériser une brique

2.5.4 *Lier les données aux briques*

La dernière étape pour obtenir un environnement interactif 3D basé sur la maquette physique consiste à lier les objets ou groupes d'objets Unity définis à l'étape 3, à une ou plusieurs données à visualiser. La liste des données disponibles pour cette maquette est automatiquement importée dans Unity à l'aide d'un fichier XML préalablement défini, ou par interrogation d'un serveur de données collectant les données captées. L'utilisateur n'a qu'à sélectionner l'objet ou le groupe d'objets dans la liste déroulante et lui associer la ou les données souhaitées.

L'ensemble du processus (cf. Figure 3.3) a été implémenté comme un Unity Asset Package. Ce package comprend tous les scripts nécessaires à l'import dans Unity ainsi que ceux nécessaires aux étapes de caractérisation des briques.

2.6 Applications

Pour illustrer notre approche, nous avons mis en place de nouvelles techniques d'interaction et notre propre environnement interactif et nous avons développé plusieurs scénarios illustrant soit le contrôle de données de consommation énergétique, soit la maintenance d'un bâtiment au cours du temps conformément aux scénarios d'usage extraits de nos entretiens et décrits en section 2.3. Nous plaçons ces différentes maquettes fabriquées à l'aide de briques Lego à l'intérieur d'une vitrine pour la protéger de toute dégradation.

2.6.1 *Conception des techniques d'interaction pour nos applications*

Visualiser des données situées proches d'une maquette physique demande aux utilisateurs de sélectionner un ou plusieurs éléments de cette maquette pour déclencher l'affichage des

données relatives aux éléments sélectionnés. Comme cela est indiqué précédemment en section 2.6, notre outil offre la possibilité de définir des lancers de rayon pour sélectionner un élément de cette maquette. Généralement, ces lancers de rayon s'effectuent à l'aide de gestes en l'air sujets à la fatigue et aux tremblements (cf. chapitre État de l'art section 2.3). D'autres solutions d'interaction proposent d'utiliser une surface de contact pour interagir avec un référent physique. Notre approche consiste à s'inspirer des deux solutions (lancer de rayon et surface de contact) pour développer nos techniques de sélection. Le rayon est alors projeté à partir de la vitrine protectrice qui sert de dispositif d'interaction et de surface de contact pour limiter les effets de la fatigue et aux tremblements (cf. Figure 3.7). Nous détaillerons en chapitre 4 (TouchGlass) deux études qui explorent l'utilisabilité et la précision d'une telle approche.

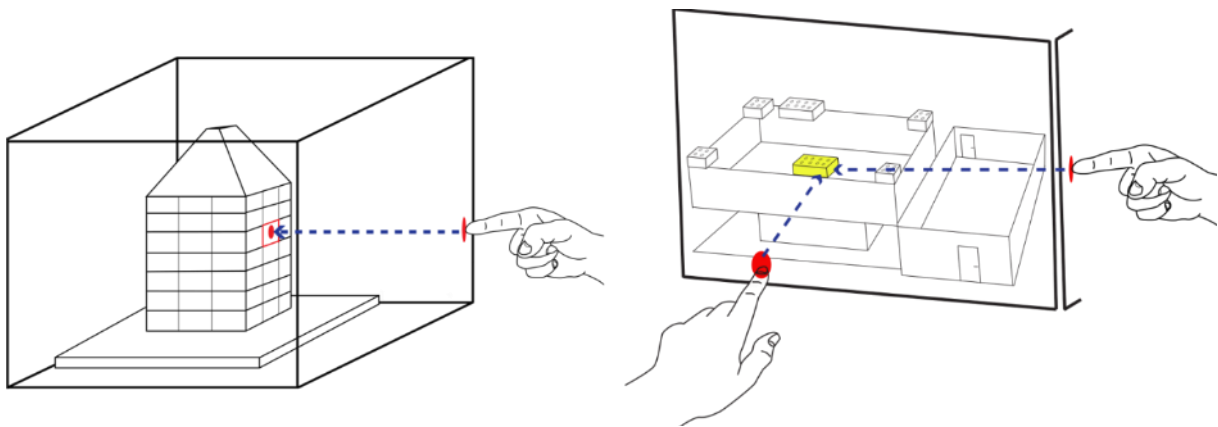


Figure 3.7 – Illustrations des visualisations de données situées sur une maquette avec (A) les techniques de sélection basées sur une vitrine et (B) les différentes visualisations ancrées et situées sur la maquette

2.6.2 Environnement matériel et logiciel

Pour construire la vitrine interactive nous avons assemblé 4 plaques de plexiglas transparentes de 785 x 440 mm liées entre elles par des pièces de jointure imprimées en 3D. Deux côtés de cette vitrine sont augmentés en utilisant des cadres infrarouges (106 centimètres de diagonale avec 40 points de contact développé par Zaag Tech¹³) attachés à la vitrine par des objets imprimés en 3D que nous avons conçus spécifiquement pour ce dispositif et des bandes adhésives Velcro¹⁴. Pour permettre la visualisation de données situées, trois picoprojecteurs d'environ 500 lumens ont été utilisés. Deux d'entre eux sont placés au-dessus des coins opposés de la vitrine. Le dernier est placé verticalement au-dessus de la vitrine afin de projeter des informations à côté de la maquette. Tous les projecteurs sont accrochés à une structure en aluminium (220 x 300 x 300 cm) en utilisant des bras Manfrotto¹⁵ (cf. Figure 3.8).

¹³ http://www.zaagtech.com/index_en.aspx

¹⁴ <https://www.velcro.fr/>

¹⁵ <https://www.manfrotto.com/fr-fr/>

Toutes les applications de démonstrations ont été développées avec Unity sur un système d'exploitation Microsoft Windows 10. Les trois picoprojecteurs sont connectés à un unique ordinateur HP EliteBook à l'aide d'un hub DisplayPort - HDMI.



Figure 3.8 – L'environnement matériel pour notre maquette augmentée

2.6.3 *Contrôle instantané de la consommation énergétique d'un campus ou d'un bâtiment*

Dans le cadre du projet neOCampus, il est important de pouvoir mesurer l'empreinte écologique du campus et de ses composantes. Pour ce faire, chaque bâtiment est équipé de capteurs permettant de mesurer la consommation instantanée du bâtiment. Ces données sont stockées sur un serveur et peuvent être affichées dans le campus à l'attention des étudiants ou du personnel (cf. Scénario 1 : Visualisation de la consommation énergétique). Afin de faciliter la compréhension et de prendre conscience de l'empreinte écologique de certains bâtiments de l'université, les différents usagers doivent pouvoir sélectionner un ou plusieurs bâtiments pour déclencher l'affichage des différentes données de consommation énergétique.

Ce scénario doit permettre d'avoir une vue d'ensemble rapide pour rendre possible la comparaison entre les données de différents bâtiments. Pour mettre en place ce scénario, nous avons donc créé une maquette 2,5D représentant le campus. Les données liées à ce modèle sont toutes en relation avec la consommation énergétique (p. ex. électricité, eau, gaz). Elles sont modélisées pour chaque bâtiment.

De plus, dans ce contexte, le laboratoire d'informatique a lancé une compétition entre les différentes équipes. Les dépenses énergétiques de chaque équipe sont calculées et affichées à

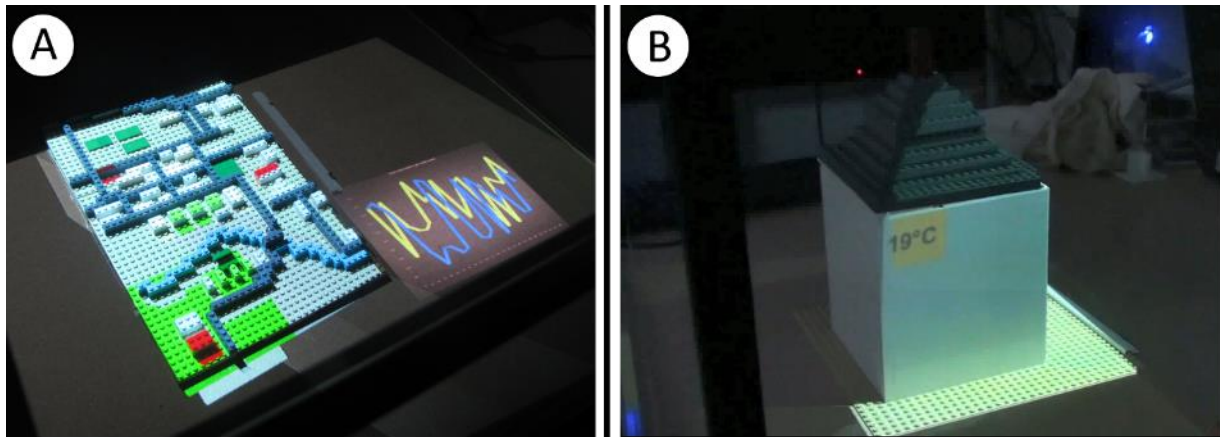


Figure 3.9 – (A) La maquette 2,5D avec une comparaison de deux bâtiments et (B) la maquette 3D fermée avec la température d'une salle visualisée

l'entrée du laboratoire afin de pouvoir les comparer. Pour ce faire, nous avons utilisé une maquette 3D fermée comme support de communication. Les données affichées sur cette maquette permettent d'obtenir en détail la consommation énergétique de chaque équipe. Les données sont donc affichées directement sur la maquette. Chaque façade est utilisée pour afficher l'électricité ou encore les flux de chaleur des salles utilisées par les chercheurs.

La maquette 2,5D (cf. Figure 3.9 A) a été construite à l'aide de 139 briques Lego : 45 briques blanches représentent les bâtiments (laboratoires, salles de classe, administration) ; 55 briques grises représentent les différentes allées et routes, 20 briques bleues permettent de représenter une rivière bordant le campus et 14 briques vertes représentent les différents espaces verts et terrain de sport. Nous avons utilisé 6 plaques de lego (16x16) pour représenter le sol. Toutes les briques mesurent une unité de hauteur (0,96 cm). Modéliser le modèle numérique a pris 18 minutes, la maquette physique 36 minutes et la liaison entre le modèle numérique et physique a pris 30 minutes.

La maquette 3D fermée (cf. Figure 3.9 B) représente notre laboratoire d'informatique en un bâtiment carré de 3 étages. 130 briques composent la maquette. Chaque façade est composée de 3x3 murs blancs. Une ligne de 12 pièces blanches permet de séparer les étages. 66 briques gris foncé modélisent le toit et 10 pièces plates blanches permettent de le solidifier. La modélisation numérique a été réalisée en 24 minutes, la construction physique en 25 et la liaison dans Unity en 16 minutes.

2.6.4 Support à la gestion patrimoniale

Ce scénario a été motivé à la suite des entretiens menés en présence des architectes (Scénario 2 : Maquette interactive pour l'affichage de données architecturales). Nous avons conçu un

scénario permettant de visualiser les coûts des différentes opérations de maintenance du bâtiment au cours du temps. En utilisant un modèle 3D ouvert (cf. Figure 3.10), nous représentons l'évolution des différents matériaux à la fois à l'intérieur mais aussi à l'extérieur du bâtiment. La maquette 3D ouverte a été réalisée avec 200 briques Lego. 38 murs permettent de définir les étages et les salles, les 162 autres permettant de représenter l'équipement à l'intérieur des salles. La modélisation en Lego Digital Designer a pris 24 minutes, la construction 28 minutes et la liaison dans Unity 18 minutes.

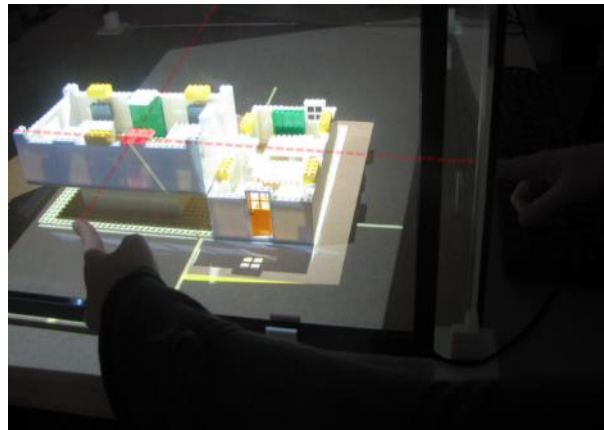


Figure 3.10 – La maquette 3D ouverte avec un dispositif lumineux sélectionné à l'aide de deux lancers de rayon

2.7 Synthèse

Nous avons présenté dans cette section une nouvelle approche modulaire et interactive permettant l'interaction avec un modèle physique 3D supportant la visualisation et la manipulation de données situées. Chaque composant d'un modèle 3D physique peut se transformer en vecteur d'informations numériques. Les informations sont rendues directement à travers la maquette. Nous avons implémenté un Unity Asset Package pour générer rapidement un modèle numérique à partir d'une maquette physique et des scripts permettant de lier les deux espaces. Cette approche a été mise en œuvre et illustrée sur trois applications implémentant différents scénarios émergents d'entretiens.

Dans ces applications, les utilisateurs pouvaient sélectionner une ou plusieurs pièces de la maquette. Cette sélection déclenchait alors une visualisation de données liées à la brique Lego. La technique de sélection se base sur un lancer de rayon à partir de la vitrine protectrice de la maquette et était inspirée de différentes techniques de lancers de rayons vues dans l'État de l'art. Cette technique fait l'objet de plusieurs études utilisateur que l'on explicitera dans le chapitre suivant (cf. chapitre TouchGlass).

Cet outil logiciel permet donc de créer un ensemble de systèmes interactifs de visualisations de données situées sur des maquettes physiques. Cependant, certains choix de conception de notre outil de visualisation de données situées (technologie utilisée, technique d'interaction, type de visualisation, etc.) peuvent être différents pour d'autres systèmes de visualisation de données situées. Dans la section suivante, nous proposons un outil conceptuel pour appréhender les différentes propriétés des systèmes de visualisation de données situées.

3 Outil Conceptuel

Nous avons vu précédemment dans l'état de l'art que des systèmes de visualisation de données situées peuvent prendre plusieurs formes aussi bien pour les dispositifs en entrée que pour les dispositifs d'affichage. Afin de mieux comprendre la diversité des systèmes de visualisation de données situées, nous proposons un outil conceptuel support au raisonnement et à la conception de l'interaction avec un référent physique. Cet outil conceptuel prend la forme d'un espace de conception. Cet espace permet de définir certaines propriétés des systèmes de visualisation de données situées, les illustrer et les comparer au travers d'exemples de la littérature. Nous avons structuré ce support conceptuel autour de cinq propriétés :

- **La localisation des données** caractérise le rendu des données situées (Willett et al., 2017) : les données sont soit situées proches du référent physique soit ancrées sur le référent physique.
- **La distance utilisateur-référent** se base sur les distances proxémiques (Ballendat et al., 2010; E. T. Hall, 1966) pour identifier l'impact de la distance entre l'utilisateur et le référent physique.
- **La superposition moteur-visuel** caractérise la continuité dans l'interaction entre l'utilisateur et le référent physique, autrement dit, l'utilisateur agit-il dans un espace qui se superpose à l'espace dans lequel sont affichées les données ?
- **La relation référent-médiateur** exprime le degré de physicalité de l'interaction en entrée entre un dispositif d'entrée, qu'on appelle le médiateur, et le référent physique.
- **La localisation du feedback** d'interaction caractérise le lieu où est rendu perceptible ce *feedback*.

Dans cette section nous commençons par présenter les cinq propriétés de notre outil. Nous illustrons ensuite les capacités illustratives puis comparatives de notre espace de conception.

3.1 L'espace de conception

Notre approche vise à étendre une partie d'un espace de conception existant, qui caractérise l'interaction en sortie pour des données situées. Pour cela, nous nous appuyons sur les travaux de Willett et collab. (Willett et al., 2017) et les étendons en introduisant quatre propriétés complémentaires destinées à caractériser des techniques d'interaction en entrée en présence d'un référent physique porteur de données situées.

3.1.1 Composantes de notre espace de conception

Willett et collab. (Willett et al., 2017), décrivent les données situées en fonction de la localisation des données par rapport à leur référent physique. Si les données sont affichées dans l'environnement du référent, elles sont **situées proches** du référent. Si les données sont affichées sur le référent, elles sont **ancrées** sur le référent. Cependant, cet espace ne décrit pas l'interaction en entrée avec des systèmes proposant des visualisations de données situées.

Nous considérons que l'interaction en entrée avec un système de visualisation de données situées est constituée de 4 composantes : l'utilisateur, le médiateur (p. ex. le type de dispositif d'interaction utilisé), le référent physique et le *feedback* d'interaction. Ces dimensions sont explicitées puis nous présentons les différentes valeurs qui les composent.

3.1.2 Distance utilisateur-référent

Edward Hall (E. T. Hall, 1966) définit la proxémie chez l'humain comme comportant 4 distances distinctes où les rapports sociaux et sensoriels évoluent : la distance intime (inférieur à 45 cm), la distance personnelle (entre 40 et 125cm), la distance sociale (entre 120 et 360 cm) et la distance publique (au-delà de 360 cm). Cette approche basée sur les relations humain-humain a été reprise par Ballendat et collab. (Ballendat et al., 2010) pour concevoir des interactions différentes selon la distance et l'orientation de l'utilisateur par rapport à un écran dans un espace public.

Nous nous réapproprions cette notion dans notre contexte pour l'appliquer à la distance utilisateur-référent. En effet, les objets physiques peuvent être placés à des distances différentes soit à cause de leur taille soit parce qu'ils sont destinés à être examinés différemment (p. ex. voir les détails d'une maquette d'un bâtiment est différent de voir la globalité d'un squelette de dinosaure). Concevoir des interactions avec un référent physique porteur de données situées nécessite donc d'appréhender cette distance pour pouvoir tirer profit au mieux de l'environnement du référent physique. Nous définissons les distances entre l'utilisateur et le

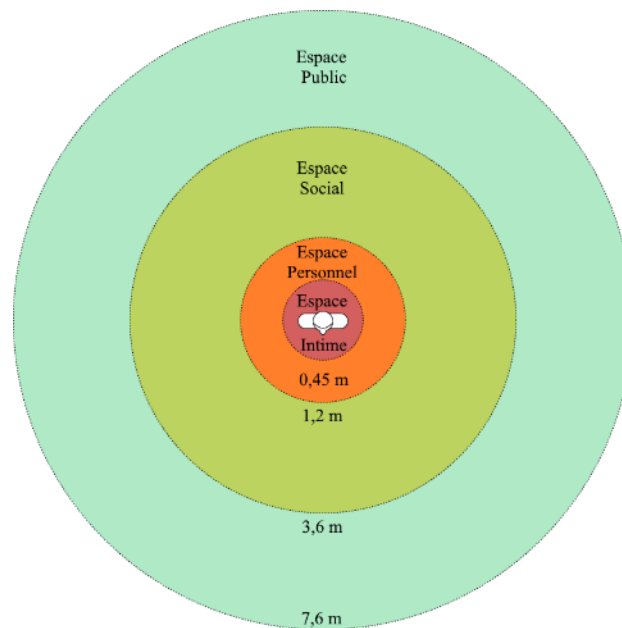


Figure 3.11 – Les quatre distances proposées par Hall (E. T. Hall, 1966). Image récupérée de la page Wikipédia - Proxemics

réfèrent physique selon les quatre distances de Hall : **intime**, **personnelle**, **publique** et **sociale** et leur équivalent en mètres (cf. Figure 3.11 et en vert sur la Figure 3.14).

3.1.3 *Superposition espace moteur-visuel*

Selon la définition faite par Casiez (Casiez, 2012), la superposition de l'espace moteur-visuel est la composante principale de l'interaction directe. Dans notre espace de conception, l'espace moteur-visuel est dit **superposé** si les quatre composantes de l'interaction en entrée impliquées avec ce système (utilisateur, médiateur, *feedback* et réfèrent) se trouvent dans l'alignement visuel utilisateur-réfèrent physique. Cependant, maintenir l'alignement de l'espace moteur-visuel peut causer des problèmes de type « *fat finger problem* » (Siek et al., 2005). Ce problème connu des interfaces tactiles entraîne de l'occultation entre l'utilisateur et sa zone d'intérêt. Ce problème se pose aussi avec un objet physique. Si l'utilisateur touche la maquette, il occulte la zone qu'il vient de toucher (cf. chapitre État de l'art section 2.1). Pour contourner ce problème, il est donc parfois nécessaire de séparer ces composantes. Par exemple, utiliser un menu sur un smartphone permet de sélectionner une partie du réfèrent physique sans occulter la visualisation de données ou du réfèrent physique. L'alignement n'est donc pas préservé, il y a **dispersion** de l'espace moteur-visuel. Nous précisons qu'une représentation numérique du réfèrent physique (p. ex. une vidéo à travers un smartphone) provoque l'occultation du réfèrent physique et ne préserve donc pas l'alignement.

Nous identifions donc 2 valeurs possibles pour cette dimension : **superposition** et **dispersion** de l'espace moteur-visuel (cf. Figure 3.14 en orange).

3.1.4 Relation référent-médiateur

L'interaction avec un système de visualisation de données situées peut se faire sur les données ou sur le référent physique. Nous avons vu dans l'état de l'art plusieurs techniques de sélection pour interagir avec ces systèmes (au contact ou à distance, directe ou indirecte, avec un support ou sans support de contact). Nous proposons avec cette dimension de classifier le système interactif en fonction du dispositif interactif qu'il utilise. Dans le contexte d'interaction avec un référent physique, nous appelons ce dispositif le médiateur car il permet de faire la passerelle entre le monde numérique et le monde physique.

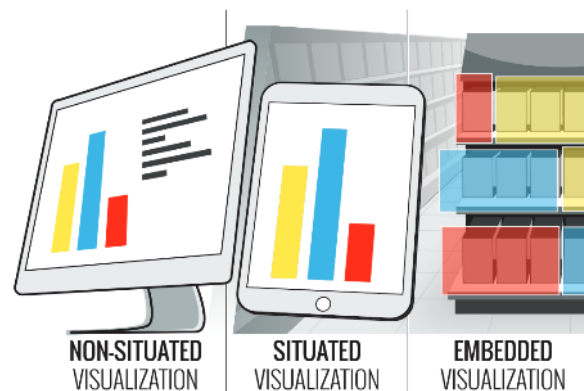


Figure 3.12 – Exemple du continuum de visualisation de données situées (Figure 10 (Willett et al., 2017))

À l'instar du continuum de données situées proposées par Willett et collab. (Willett et al., 2017), cette propriété permet de décrire l'ancrage spatial dans l'espace du référent physique du dispositif d'interaction en entrée. Ce continuum propose deux valeurs (**situées** et **ancrées**) auxquelles s'ajoutent les données non situées non présentes dans notre espace (c'est-à-dire affichées dans un contexte différent du référent physique). Autrement dit, la visualisation de données se caractérise par rapport à son ancrage par rapport au référent. Ce continuum est résumé en Figure 3.12 dans le cas d'une visualisation d'un stock de produits dans un magasin. La propriété de la relation référent-médiateur propose de décrire la technique d'interaction en entrée par rapport au référent physique. La propriété est illustrée en Figure 3.13 et décrite selon 4 niveaux :

Niveau 1 – **virtuel pur** : La position et l'orientation du médiateur par rapport au référent physique n'influencent pas l'interaction. Par exemple, l'appui sur une touche du clavier ou le déplacement de la souris ne dépend pas de l'orientation et la position du dispositif par rapport au référent.

Niveau 2 – **extension virtuelle** : le médiateur offre un support d'interaction en entrée sur une représentation numérique (p. ex. l'utilisateur touche une représentation numérique d'une

maquette en réalité augmentée rendue par la caméra d'un smartphone). L'interaction passe par une représentation numérique du référent physique, typiquement son image, un plan, une vidéo en direct. Cette représentation dépend de l'orientation et la position du dispositif par rapport au référent.

Niveau 3 – **extension physique** : le médiateur permet d'étendre le corps de l'utilisateur. L'orientation et la position du médiateur par rapport au référent définissent l'interaction. Par exemple, pour un contrôleur tenu en main et permettant de lancer un rayon, le système interprète la position et l'orientation d'une partie du corps de l'utilisateur pour interagir avec le référent physique.

Niveau 4 - **physique pure (ou pas de médiateur)** : il n'y a pas de relation. L'interaction entre l'utilisateur et le référent ne s'appuie sur aucun intermédiaire : seul le corps de l'utilisateur (membres, postures, etc.) est utilisé.

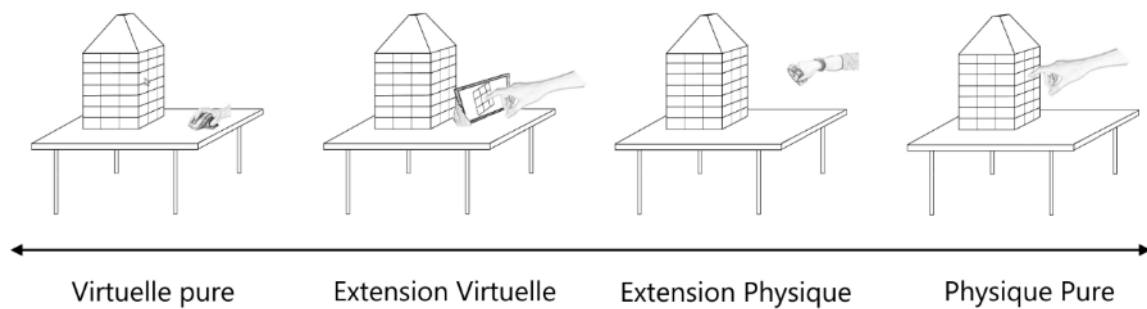


Figure 3.13 – Continuum de la relation référent-médiateur

Nous avons donc 4 niveaux de relation : **virtuelle pure, extension virtuelle, extension physique, physique pure** (cf. Figure 3.13 et en bleu sur la Figure 3.14).

3.1.5 Localisation du feedback

Le *feedback* d'interaction est important car il permet à l'utilisateur d'être guidé durant sa tâche en lui indiquant les étapes nécessaires pour la compléter (p. ex. la position d'un pointeur de souris pour une tâche de pointage) ou quand la tâche est terminée (p. ex. surbrillance d'un clic sur un bouton). Cependant, certaines solutions n'apportent pas de *feedback* pour aider l'utilisateur durant son interaction (p. ex. un toucher sur l'icône d'une application sur smartphone). Quand un *feedback* d'aide à l'interaction est apporté à l'utilisateur, il est important de le caractériser. En nous inspirant de l'approche de Willett et collab. (Willett et al., 2017), nous caractérisons le *feedback* d'aide à l'interaction selon sa localisation. Nous retenons trois localisations possibles pour le *feedback* d'aide à l'interaction (cf. Figure 3.14 (en violet)) : sur

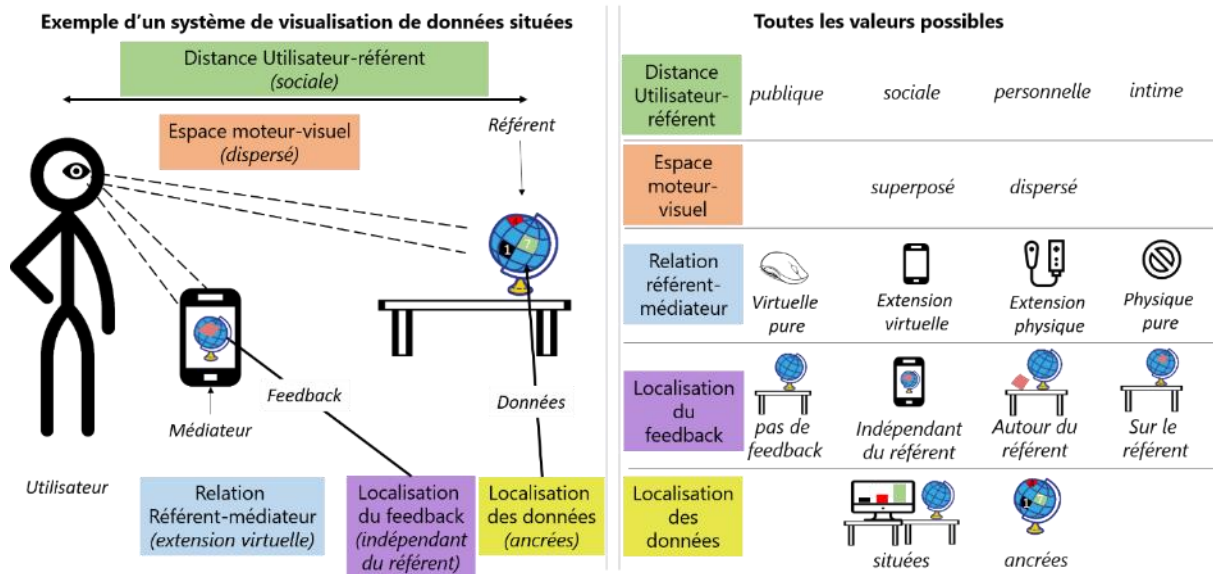


Figure 3.14 – Illustration (gauche) d'un exemple de système interactif de visualisation de données situées à une distance **sociale** avec un espace moteur-visuel **dispersé**, un médiateur ayant une relation de type **extension virtuelle**, un retour visuel affiché **indépendamment** du référent physique et des données **ancrées** et (droite) les propriétés de notre outil conceptuel

le référent physique (p. ex. un pointeur de souris sur le référent physique), autour du référent physique (p. ex. projection du pointeur à côté du référent physique) ou indépendamment du référent physique (p. ex. pointeur sur l'écran d'un smartphone).

3.2 Illustration du pouvoir descriptif de notre espace

Afin d'illustrer le pouvoir descriptif de notre outil (Beaudouin-Lafon, 2004), nous avons recensé les techniques d'interaction avec un référent physique porteur de données dans les communications de la conférence CHI de l'année 2013 à l'année 2018. Nous avons étendu notre recherche aux références citées dans ces articles. Au total, 49 techniques d'interaction ont été sélectionnées et classifiées. Nous les avons ensuite analysées selon les propriétés de notre outil conceptuel.

Dans cette section nous proposons une analyse synthétique de ces travaux pour étudier le pouvoir descriptif de notre outil. Nous sélectionnons un sous-ensemble de ces techniques pour illustrer chaque valeur des propriétés de notre outil conceptuel.

3.2.1 Distance utilisateur-référent

La plupart des interactions avec des systèmes de visualisation de données situées par vidéoprojection permettent de manipuler le référent à une distance intime. Par exemple dans Tangible Landscape, Millar et collab. (Millar et al., 2018) utilisent un relief topologique pour faire appréhender à des élèves des processus scientifiques tels que la géomorphologie ou l'hydrologie (cf. Figure 3.15 A). Le groupe d'étudiants peut venir sélectionner un point d'intérêt

ou modifier directement le relief et voir ainsi son évolution. Les données géomorphologiques ou hydrologiques se mettent à jour en temps réel. Cette distance est la plus représentée dans notre sélection avec 29 techniques (59%) qui utilisent un système interactif à une distance **intime**.

Certains objets physiques ne peuvent pas être appréhendés de très près. En effet, certains référents physiques doivent être protégés de tout contact avec les utilisateurs. Roberts et collab. (Roberts et al., 2018) ont étudié l'interaction avec une tablette fixée devant une vitrine de musée contenant un objet physique fragile. Les informations relatives aux référents physiques sont alors présentées sur l'écran de la tablette sous la forme de texte, de *timelines* ou de questions. L'utilisateur est à une distance **personnelle** (cf. Figure 3.15 B) tout comme 11 techniques (23%) de notre sélection.

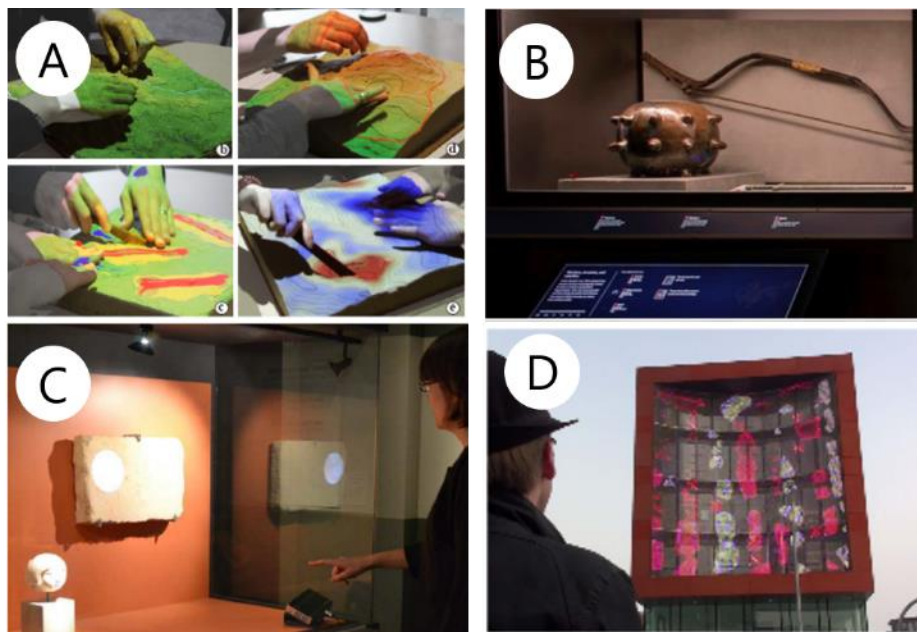


Figure 3.15 – Illustration de la distance utilisateur-référent avec (A) une distance **intime**, (B) une distance **personnelle**, (C) une distance **sociale** et (D) une distance **publique**

Avec *The Revealing Flashlight* (Ridel et al., 2014), Ridel et collab. proposent d'augmenter des artefacts archéologiques en leur projetant une visualisation. L'utilisateur choisit, en pointant du doigt ou avec un contrôleur, la partie du référent physique dont il souhaite voir les données. À la manière d'une lampe torche, l'utilisateur éclaire le référent physique. Les données archéologiques généralement invisibles car masquées par le temps sont mises en évidence par vidéoprojection sur l'artefact archéologique. Dans ce projet, l'utilisateur se trouve donc à une distance **sociale** (cf. Figure 3.15 C). Dans notre sélection, 4 techniques (8%) étudient cette distance.

Enfin, visualiser et interagir avec des données projetées sur des bâtiments nécessite d'avoir une vision globale et donc de se positionner à une distance **publique**. Böhmer et collab. (Böhmer et al., 2011) utilisent le bâtiment comme plateforme de jeu pour un utilisateur se situant de l'autre côté de la rue (cf. Figure 3.15 D). La forme du bâtiment délimite la zone de jeu et de projection d'informations. 5 techniques (10%) étudient cette distance dans notre sélection.

3.2.2 *Superposition de l'espace moteur-visuel*

Inner Garden développé par Roo et collab. (Roo et al., 2017) projette sur un relief construit dans un bac à sable des informations topologiques. L'utilisateur manipule le sable et les données topologiques s'adaptent en conséquence. Si le sable forme un relief, une montagne sera projetée, sinon une couleur bleue sera projetée pour représenter la mer. L'espace moteur-visuel est donc **superposé** (cf. Figure 3.16 A). Dans notre sélection, 32 techniques (65%) superposent l'espace moteur-visuel. Au contraire, 17 projets (35%) dispersent l'espace moteur-visuel.

Par exemple, dans une version de leur prototype Leithinger et collab. (Leithinger et al., 2013) proposent l'utilisation d'une tablette pour afficher le *feedback* d'aide à l'interaction. Leur système permet de visualiser des données géospatiales relatives à la forme de leur écran déformable, l'alignement visuel est alors rompu par la tablette. L'espace moteur-visuel est **dispersé** (cf. Figure 3.16 B).



Figure 3.16 – Illustration de la superposition de l'espace moteur-visuel avec (A) un espace moteur-visuel **superposé** et (B) un espace moteur-visuel **dispersé**

3.2.3 *Relation référent-médiateur*

Les différents niveaux de relation permettent de définir le caractère physique de l'interaction (du virtuelle pure au physique pure). Dans le projet Augmented Studio (Hoang et al., 2017), Hoang et collab. utilisent un clavier et une souris pour venir interagir avec des données morphologiques (p. ex. os, muscles, tendons, etc.) directement projetées sur le corps d'un être humain ; la relation est donc **virtuelle pure** (cf. Figure 3.17 A). 9 techniques (18%) proposent ce type de relation.

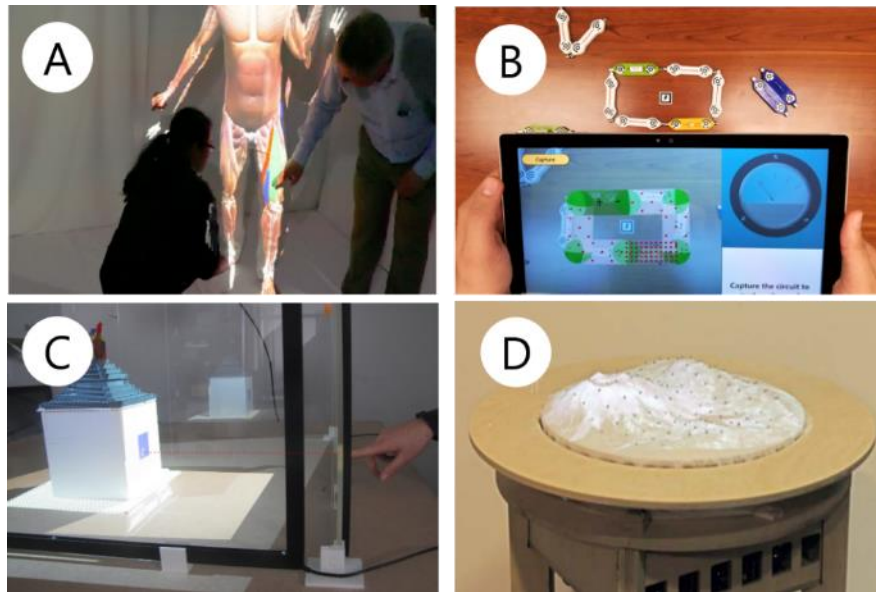


Figure 3.17 – Illustration de la relation référent-médiateur avec (A) une relation **physique pure**, (B) une **extension virtuelle**, (C) une **extension physique** et (D) une relation **virtuelle pure**.

Beheshti et collab. (Beheshti et al., 2017) ont développé le système Looking Inside The Wires utilisant de la réalité augmentée au travers d'une tablette pour faire apprendre les rudiments des circuits électriques. Les données électriques (p. ex. mouvement des électrons) sont représentées sur le modèle numérique à l'aide de la tablette et l'utilisateur peut interagir en touchant l'écran de la tablette. La relation est donc une **extension virtuelle** (cf. Figure 3.17 B). Seulement, 4 techniques (8%) utilisant ce type de relation ont été proposées.

Dans le chapitre suivant, nous proposons d'utiliser une vitrine tactile protégeant les objets physiques pour interagir avec ces objets (Cabric et al., 2019) pour, à terme, visualiser des données relatives aux référents (p. ex. âge d'une roche volcanique, lieu de découverte d'un fossile, température d'une salle, etc.). Cette vitrine ne permet pas d'afficher une représentation numérique du référent physique. L'interaction se fait par projection du doigt par lancer de rayon sur le référent. La relation est donc une **extension physique** (cf. Figure 3.17 C). L'utilisation d'un médiateur pour étendre physiquement le référent est présente dans 19 techniques de notre sélection (39%).

Enfin, le corps est utilisé comme dispositif d'interaction en entrée dans 17 techniques (35%). Par exemple, Leithinger et Ihsii (Leithinger & Ishii, 2010) demandent à l'utilisateur de modifier un relief 3D en le touchant. Le relief est représenté à l'aide d'un écran déformable recouvert d'un lycra permettant la projection de données topologiques. Le corps humain est alors le dispositif d'interaction et il n'y a pas de médiateur. La relation est d'ordre **physique pure** (cf. Figure 3.17 D).

3.2.4 Localisation du feedback

Les différentes localisations du *feedback* permettent de définir la proximité entre le référent physique et le rendu permettant d'aider l'utilisateur à interagir. Kim et collab. (H.-J. Kim et al., 2016), permettent avec miniStudio de prototyper des espaces ubiquitaires en utilisant des maquettes augmentées en papier. Chaque partie de l'environnement (p. ex. mur de données, personnes, écrans) est modélisée puis construite. Le déplacement de chaque objet permet de manipuler directement les données projetées sur la maquette (p. ex. sur la Figure 3.18 A des données démographiques liées à une carte). Il n'y a donc **pas de feedback** aidant l'interaction. Dans notre sélection, 21 techniques (43%) ne proposent aucun *feedback*.

ElSayed et collab. (ElSayed et al., 2015) étudient l'interaction avec des produits dans un magasin avec une tablette en réalité augmentée. La tablette permet ici d'afficher les données propres à chaque produit (p. ex. prix, composition, réduction, etc.). Le ou les produits sélectionnés sont entourés par un cercle de couleur sur la tablette. Le *feedback* est donc **indépendant du référent** (cf. Figure 3.18 B). 11 techniques de notre sélection d'articles (22%) proposent ce type de *feedback*.

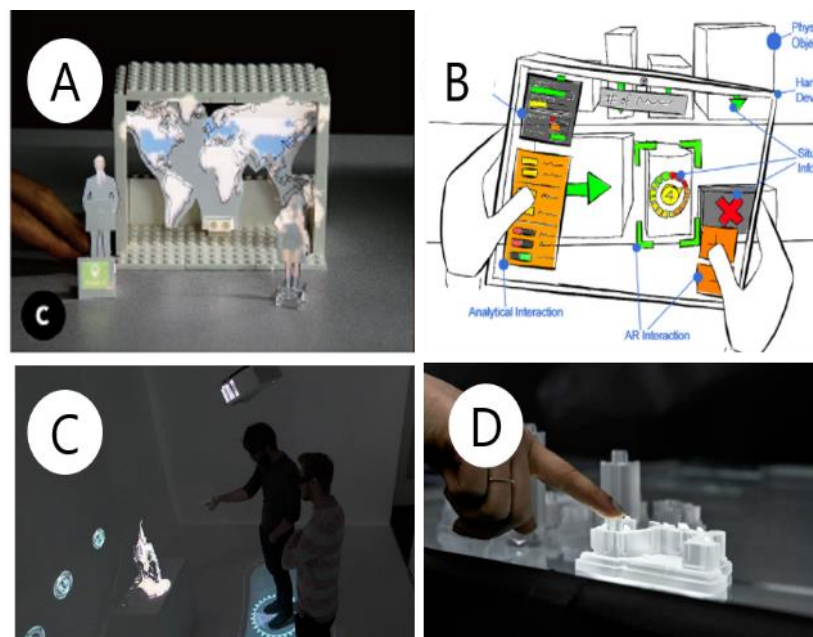


Figure 3.18 – Illustration de la propriété de localisation du feedback avec (A) **pas** de feedback, (B) un feedback **indépendant** du référent physique, (C) un feedback **autour** du référent physique et (D) un feedback **sur le référent physique**

Schmidt et collab. (S. Schmidt et al., 2018) visualisent des données sur une maquette physique de dinosaure. La visualisation est déclenchée quand l'utilisateur se trouve à une position précise autour de la maquette. Cette position est inscrite au sol par des cercles vidéoprojetés. Ce *feedback* se trouve **autour du référent** (cf. Figure 3.18 C). 12 techniques (25%) affichent un *feedback* autour du référent physique.

Au contraire, Walker et collab. (Walker et al., 2017) utilisent une carte tangible pour évaluer la relation spatiale avec des données provenant de bâtiments représentés physiquement. Afin de guider l'utilisateur et voir quel bâtiment est sélectionné, un *feedback* lumineux apparaît sur le bâtiment et donc **sur le référent** physique (cf. Figure 3.18 D). Seules 5 techniques (10%) utilisent un *feedback* sur le référent physique.

Dans cette section, nous avons illustré chacune des propriétés et leurs différentes valeurs au travers d'exemple de la littérature. Notre espace de conception semble donc être capable d'illustrer les systèmes de visualisation de données situées interactifs en fonction de leurs interactions avec le référent. Dans la section suivante, nous présentons notre espace de conception comme un outil de comparaison de systèmes selon les propriétés présentées dans cette section (cf. Figure 3.19).

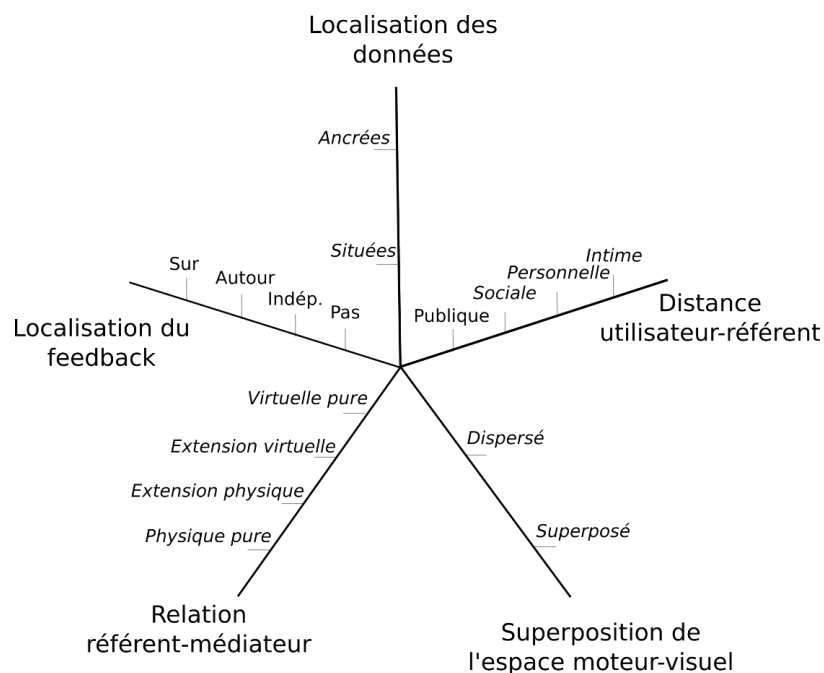


Figure 3.19 – Les cinq propriétés de notre outil conceptuel

3.3 Illustration du pouvoir comparatif de notre espace

Dans cette section, nous illustrons à partir d'exemples de la littérature le pouvoir comparatif de notre espace de conception (Beaudouin-Lafon, 2004) en comparant des systèmes deux à deux. Ces comparaisons de systèmes permettent d'analyser et de mieux comprendre les choix de conception des différents systèmes présents dans notre sélection d'articles. Nous montrons au travers de deux comparaisons, comment certains systèmes issus d'un même contexte d'utilisation et qui paraissent proches peuvent se différencier en matière d'interaction en entrée.

Ensuite, nous comparons deux systèmes qui ont été développés dans des buts et des contextes d'utilisation différents. Cependant, nous montrons au travers de notre espace de conception que ces deux systèmes qui semblent éloignés sont finalement proches en matière d'interaction. Nous finissons par comparer des systèmes qui se différencient uniquement selon la localisation des données afin de mettre en perspective l'outil conceptuel de l'interaction en entrée que nous avons défini et l'espace de conception d'interaction en sortie proposé par Willett et collab. (Willett et al., 2017).

3.3.1 Interagir avec des bâtiments

Le domaine du Media Architecture étudie les bâtiments comme support à l'interaction (en entrée ou en sortie). Les recherches dans ce domaine ont pu mettre en lumière un ensemble d'interactions avec des bâtiments physiques permettant de visualiser des informations sur leurs façades. Il s'agit donc de **données ancrées**, affichées à une **distance publique**. Néanmoins et malgré des contextes très proches, les interactions avec des bâtiments peuvent être différentes.

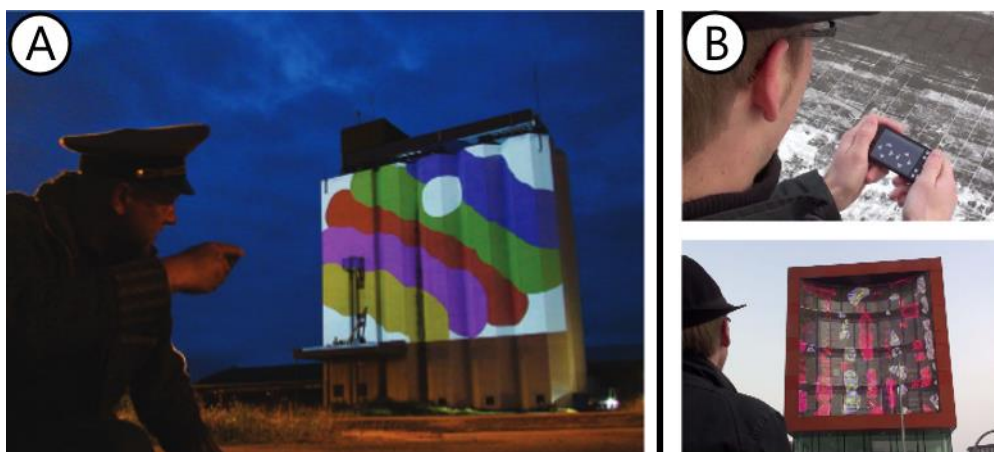


Figure 3.20 – Les systèmes interactifs sur un bâtiment avec (A) MobiSpray (Scheible & Ojala, 2009) et (B) les travaux de Böhmer (Böhmer et al., 2011)

Les projets d'interaction avec des données ancrées sur des bâtiments physiques permettent de sensibiliser les passants ou peuvent encore être utilisés comme œuvres d'art. Ces interactions se déroulent à des distances publiques permettant de voir l'entièreté du bâtiment. Dans le projet MobiSpray (Scheible & Ojala, 2009), Scheible et collab. permettent à un ou plusieurs utilisateurs de dessiner des informations (traits de couleurs) sur un bâtiment. De manière similaire, Böhmer et collab. (Böhmer et al., 2011) permettent de jouer avec des données projetées sur un bâtiment physique. Dans ces deux projets, l'interaction se fait à l'aide d'un téléphone. Dans MobiSpray, ce téléphone est utilisé comme un pointeur laser physique (cf. Figure 3.20 A). La relation est une **extension** physique. En revanche, Böhmer et collab. (Böhmer et al., 2011) utilisent le tactile du smartphone pour appliquer des commandes dans un

menu. Cette relation est donc **virtuelle pure** (cf. Figure 3.20 B). L'utilisateur peut s'affranchir du contexte du référent physique pour interagir.

Bien que les données soient ancrées sur le bâtiment, **la localisation du feedback** est différente. Dans le projet MobiSpray (Scheible & Ojala, 2009) les concepteurs ont fait le choix de ne pas utiliser de *feedback* d'aide à l'interaction. Le résultat de l'interaction est directement affiché sur le référent physique. Au contraire, Böhmer et collab. (Böhmer et al., 2011) utilisent le smartphone pour venir proposer un *feedback* **indépendant** du référent physique à l'utilisateur. Positionner ce *feedback* sur le smartphone permet d'avoir un meilleur contrôle de l'action mais entraîne une **dispersion** de l'espace moteur-visuel, au contraire de MobiSpray qui **superpose** les espaces de l'utilisateur, du médiateur et du référent (cf. Figure 3.21).

Malgré des contextes d'utilisation très proches, les deux projets utilisent donc des solutions d'interaction en entrée différentes selon trois propriétés.

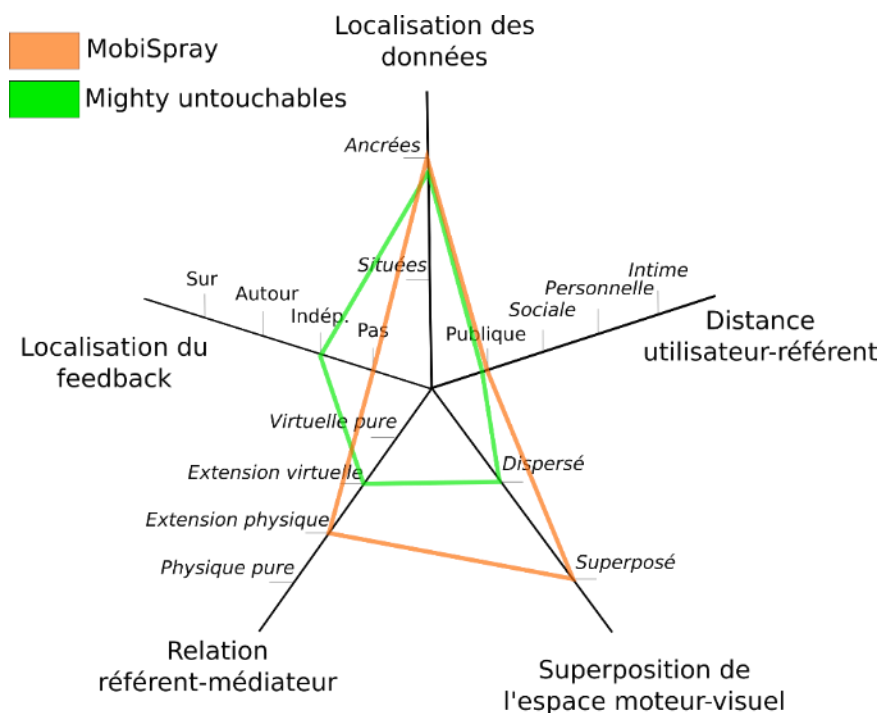


Figure 3.21 – Comparaison de systèmes interactifs sur un bâtiment avec (orange) MobiSpray (Scheible & Ojala, 2009) et (vert) les travaux de Böhmer (Böhmer et al., 2011)

3.3.2 Interagir avec des petits objets physiques

Interagir avec des petits objets nécessite d'avoir une interaction précise. Malgré cette contrainte, les techniques d'interaction conçues pour ces systèmes peuvent être différentes. Les projets TouchGlass (Cabric et al., 2019) et Point-and-Shake (Freeman et al., 2018) étudient une tâche de pointage se basant sur un lancer de rayon pour sélectionner des cibles physiques. Dans le

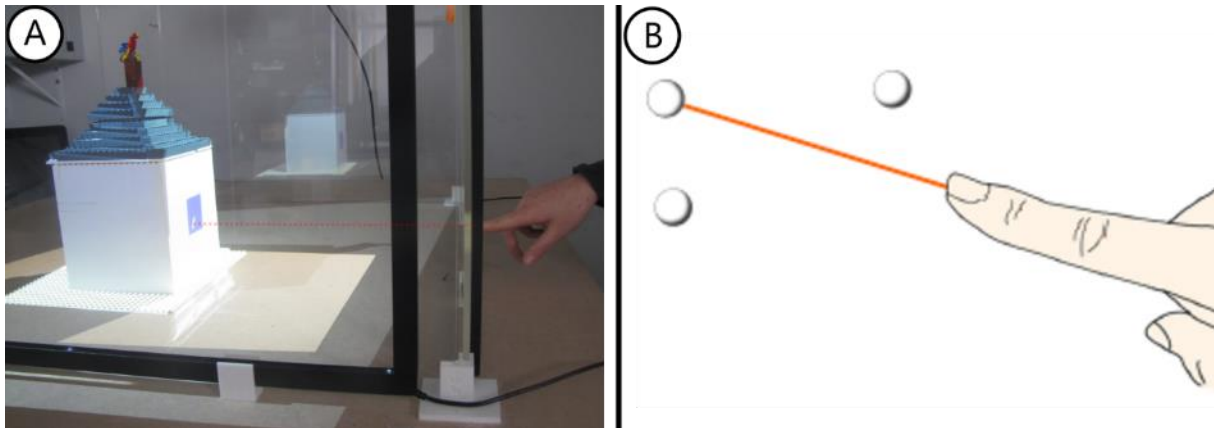


Figure 3.22 – Les systèmes interactifs sur des petits objets (A) TouchGlass (Cabric et al., 2019) et (B) Point-and-shake (Freeman et al., 2018)

projet TouchGlass, les données sont **situées** à côté de la maquette contrairement à Point-and-Shake qui représente une donnée (la couleur ou la hauteur de la sphère) sur la cible physique.

Cependant, dans notre projet TouchGlass (Cabric et al., 2019), nous utilisons une contrainte environnementale comme support à l'interaction : une vitrine. La vitrine devenue médiateur va permettre une **extension physique** entre elle et le référent. Dans Point-and-Shake, l'utilisateur pointe des sphères en utilisant uniquement son doigt. L'orientation et l'inclinaison sont captées à l'aide d'un LeapMotion, il n'y a donc pas de médiateur (cf. Figure 3.22).

Ces deux systèmes ont aussi étudié la distance entre le référent et l'utilisateur. Dans TouchGlass, les cibles peuvent se trouver jusqu'à 55 centimètres de distance, l'interaction se déroule donc à une **distance personnelle**, contrairement à Point-and-Shake qui se limite à une distance de 4 centimètres (c'est-à-dire **intime**) entre la cible et l'utilisateur. Ces systèmes se

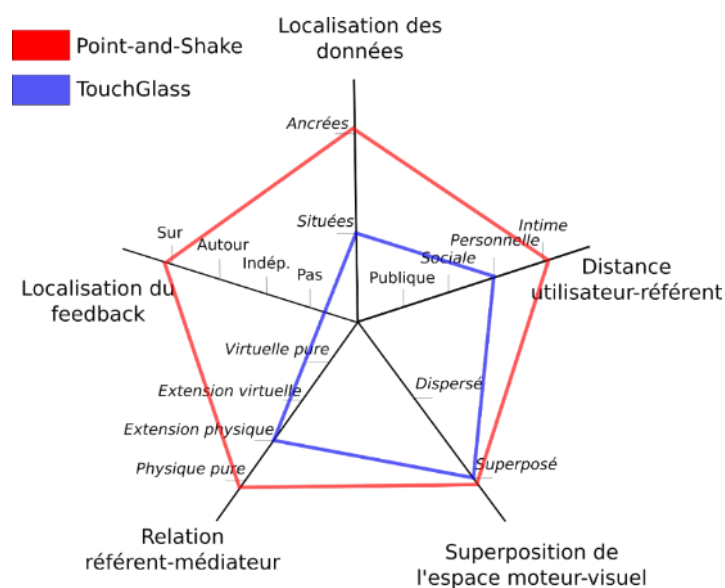


Figure 3.23 – Comparaison de systèmes interactifs sur des petits objets (A) TouchGlass (Cabric et al., 2019) et (B) Point-and-Shake (Freeman et al., 2018)

différencient aussi par le *feedback* apporté à l'utilisateur. TouchGlass s'intéresse à l'interaction **sans feedback** dans une tâche de pointage alors que Point-and-Shake donne une validation à l'utilisateur en faisant vibrer les sphères. Ce *feedback* se trouve **sur le référent**.

Finalement, en matière d'interaction, ces systèmes ne se ressemblent qu'au niveau de leur espace moteur-visuel. (cf. Figure 3.23). Malgré un besoin de précision dû à la petite taille des cibles, nous avons vu que les projets TouchGlass (Cabric et al., 2019) et Point-and-Shake (Freeman et al., 2018) utilisent des interactions différentes.

3.3.3 Interagir avec une projection sur un objet physique

Interagir avec un buste humain suivi en temps réel et interagir avec un relief topologique posé sur une surface plate semble être des projets nécessitant des techniques d'interaction différentes. Nous verrons dans cette section les points communs de ces projets en matière de techniques d'interaction.

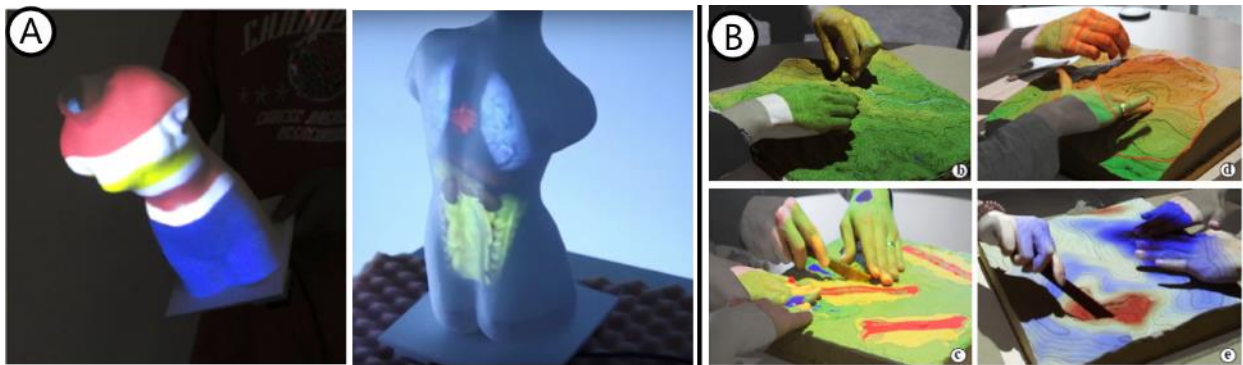


Figure 3.24 – Les systèmes de projection sur des objets physiques avec (A) Pmomo (Zhou et al., 2016) et (B) Tangible Landscape (Millar et al., 2018)

Le projet Pmomo (Zhou et al., 2016) permet d'interagir avec un objet physique pouvant se déplacer dans l'espace. L'objet est suivi dynamiquement et la projection des **données ancrées** s'adapte à la position et l'orientation (cf. Figure 3.24 A). Le référent physique est un prototype de buste humain interactif. L'utilisateur peut modifier le type de données à afficher en touchant directement l'objet (**physique pure**). L'utilisateur se trouve alors à une distance **intime**. **Aucun feedback** n'indique à l'utilisateur où il doit sélectionner pour modifier le type de données affichées. Enfin, l'espace d'interaction, le référent physique et les données numériques sont **superposés**.

Dans Tangible Landscape (Millar et al., 2018), Millar et collab. ont conçu un système permettant la découverte de phénomènes physiques sur un relief 3D. Les étudiants peuvent venir activer l'apparition de phénomènes naturels en modifiant le référent physique. La

visualisation de **données ancrées** sur le relief s'adapte en direct aux modifications topologiques (cf. Figure 3.24). L'utilisateur peut modifier le relief 3D à une distance **intime** en utilisant leurs mains (**relation physique pure**). La visualisation de données s'adapte en temps réel aux modifications **sans proposer de feedback**. Enfin, l'utilisateur interagit dans le même espace que les données numériques et le référent physique. L'espace moteur-visuel est donc **superposé**.

Malgré des contextes différents ces projets possèdent les mêmes caractéristiques d'interaction : **espace moteur-visuel superposé, distance intime, physique pure, absence de feedback** et les données sont **ancrées** (cf. Figure 3.25).

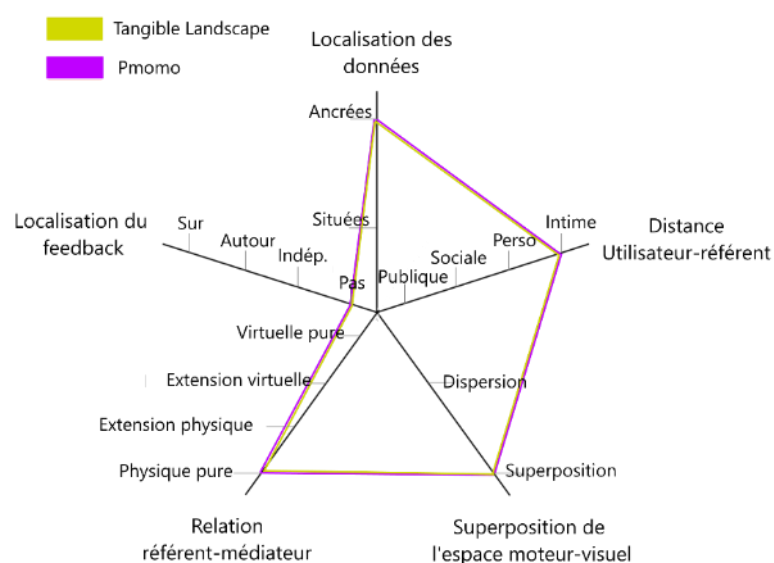


Figure 3.25 – Comparaison de systèmes de projection sur des objets avec (jaune) Tangible Landscape (Millar et al., 2018) et (rose) Pmomo (Zhou et al., 2016)

3.3.4 Exploration de données urbaines

L'exploration de données urbaines peut être utilisée pour simuler des évolutions démographiques (Grignard et al., 2018) ou sensibiliser des utilisateurs à la consommation énergétique des bâtiments (Alonso et al., 2018). Nous présentons dans cette section, deux projets qui sont proches pour l'interaction en entrée, mais qui sont différents pour la technique de visualisation utilisée.

La présentation de données situées sur des objets physiques est propice à l'utilisation de maquettes miniatures permettant de représenter des villes ou des quartiers (cf. chapitre État de l'art section 1.2). C'est le cas de CityScope (Alonso et al., 2018) et du projet Tangible Map (Walker et al., 2017) (cf. Figure 3.26).

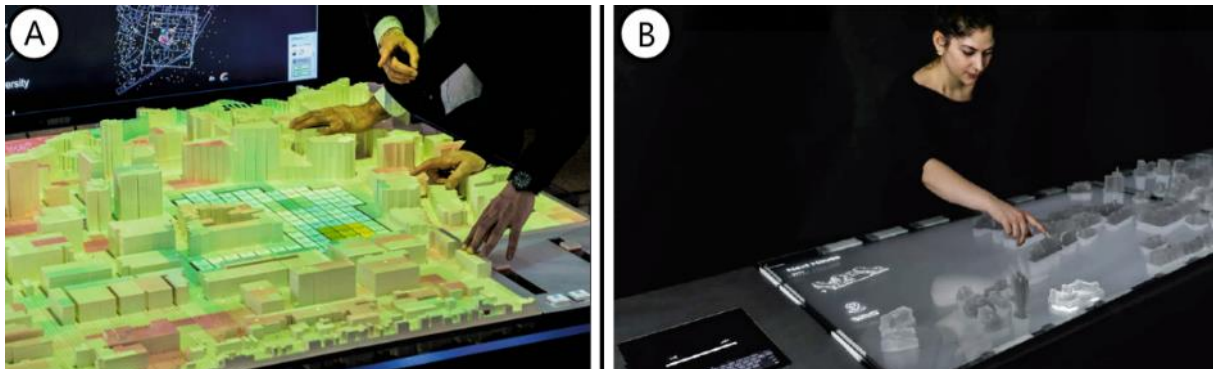


Figure 3.26 – Les systèmes d’exploration de données urbaines avec (A) CityScope (Alonso et al., 2018) et (B) TangibleMap (Walker et al., 2017)

Ces deux projets ont des caractéristiques d’interaction en entrée commune. L’interaction se fait au contact (**distance intime**), **sans médiateur**, avec un espace moteur-visuel **superposé** et en utilisant un **feedback sur le référent**. Cependant, ces systèmes ne présentent pas les données de la même manière. En effet, dans le projet Tangible Map, les données sont affichées à côté de la carte tangible. Elles sont donc situées **proches du référent** physique. Au contraire, le projet CityScope projette les données sur les éléments composant la maquette physique. Les données sont donc **ancrées sur le référent**.

La différence majeure est donc ici au niveau du rendu des données affichées (cf. Figure 3.27). Une visualisation ancrée sur le référent permet une meilleure mise en contexte, mais limite cependant la quantité d’informations qui peut être affichée (Willett et al., 2017). De plus, afficher les données sur la maquette entraîne de l’occultation quand l’utilisateur souhaite interagir avec la maquette ou les données (cf. chapitre État de l’art section 2.1)

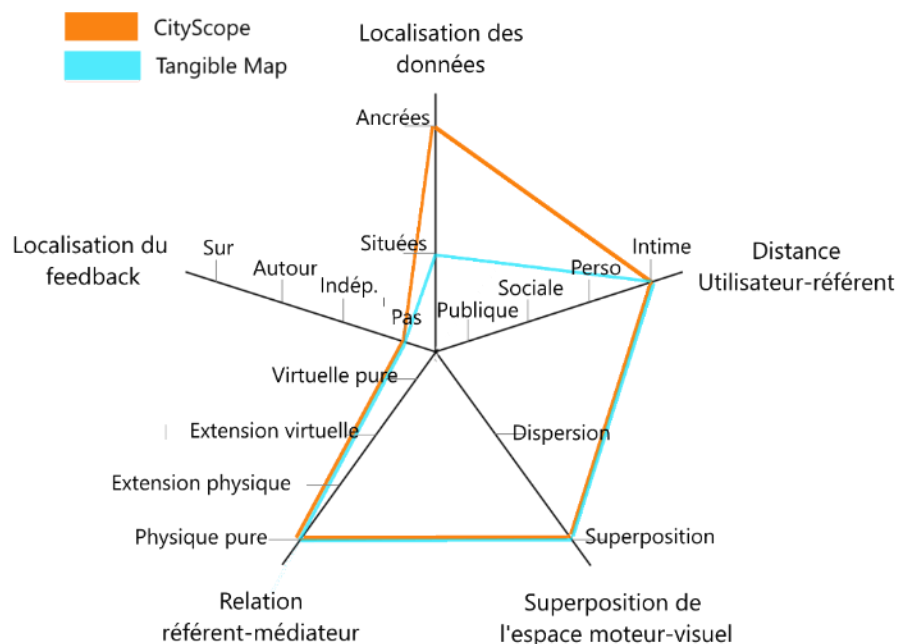


Figure 3.27 – Comparaison de systèmes d’exploration de données urbaines avec (en orange) le projet CityScope (Alonso et al., 2018) et (en bleu) le projet Tangible Map (Walker et al., 2017)

Cet espace de conception nous permet donc d'illustrer et de comparer plusieurs systèmes interactifs de visualisation de données situées où le référent physique est un objet à l'aide de cinq propriétés : la localisation des données, la distance utilisateur-référent, la superposition moteur-visuel, la relation médiateur-référent et la localisation du *feedback*. Ces cinq propriétés ont été choisies suite à l'analyse d'une série d'articles de la littérature. Nous pensons que cet espace de conception peut être une valeur ajoutée pour les concepteurs de systèmes interactifs de visualisation de données situées avec un objet physique, car il permet à la fois de se positionner et de tirer profit de travaux antérieurs potentiellement similaires.

4 Synthèse

Dans ce chapitre nous avons présenté le contexte dans lequel la thèse s'inscrit et les spécificités en matière de systèmes de visualisation de données situées. En effet, promouvoir le campus intelligent nécessite de prendre en compte plusieurs granularités à la fois au niveau des données (p. ex. les consommations énergétiques d'un campus d'un côté et de l'autre la température d'une seule salle) et au niveau des référents de ces données (p. ex. la maquette réduite d'un campus d'un côté et le capteur collectant la donnée de l'autre). Cette diversité a nécessité la création de deux outils : un logiciel et un conceptuel.

L'outil logiciel implémenté en Unity permet à des concepteurs de créer, à partir de briques Lego, un système de visualisation de données interactives avec plusieurs types de maquettes servant de référent physique. La création d'un package Unity permet au travers d'une interface graphique personnalisée incorporée à l'éditeur Unity, de faciliter la création du système.

Enfin, l'outil conceptuel prend la forme d'un espace de conception qui permet d'analyser les systèmes interactifs de données situées par le prisme de l'interaction en entrée. Il vient donc en complément du support d'analyse approfondi par Willet et collab. (Willett et al., 2017) qui détaillent ces types de systèmes du point de vue de l'interaction en sortie. Cet outil conceptuel prend en compte quatre dimensions de l'interaction en entrée : la distance utilisateur-référent, la superposition moteur-visuel, la relation médiateur-référent et la localisation du *feedback*. Au travers d'une analyse de la littérature, nous montrons que cet espace peut être utilisé pour décrire des systèmes mais aussi les comparer. Nous pensons que cet outil peut être un complément intéressant pour tous les concepteurs de systèmes permettant de s'inspirer de travaux précédents ou de faciliter la conception d'un système de visualisation de données situées. Nous avons montré, au travers de plusieurs exemples, le pouvoir descriptif et

comparatif de notre outil conceptuel. Il est commun d'évaluer un tel espace selon son pouvoir descriptif, comparatif et génératif. D'un point de vue génératif, cet espace permet de s'appuyer sur les propriétés présentées pour concevoir de nouveaux systèmes interactifs de visualisations de données situées. Ce pouvoir génératif devra être étudié prochainement.

Chapitre 4 TouchGlass

Contenu du Chapitre

1 Conception d'une technique de sélection depuis une vitrine.....	82
1.1 Problèmes et contraintes liés au lancer de rayon depuis une vitrine comme technique de sélection à distance.....	83
1.2 Facteurs influençant les techniques de sélection absolues pour la sélection d'objets distants avec une surface tactile comme support de contact.....	85
1.3 Synthèse	87
2 Étude informative	87
2.1 Protocole expérimental	88
2.2 Résultats	89
2.3 Conclusion.....	90
3 Étude de la sélection : conception	90
3.1 Tâche et instructions.....	91
3.2 Conditions	92
3.3 Participants	93
3.4 Conception et procédure.....	93
3.5 Matériel	93
3.6 Données collectées et traitement des données	94
4 Résultats de l'étude de sélection	95
4.1 Taux de réussite de la sélection de cible.....	95
4.2 Temps de réalisation des essais	96
4.3 Sélection au premier contact : centroïde, offset et spread	97
5 Discussion	100
5.1 Recommandations de conception	101
5.2 Perspectives	102
6 Synthèse.....	102

La visualisation de données situées requiert que celles-ci soient affichées proches voire sur leur référent physique. Ce référent peut être un lieu ou un objet physique. Ces objets physiques tels que les objets d'histoire naturelle, les œuvres d'art ou les maquettes architecturales sont souvent placés derrière une vitrine pour les protéger de la poussière ou du vandalisme (Mäkelä et al., 2017). La plupart de ces référents ne sont pas interactifs, et le niveau d'information qui peut être affiché est donc généralement limité à de petites notes textuelles (cf. Figure 4.1). Les rares interactions dans ce contexte sont souvent basées sur une entrée tactile sur un écran séparé (Roberts et al., 2018), ce qui éloigne l'utilisateur de l'objet exposé.

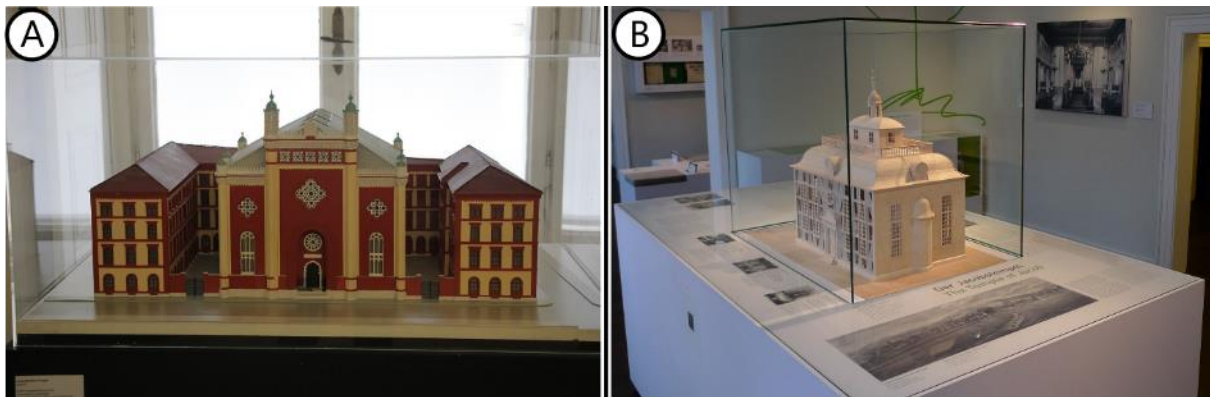


Figure 4.1 – Exemples d'objets physiques protégés par une vitrine avec des données textuelles : (A) la synagogue de Vienne-Leopoldstadt au musée juif autrichien d'Eisenstadt et (B) la synagogue de Seesen au Städtischen Museum de Seesen (images récupérées de Wikimedia)

Bien que ces objets soient placés derrière une vitrine, peu de projets ont exploré la manière d'utiliser ces vitrines pour interagir avec les objets qui s'y trouvent. Les solutions précédentes incluent l'utilisation de la réalité augmentée par le biais de smartphones ou de tablettes (Spindler & Dachsel, 2009). Ces solutions nécessitent alors que l'utilisateur installe une application dédiée. Des recherches antérieures ont proposé une variété d'approches ne nécessitant pas l'installation d'une application dédiée pour promouvoir une expérience plus engageante avec les objets exposés. Par exemple, Gehring et collab. (Gehring et al., 2012) ont proposé d'utiliser des gestes en l'air, détectés avec un Kinect, pour pointer un objet depuis une vitrine. La métaphore de la lampe de poche a été étudiée dans des expositions publiques (J. Green et al., 2014; Ridel et al., 2014) par le biais d'un système qui augmente les objets physiques du musée en utilisant la vidéoprojection. Pour pointer, il faut utiliser un contrôleur à 6 degrés de liberté ou d'un Leap Motion. La principale limitation de ces approches est qu'elles nécessitent d'augmenter les vitrines avec des capacités d'affichages. De plus, les gestes en l'air souffrent d'un manque de délimiteur, ce qui peut conduire à de faux positifs (c'est-à-dire des interactions que l'utilisateur ne veut pas réaliser). Enfin, la nécessité de suivre spatialement la main ou les doigts ne peut généralement être réalisée que sur un volume réduit.

L'interaction au contact de la vitrine présente plusieurs avantages : 1) elle offre un délimiteur physique pour l'interaction (toucher la vitrine), matérialisant le lieu de l'interaction et évitant les activations involontaires ; 2) elle ne détourne pas l'attention visuelle de l'utilisateur du contenu physique affiché à l'intérieur de la vitrine et 3) elle est techniquement robuste, peu couteuse et facile à mettre en œuvre sur les vitrines en verre existantes grâce aux technologies tactiles actuelles (p. ex., les cadres infrarouges). Dans ce chapitre, nous explorons comment sélectionner un objet physique derrière une vitrine en utilisant une interaction tactile 2D sur la vitrine. Nous analysons d'abord les différentes approches « *walk-up-and-use* » (Lewis et al., 1990; Marshall et al., 2011; Wigdor et al., 2009) que les utilisateurs emploieraient face à une vitrine tactile. Dans ce but, nous avons mené une étude explorant la manière spontanée des utilisateurs de sélectionner un objet depuis une vitrine. Cependant, contrairement à l'interaction directe classique sur les écrans tactiles, cette configuration introduit une séparation physique entre l'objet à sélectionner (présenté dans la vitrine) et la surface tactile (c'est-à-dire la vitrine). Cette séparation peut affecter la capacité de l'utilisateur à toucher la surface à un endroit qui est associé à la cible derrière la vitrine.

Dans un deuxième temps, nous évaluons l'effet de la taille de la cible, de la distance de la cible, de la présence de références spatiales et de la position de la tête de l'utilisateur par rapport à la vitrine, sur la sélection de la cible en utilisant un lancer de rayon.

1 Conception d'une technique de sélection depuis une vitrine

Ce travail porte sur l'utilisation d'une vitrine interactive pour sélectionner un objet distant, c'est-à-dire un objet hors de portée de l'utilisateur. La sélection tactile directe et indirecte est largement étudiée dans différents contextes (grandes surfaces (D. Schmidt et al., 2009), interaction avec un seul pointeur (Sears & Shneiderman, 1991), saisie avec un stylo (Forlines & Balakrishnan, 2008), etc.). L'originalité de notre approche est que l'objet à sélectionner n'est pas seulement derrière la surface tactile, mais physiquement distant de celle-ci. Dans un tel contexte, le lancer de rayon est l'une des techniques les plus utilisées pour la sélection avec une relation spatiale entre le dispositif et le référent, tandis que l'utilisation d'un support de contact (c'est-à-dire une surface tactile) séparée est l'une des techniques les plus fréquentes pour la sélection sans relation spatiale (cf. chapitre État de l'art section 2.2). Nous cherchons à tirer profit de ces deux approches pour sélectionner en utilisant la vitrine comme surface de contact.

Cependant, combiner ces deux approches dans le monde physique lève inévitablement des problèmes et contraintes liés au lancer de rayon depuis une vitrine d'une part et liés à l'usage d'une surface tactile comme support de contact pour sélectionner un objet à distance d'autre part.

Dans un premier temps, nous détaillons les problèmes et les contraintes d'une technique basée sur un lancer de rayon depuis une vitrine selon notre contexte. Ces lancers de rayons doivent ensuite être contrôlés par des actions de l'utilisateur. Communément, un lancer de rayon est défini selon la position et l'orientation du dispositif (contrôleur ou main). La technique est alors qualifiée d'absolue (Hinckley & Wigdor, 2002). Dans un second temps, nous présentons une définition des techniques de sélection absolues ainsi que les facteurs influençant la sélection d'objets avec les techniques de sélection absolues.

1.1 Problèmes et contraintes liés au lancer de rayon depuis une vitrine comme technique de sélection à distance

La plupart des techniques de sélection d'objets dans l'espace sont liées au lancer de rayon, qui est la technique standard pour sélectionner des objets dans des environnements virtuels (Argelaguet & Andujar, 2013; Debarba et al., 2013; Mine, 1995). Pour contrôler la direction du rayon dans le monde réel, la plupart des approches utilisent soit le doigt (Matulic & Vogel, 2018; Song et al., 2000), soit un objet portable (Gallo et al., 2008; Hincapié-Ramos, Guo, & Irani, 2014; Teather & Stuerzlinger, 2011). Ces techniques ont été utilisées dans les affichages volumétriques et les systèmes de réalité augmentée (Grossman et al., 2005; Grossman & Balakrishnan, 2006; Seonho Lee et al., 2013), qui peuvent afficher le rayon entier. Cependant, l'utilisation de ces techniques de lancer de rayon dans le monde réel et au contact d'une vitrine pose deux problèmes majeurs : l'effet de parallaxe et le manque de retour visuel pour afficher le rayon en l'air ou le curseur résultant. Ces deux problèmes sont évoqués dans les deux sous-sections suivantes.

1.1.1 Effet de parallaxe

Dans le contexte des surfaces tactiles, l'effet de parallaxe a été défini comme « *une différence entre l'emplacement perçu d'un élément particulier de l'interface utilisateur et l'emplacement réel du toucher de l'utilisateur sur la surface* » (Khamis et al., 2018). L'effet de parallaxe est lié à l'incapacité pour l'œil humain de se concentrer simultanément sur deux profondeurs différentes (cet effet est aussi appelé parallaxe binoculaire), par exemple le bout du doigt et l'objet distant (cf. Figure 4.2). Migge et Kunz (Migge & Kunz, 2010) définissent les erreurs

dues à cette parallaxe comme des « erreurs de parallaxe ». Un exemple de cet effet peut être trouvé sur des écrans interactifs publics protégés par une vitre tactile épaisse (p. ex. un distributeur automatique de billets). Dans ce contexte, Khamis et collab. (Khamis et al., 2018) étudient l'écart entre le toucher perçu et le pointeur (c'est-à-dire le toucher reconnu par le système). Ils explorent l'effet de plusieurs méthodes de correction pour augmenter la précision du toucher, ce qui a amélioré de manière significative la précision des tâches de sélection. Dans un contexte différent, Lee et collab. (J. H. Lee & Bae, 2013) conçoivent un curseur binoculaire pour améliorer les performances de sélection. Dans les deux cas, les solutions proposées nécessitent l'affichage d'un retour visuel pour atténuer l'effet de parallaxe. Cependant, dans notre contexte, la vitrine n'a aucune capacité d'affichage.

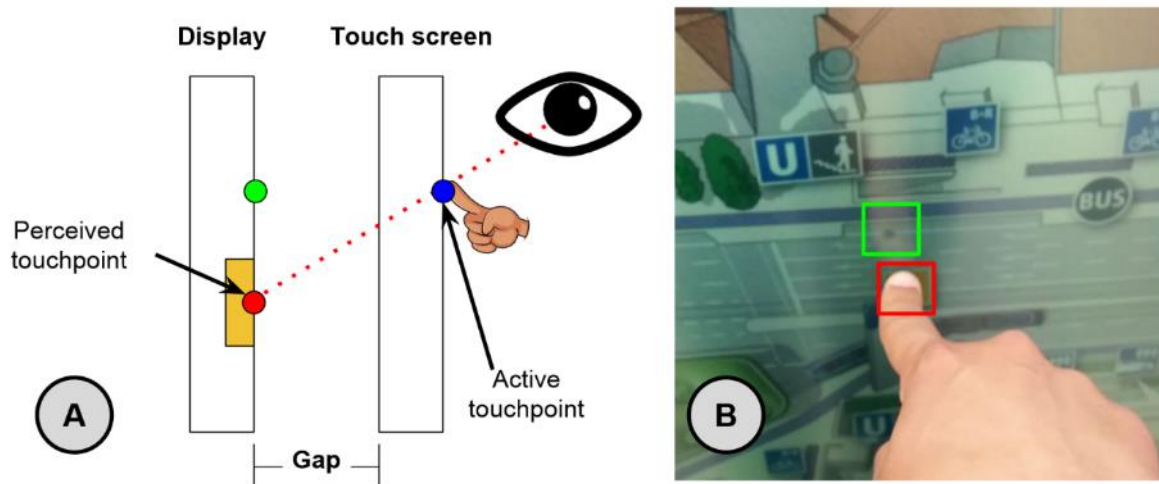


Figure 4.2 – (A) Explication et (B) exemple de parallaxe avec en rouge la zone perçue par l'utilisateur et en vert la zone réellement sélectionnée (Khamis et al., 2018)

1.1.2 Manque de retour visuel

L'affichage de retour visuel à l'intérieur de la vitrine est difficile. En effet, certains objets physiques ne doivent pas être exposés à une lumière directe (p. ex. une peinture dans un musée). L'affichage de retour visuel pour un lancer de rayon se traduit généralement par une droite numérique virtuellement tracée dans l'air entre l'origine et la cible. Rendre perceptible cette droite dans l'air hors des contextes de visualisation immersive (RA, RV) est problématique. Plusieurs solutions ont été explorées pour proposer un usage de lancer de rayon sans l'affichage de ce rayon. Par exemple, Bandyopadhyay et collab. (Bandyopadhyay et al., 2001) utilisent un stylet comme dispositif de sélection pour sélectionner et interagir avec des objets physiques. Le système composé de deux vidéoprojecteurs permet aux utilisateurs de sélectionner différentes commandes (p. ex. des couleurs) sur une tablette physique et ensuite les appliquer à un objet ou une partie de l'objet physique (cf. Figure 4.3 A). Le rendu est fait automatiquement à l'aide de

détecteurs infrarouges. Myers et collab. (Myers et al., 2002) étudient plusieurs pointeurs laser pour sélectionner des cibles (cf. Figure 4.3 B). Bien que ces approches fonctionnent généralement bien avec un objet de grande taille, elles peuvent être limitées lorsqu'il s'agit de petits objets, comme des insectes dans un musée ou des briques Lego, ou lorsqu'il n'y a pas de surface de projection entre deux objets.

Récemment, Freeman et collab. (Freeman et al., 2018) étudient la sélection de petits objets sphériques en lévitation à l'aide de gestes de l'index en l'air (cf. Figure 4.3 C). La technique consiste en un lancer de rayon, sans aucun retour visuel du rayon. Cependant, comme l'étude porte sur une très petite distance entre le doigt et la cible (de 5 cm à 8 cm), leurs résultats sont difficiles à étendre à notre contexte d'utilisation où les objets sont généralement plus éloignés de la vitrine.



Figure 4.3 – Exemples d'interaction sans retour visuel avec de gauche à droite, (A) Dynamic Shader Lamps (Bandyopadhyay et al., 2001), (B) sélection avec des pointeurs laser (Myers et al., 2002) et (c) sélection de sphères physiques en lévitation (Freeman et al., 2018)

1.2 Facteurs influençant les techniques de sélection absolues pour la sélection d'objets distants avec une surface tactile comme support de contact

Il existe deux types de techniques de sélection tactile à distance : absolue (c'est-à-dire que le dispositif d'entrée envoie sa position (x, y)) ou relative (c'est-à-dire que le dispositif d'entrée envoie son déplacement (Δ_x, Δ_y)) (Hinckley & Wigdor, 2002). Les techniques relatives nécessitent de visualiser en permanence la position actuelle du pointeur. Dans notre contexte, l'affichage d'un pointeur n'est pas viable (comme expliqué dans la section précédente), cette section se concentre donc uniquement sur les techniques de sélection absolues.

Des travaux antérieurs ont exploré les principaux facteurs influençant la sélection absolue. Un premier facteur important est la présence d'une référence spatiale tactile. Par exemple, des repères spatiaux placés à l'arrière de l'appareil contribuent à augmenter la précision. Avec HaptiCase (Corsten et al., 2015), Corsten et collab. proposent un système de repères tactiles spatiaux placés à l'arrière d'un smartphone (cf. Figure 4.4). L'utilisateur peut sentir sous son

index les repères spatiaux tout en gardant leurs pouces au contact du smartphone et l'utilisateur peut sélectionner plusieurs cibles placées sur un écran en effectuant un pincement (c'est-à-dire serrer le pouce et l'index). Ces repères spatiaux ont permis d'augmenter la précision d'une sélection sans regarder l'écran. Ces repères spatiaux ont montré leur efficacité pour d'autres tâches. Fruchard et collab. (Fruchard et al., 2017) utilisent les repères spatiaux pour augmenter les performances d'une sélection sur un ordinateur portable avec un pavé tactile.

Un deuxième facteur est la taille relative et les ratios entre le dispositif d'entrée et le dispositif d'affichage, qui peuvent influencer négativement le taux de réussite de la sélection de la cible (Gilliot et al., 2014). En effet, Gilliot et collab. montrent que les utilisateurs font moins d'erreurs si le ratio entre dispositif d'entrée et dispositif de sortie est identique. Dans HaptiCase (Corsten et al., 2015), présenté précédemment, Corsten et collab. utilisent le même ratio entre dispositif d'entrée et dispositif de sortie.

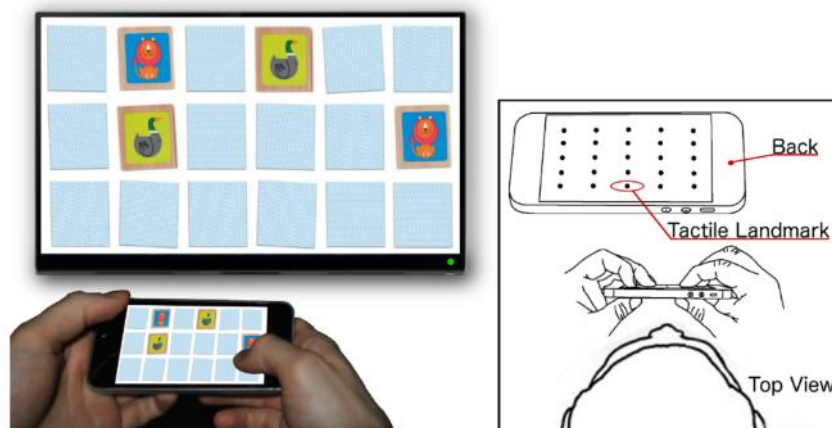


Figure 4.4 – Le concept d'HaptiCase (Corsten et al., 2015), avec (à gauche) la tâche réalisée dans l'étude et (à droite) la présentation du dispositif avec les repères spatiaux placés derrière l'écran

Enfin, un troisième facteur important est la présence de retour visuel sur le dispositif d'entrée (Gilliot et al., 2014). Il est en effet préférable, même si le dispositif d'entrée n'a aucune capacité d'affichage, que l'utilisateur puisse le regarder. Des études antérieures sur l'entrée tactile pour les grands écrans projetés ont révélé que, même sans retour visuel, les utilisateurs peuvent sélectionner 90% des cibles (Pietroszek & Lank, 2012). Cela signifie que la sélection absolue est toujours un bon candidat lorsqu'aucun retour visuel n'est fourni.

Cependant, ces trois études se sont concentrées sur l'interaction avec des écrans tactiles 2D alors que dans notre cas, nous étudions l'utilisation d'une surface en verre transparent avec un grand espace physique entre la vitrine et la cible.

1.3 Synthèse

D'une part, le lancer de rayon est une technique de sélection populaire pour atteindre des objets distants quand il y a une relation spatiale entre le dispositif et le référent, mais il peut introduire un effet de parallaxe et souffrira du manque de retour visuel dans le monde réel. D'autre part, les techniques basées sur un support de contact semblent très efficaces même sans retour visuel, mais ont tendance à rompre le flux d'interaction car elles divisent l'attention de l'utilisateur entre différents dispositifs d'entrée et de sortie.

Dans notre travail, nous explorons comment combiner ces deux approches pour tirer parti de leurs avantages respectifs, ce qui aboutit au concept de lancer de rayon à partir d'une surface en verre. Dans un contexte public, l'utilisation d'une technique de lancer de rayon devrait rendre l'interaction facile à découvrir et à apprendre. La vitrine peut être considérée comme une surface d'entrée, qui en fournissant un délimiteur physique, devrait faciliter la sélection sans retour visuel et, contrairement aux gestes en l'air, permettrait à l'utilisateur d'être moins sujet à la fatigue.

À notre connaissance, ce concept de lancer de rayon à partir d'une surface en verre n'a pas été étudié auparavant et constitue une configuration d'interaction différente de celles citées dans l'état de l'art : la surface tactile est transparente et l'objet à sélectionner est à une certaine distance derrière elle. Par conséquent, nous avons adopté une approche en deux étapes pour explorer ce nouvel environnement. D'abord, nous avons mené une étude informative pour trouver l'interaction la plus naturelle dans un contexte public. Ensuite, nous avons exploré plus en détail l'approche la plus utilisée dans une expérience contrôlée pour évaluer sa performance.

2 Étude informative

Si l'utilisation d'une vitrine pour sélectionner un objet physique placé derrière la vitrine présente plusieurs avantages, comme nous l'avons déjà mentionné, le lancer de rayon nécessite la définition d'une origine et d'une direction ou d'une cible (cf. chapitre État de l'art section 2.2.2). Il existe donc plusieurs façons de lancer un rayon à partir de la surface de la vitrine (cf. Figure 4.5) : en utilisant l'orientation du doigt (c'est-à-dire le toucher sur la vitrine définit l'origine et l'orientation du doigt sa direction), en visant l'objet avec le doigt (c'est-à-dire l'œil définit l'origine du rayon et la droite « œil-doigt » indique sa direction) ou en plaçant le doigt en face de l'objet (c'est-à-dire le toucher sur la vitrine définit l'origine et l'orientation est toujours orthogonale à la vitrine). Cependant, une technique d'interaction appropriée pour un contexte

public doit être aussi explicite que possible (Marquardt & Greenberg, 2012), afin que les passants interagissent avec la vitrine avec un minimum d'instructions.

Pour trouver l'interaction la plus appropriée, nous avons mené une étude informative dans un environnement public avec des passants. L'objectif de cette étude était d'explorer comment les participants utiliseraient intuitivement la vitrine pour sélectionner un objet derrière elle.

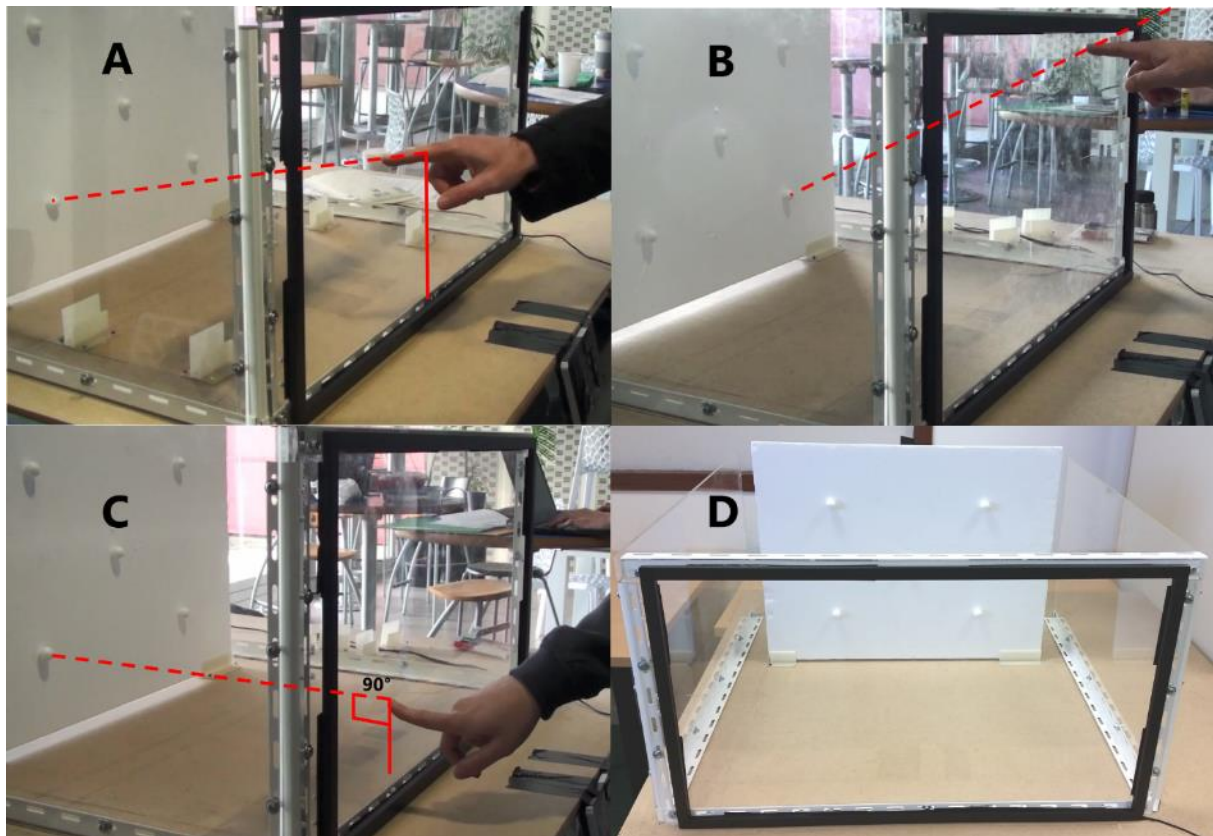


Figure 4.5 – Les trois types de lancers de rayon avec (A) le rayon orienté selon l'orientation du doigt, (B) un lancer de rayon oeil-doigt et (C) un lancer orthogonal. Une vue d'ensemble du dispositif utilisé pour l'étude (D)

2.1 Protocole expérimental

2.1.1 Tâche et instructions

Cinq sphères physiques blanches d'un diamètre de 2 cm étaient positionnées sur un carton placé à 55 cm derrière la vitrine. Quatre d'entre elles définissaient les coins d'un rectangle de 33x26 cm de côté, tandis que la cinquième était positionnée au milieu du rectangle (cf. Figure 4.5).

Les participants ont reçu l'instruction suivante : « *Veillez sélectionner chacune des cinq sphères blanches, en utilisant la vitrine en verre devant vous comme support à l'interaction* ». Avant de sélectionner une cible, les participants devaient préciser oralement la cible qu'ils comptaient sélectionner. Les participants devaient toutes les sélectionner dans un ordre libre.

2.1.2 *Participants*

Nous avons recruté 46 participants (39 hommes et 7 femmes) âgés de 19 à 62 ans ($M= 33$). Tous les participants ont été recrutés dans le hall du laboratoire. Les participants ont été conviés à l'expérimentation dès lors qu'ils s'approchaient du dispositif. 13 étaient des chercheurs et chercheuses de l'université, 26 étaient des étudiants et étudiantes en informatique, cinq étaient des ingénieurs et deux étaient des employés administratifs.

2.1.3 *Matériel*

La vitrine a été installée pendant 10 heures dans le hall de notre laboratoire. La vitrine (78,5 x 44 cm) était placée sur une table de 90 cm de hauteur. Les utilisateurs devaient se placer au milieu de la vitrine et à une distance où toutes les cibles étaient visibles et la vitrine facilement accessible avec le bras (cf. Figure 4.5 D).

2.1.4 *Données recueillies*

Toute l'expérience a été filmée. Une fois tous les essais terminés, nous avons demandé aux participants de fournir une représentation graphique de la façon dont leur doigt était positionné et orienté pour effectuer chaque tâche de sélection. Sur la base des enregistrements vidéos et des représentations fournies par les participants, nous avons regroupé les résultats en trois catégories, correspondantes aux approches possibles pour lancer un rayon depuis la vitrine :

1. Le lancer de rayon orienté
2. Le lancer de rayon œil-doigt
3. Le lancer de rayon orthogonal.

2.2 Résultats

Parmi les 46 participants, 40 d'entre eux ont utilisé un lancer de rayon pour sélectionner les objets, comme prévu, tandis que 4 participants ont utilisé des entrées gestuelles et 2 participants ont mal compris les instructions (c'est-à-dire qu'ils ont essayé d'atteindre les objets en évitant la vitrine).

Nos résultats révèlent que parmi ces 40 participants, 5 participants (12,5%) ont lancé un rayon orienté (c'est-à-dire en utilisant l'orientation de leur doigt) pour sélectionner les cibles ; 5 participants (12,5%) ont lancé un rayon œil-doigt (c'est-à-dire en alignant l'œil, le doigt et l'objet) ; tandis que 30 participants (75%) ont lancé un rayon orthogonal à la surface de la vitrine, comme l'illustre la Figure 4.6.

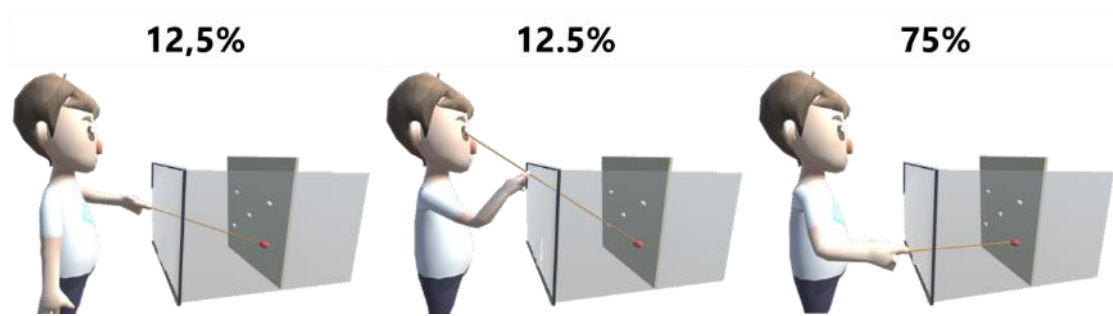


Figure 4.6 – Pourcentage de participants ayant utilisé chaque type de lancer de rayon : lancer de rayon orienté (gauche), lancer de rayon œil-doigt (centre) et lancer de rayon orthogonal (droite)

2.3 Conclusion

75% des participants qui ont utilisé un lancer de rayon ont adopté un rayon orthogonal comme la technique la plus « *walk-up and use* » dans un contexte public. Une explication possible de ce résultat est que cette approche élimine l'effet de parallaxe puisque le doigt, l'œil et l'objet n'ont pas besoin d'être alignés. Par conséquent, dans nos travaux ultérieurs, nous prenons en compte cette interaction et n'essayons pas de compenser l'effet de parallaxe. Incidemment, cette interaction ne nécessite que le suivi de la position du doigt de l'utilisateur sur la vitrine, ce qui rend son implémentation simple en utilisant un cadre infrarouge. Le principal défi consiste alors à évaluer si cette forme d'interaction est efficace en matière de précision. Pour aborder cette question, nous réalisons une étude de sélection contrôlée.

3 Étude de la sélection : conception

L'objectif de cette étude est d'examiner l'impact de quatre facteurs sur la sélection d'objets derrière une vitrine en utilisant un rayon projeté perpendiculairement à la vitre. Ces facteurs sont :

1. La taille de la cible
2. La distance entre la cible et la vitrine
3. La présence de références spatiales
4. La position de la tête de l'utilisateur

De plus, puisque l'interaction dans les espaces publics doit être facile et rapide à utiliser, nous avons également étudié la position du rayon lors du premier contact avec la vitrine tactile (c'est-à-dire avant tout ajustement) en matière de déviation par rapport à la cible. En effet, traditionnellement sur un écran 2D le pointage d'un élément est décomposé en deux phases : un premier mouvement rapide pour sélectionner la cible et, en cas d'échec de ce premier

mouvement, un deuxième mouvement permettant de corriger la sélection (Meyer et al., 1988). Il nous apparaît donc intéressant d'étudier ces deux phases pour la sélection d'objets physiques depuis une vitrine. Enfin, comme notre objectif principal est d'explorer les performances d'un nouveau paradigme de sélection ancré dans un contexte spécifique, nous ne l'avons pas comparée à une technique de référence. En effet, aucune technique de référence ne pourrait s'adapter à notre contexte sans retour visuel avec une approche « *walk-up and use* ».

3.1 Tâche et instructions

La tâche consistait à sélectionner une cible physique sphérique positionnée sur un panneau à une certaine distance de la vitrine, sans aucun retour visuel pouvant guider la sélection. Un retour visuel indiquait par contre quelle cible sélectionner. Les utilisateurs devaient appuyer sur un bouton du clavier pour commencer un essai, puis toucher la vitrine à la position prévue de la cible, déplacer le doigt pour ajuster leur sélection si cela était nécessaire (c'est-à-dire si le pointage au premier contact n'était pas sur la cible), et retirer le doigt de la vitrine pour terminer l'essai. Un retour visuel indiquait alors si la cible avait bien été atteinte par le rayon.

Le clavier pour commencer un nouvel essai était placé en dehors de la vitrine (cf. Figure 4.7). Les participants devaient se concentrer sur autre chose que la vitrine. Cette perturbation est destinée à limiter considérablement tout effet d'apprentissage. Pour éviter tout impact de la taille du doigt de l'utilisateur et pour évaluer le pire scénario, nous avons utilisé un diamètre de rayon de 0,2 mm. Nous discutons de l'impact de la taille des rayons dans la section Discussion.

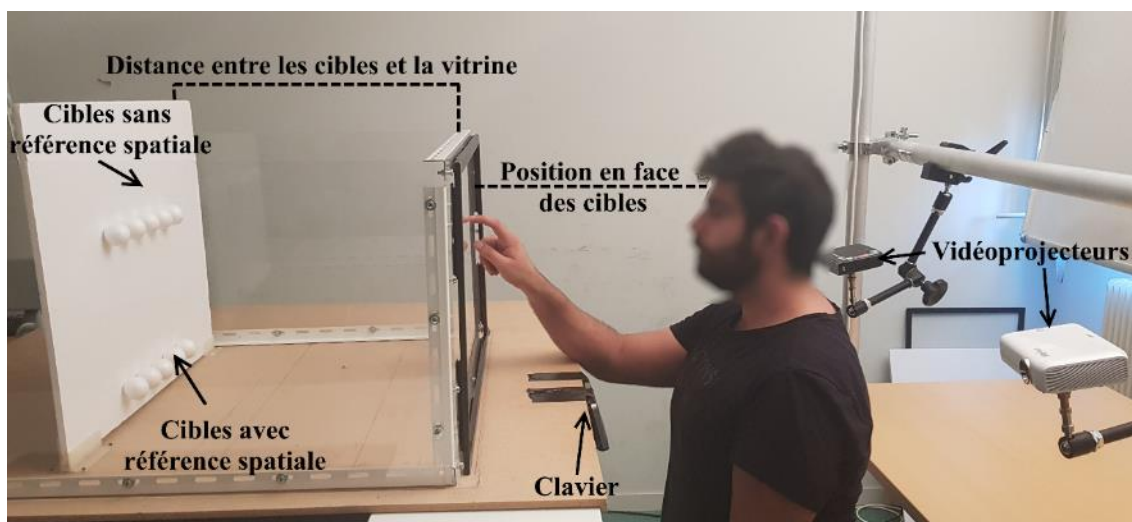


Figure 4.7 – Matériel utilisé de notre expérimentation

Nous avons demandé aux participants de sélectionner la cible aussi précisément que possible du premier coup (c'est-à-dire que le premier contact avec la vitrine devait sélectionner la cible)

et, s'ils n'atteignaient pas la cible, de déplacer le doigt aussi vite que possible pour ajuster leur sélection. Ils devaient utiliser l'index de leur main dominante. Comme nous l'avons mentionné précédemment, la projection, et donc le retour d'information dynamique, n'est pas toujours possible dans les applications dans un contexte public. Par conséquent, le seul retour visuel d'interaction utilisé dans notre étude est un ensemble de 4 points blancs entourant la prochaine cible à sélectionner. Les points devenaient verts ou rouges pour indiquer le succès ou l'échec de la sélection. Ce retour visuel permettait à l'utilisateur de rester concentré sur la cible tout au long de l'expérience.

Bien sûr, dans un contexte public, la sélection entraînerait l'affichage d'informations liées à l'objet pour l'utilisateur. Ce retour est généralement affiché sur un écran séparé, mais il pourrait également être présenté de manière plus directe en utilisant un son spatialisé ou une projection à l'intérieur de la vitrine. Quoi qu'il en soit, l'exploration de la manière de rendre cette information n'entre pas dans le cadre de notre travail.

3.2 Conditions

Nous avons évalué différentes tailles de cible, distances de la cible par rapport à la vitrine, proximité de la cible par rapport à une référence spatiale et position de la tête de l'utilisateur par rapport à la vitrine (et donc aux cibles). Nous avons défini quatre tailles de cibles, de 1 à 4 cm de diamètre, afin de couvrir le spectre des plus petites cibles trouvées dans un contexte public. Ces tailles correspondent aussi aux différentes tailles des briques Lego utilisées dans notre outil logiciel (cf. chapitre Contexte et Outils section 2.3.2). Des tailles similaires ont également été examinées dans des travaux antérieurs sur les cibles pour le toucher (A. D. Hall et al., 1988; Vogel & Baudisch, 2007; F. Wang & Ren, 2009), jugées atteignables de 10,5 mm à 26 mm.

Nos cibles ont été placées à trois distances différentes de la vitrine (15, 35 et 55 cm) pour représenter des distances proches, moyennes et lointaines.

Pour chaque distance et taille de cible, nous avons positionné 10 cibles sphériques (cf. Figure 4.7) sur un panneau. Les cibles étaient alignées sur 2 lignes horizontales, soit 5 cibles par ligne, afin d'explorer l'impact d'une référence spatiale sur la sélection. Cette condition exigeait de fixer les cibles aux mêmes endroits. La ligne supérieure se trouvait légèrement au-dessus du milieu de la vitrine (c'est-à-dire sans aucune référence spatiale), tandis que la ligne inférieure était proche du bas de la vitrine (agissant comme une référence spatiale).

Enfin, l'utilisateur était alternativement assis, c'est-à-dire avec la tête en face des cibles, ou debout, c'est-à-dire avec la tête au-dessus des cibles. Nous avons veillé à ce que tous les participants positionnent leur tête et leur corps à la même hauteur en utilisant un marchepied.

3.3 Participants

Nous avons recruté 12 participants âgés de 21 à 48 ans ($M= 29,5$, $SD= 8,26$). 11 participants ont été recrutés à l'université et étaient des étudiants de premier cycle (4), des doctorants (4), des enseignants-chercheurs (3) et un participant était ingénieur dans une entreprise privée. Dix participants étaient droitiers et deux étaient gauchers.

3.4 Conception et procédure

L'expérience suivait un plan intra-participant $4 \times 3 \times 2 \times 2$ avec la taille de la cible (sphère de 1, 2, 3 ou 4 cm de diamètre), la distance de la cible à la vitrine (15, 35 et 55 cm), la référence spatiale (avec ou sans) et la position de la tête de l'utilisateur (en face ou au-dessus) comme facteurs. La position de la tête de l'utilisateur était contrebalancée entre les participants, et la distance et la taille étaient ordonnées de manière aléatoire de sorte que deux essais consécutifs avec la même distance ou taille ne pouvaient pas avoir lieu. Le reste des essais était ordonné de manière aléatoire. L'ensemble de l'expérience durait entre 70 et 110 minutes.

Nous avons créé un ensemble spécifique de 10 cibles (contenant les différentes tailles) pour l'entraînement. La session d'entraînement consistait à sélectionner chaque cible de cet ensemble pour chaque distance et hauteur (c'est-à-dire $10 \times 3 \times 2 = 60$ essais). Cet entraînement durait environ 10 minutes et représente 7,6% de l'ensemble de l'expérience.

Au total, nous avons recueilli 4 tailles \times 3 distances \times 2 références spatiales \times 2 positions de la tête de l'utilisateur \times 5 cibles \times 3 répétitions = 720 essais par participant et 8640 essais au total.

3.5 Matériel

Nous avons construit une vitrine en utilisant quatre plaques de plexiglas de 78,5 cm \times 44 cm \times 0,5 cm (LxHxP). Pour le suivi des doigts, nous avons utilisé un cadre infrarouge (ZaagTech X series v7, 70cm \times 40cm, avec 40 points de contact, résolution de 32 768*32 768, temps de réponse tactile de 7ms-13ms), connecté par USB à l'ordinateur exécutant l'étude. La surface d'interaction tactile disponible était de 70 cm \times 40 cm. Le cadre a été fixé sur un côté de la vitrine. Nous avons également renforcé la plaque de plexiglas en ajoutant une structure métallique tout autour pour éviter toute déformation ou tout mouvement de la plaque pendant l'interaction. Nous avons utilisé deux vidéoprojecteurs, reliés par un hub DisplayPort - HDMI,

pour nous assurer que le point indiquant la cible à sélectionner s'affichait même si l'utilisateur obstruait l'un d'eux. Nous avons veillé à ce qu'il n'y ait pas de reflet lumineux sur la vitrine qui pourrait gêner l'interaction. Entre chaque participant, la vitrine a été nettoyée pour éliminer toutes les traces de doigts. Dans un contexte public, un spray anti-empreintes digitales peut être utilisé pour limiter toute salissure au toucher.

Nous avons implémenté l'ensemble du dispositif sur un ordinateur portable HP EliteBook fonctionnant sous Microsoft Windows 10. Notre logiciel d'expérimentation a été implémenté en C# et Unity (version 2018 1.5). Les événements tactiles ont été envoyés à l'application en utilisant le protocole TUIO. Nous avons modélisé l'environnement physique dans Unity pour nous assurer que les informations projetées pour l'expérimentation s'affichaient autour de la cible.

3.6 Données collectées et traitement des données

Nous avons enregistré tous les événements tactiles, de l'appui sur le clavier jusqu'au retrait du doigt de la vitrine. Comme nous voulions comprendre comment les utilisateurs effectuaient le premier contact avec la vitrine, nous avons enregistré pour chaque essai la position et le temps du premier contact et du retrait du doigt (c'est-à-dire le dernier contact avec la vitrine).

Nous avons affiné l'analyse des points de contact avec la vitrine à l'aide de deux mesures (cf. Figure 4.8) déjà utilisées (Holz & Baudisch, 2010; Roudaut et al., 2011) : *l'offset* et le *spread*. Ces deux métriques reposent sur le calcul du centroïde, c'est-à-dire la position moyenne des points de premier contact pour chaque cible et chaque condition.

L'*offset* est défini comme la distance euclidienne (c'est-à-dire la valeur absolue) entre le centre de la cible et le centroïde, pour chaque condition. Dans cette section, nous calculons également *l'offsetX* et *l'offsetY*, qui représentent la distance relative en x et y entre le centroïde et le centre de la cible (c'est-à-dire une valeur positive ou négative).

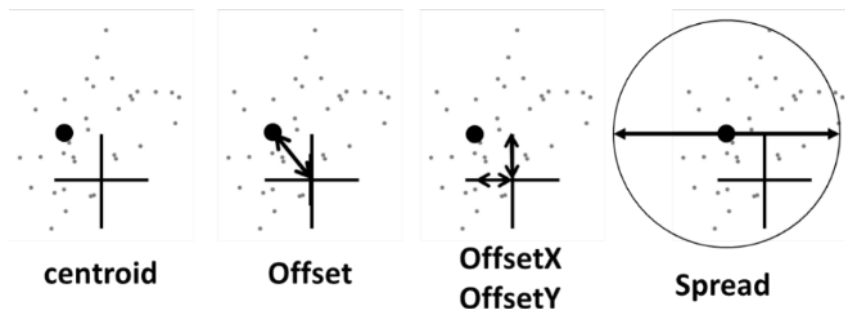


Figure 4.8 – Illustration des paramètres utilisés pour analyser les premiers contacts

Le *spread* représente la distribution des points de contact autour du centroïde, calculé comme le diamètre du plus petit cercle contenant 95% de tous les points de contact.

En ce qui concerne l'analyse des données, nous avons choisi de nous appuyer sur des techniques d'estimation avec des intervalles de confiance à 95%, comme le recommande l'APA (American Psychological Association, 2020). Tous les IC sont des intervalles de confiance de 95 % BCa bootstrap. Cette technique d'estimation a été utilisée dans de nombreux travaux récents dans la communauté de l'Interaction Humain-Machine (Danyluk et al., 2019; Desai et al., 2019; Jansen et al., 2019; Keddisseh et al., 2021; Kuhlman et al., 2019; Mylavarapu et al., 2019; Simon et al., 2019).

4 Résultats de l'étude de sélection

Nous divisons l'analyse en deux parties. Tout d'abord, nous présentons les résultats concernant la tâche de sélection de la cible en matière de taux de réussite et de temps de réalisation de la tâche. Ensuite, nous analysons le geste de premier contact en matière d'*offset* et de *spread*, comme expliqué précédemment.

4.1 Taux de réussite de la tâche de sélection

Nous avons commencé par analyser les résultats pour chaque taille de cible, indépendamment de la distance de la cible, de la position de la tête de l'utilisateur ou de la référence spatiale. Sans surprise, les cibles de 1 cm étaient les plus difficiles à atteindre avec seulement 55,8% (IC [48,7% ; 65,1%]) d'essais réussis (cf. Figure 4.9). Les cibles de 2 cm ont été plus faciles à acquérir mais restent généralement difficiles, avec un taux de réussite de 78,1% (IC [75,8% ; 80,6%]). Les cibles de 3 cm et 4 cm ont un taux de réussite supérieur à 90% (cibles de 3 cm : 90,8%, IC [89,5% ; 92,4%] ; cibles de 4 cm : 93,2%, IC [90,3% ; 95,3%]). Pour ces deux tailles de cibles plus grandes, nous avons observé au cours de l'expérience qu'après quelques essais réussis, certains utilisateurs exécutaient la tâche trop rapidement, c'est-à-dire qu'ils retiraient le doigt avant que le retour visuel indiquant la réussite de l'essai n'apparaisse, ce qui entraînait quelques sélections erronées.

En ce qui concerne la distance entre la vitrine et les cibles, indépendamment des autres facteurs, il apparaît que les cibles placées à 15 cm étaient toujours plus faciles à sélectionner (taux de réussite : 82,5% IC [80,1% ; 85,4%]) que celles placées à 55cm (taux de réussite : 76,6% IC [73,1% ; 80,4%]). Cependant, aucune conclusion claire ne peut être faite lorsque l'on compare

la condition 15cm à la condition 35cm (79,3% IC [75,7% ; 83%]). Ce résultat est valable pour toutes les tailles de cibles, à l'exception de la cible de 1 cm, qui est aussi difficile à sélectionner pour les trois distances (c'est-à-dire que tous les IC se chevauchent, cf. Figure 4.9).

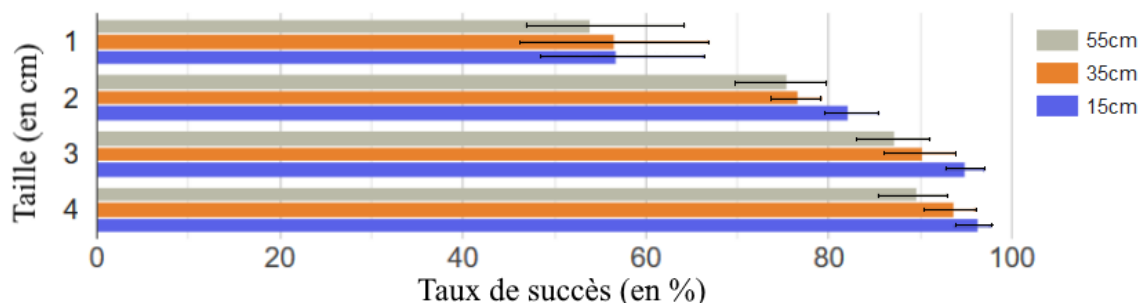


Figure 4.9 – Taux de réussite moyen en % (avec IC 95 %) selon la taille de la cible et la distance à la vitrine

En ce qui concerne la présence d'une référence spatiale, les résultats établissent clairement que les cibles avec une référence spatiale (c'est-à-dire celles proches de la ligne du bas) sont plus faciles à sélectionner (taux de réussite : 83%, IC [80,5% ; 86,1%]) que celles sans référence spatiale (taux de réussite : 76%, IC [73% ; 79,8%]). Cela pourrait s'expliquer par le fait que la référence spatiale, c'est-à-dire la partie inférieure de la vitrine, facilite la précision de la sélection. Ce résultat est en accord avec des travaux antérieurs sur la sélection tactile absolue (Pietroszek & Lank, 2012).

D'après les données recueillies, la position de la tête de l'utilisateur ne semble pas influencer le taux de réussite (au-dessus : 79,2%, IC [75,9% ; 82,7%] ; en face : 79,8% IC [76,9% ; 83,3%]). Ce résultat est assez surprenant, car nous nous attendions à ce que la condition où l'utilisateur est en face de la vitrine favorise la sélection des cibles.

4.2 Temps de réalisation des essais

En ce qui concerne le temps de réalisation des essais pour les différentes tailles de cibles, les résultats soulignent la difficulté de sélectionner les cibles de 1 cm. Dans ce cas, les utilisateurs ont mis presque 10 s pour réaliser la tâche (9,4 s, IC [8,1 ; 10,4]). Pour les autres tailles de cibles (cf. Figure 4.10), le temps de réalisation diminue drastiquement lorsque la taille de la cible augmente (cibles de 2 cm : 4,6 s IC [4,0 ; 5,1] ; cibles de 3 cm : 3 s IC [2,6 ; 3,3] ; cibles de 4 cm : 2,3 s IC [2,1 ; 2,6]).

Pour la distance de la cible, nous avons observé le même effet. Bien que le taux de réussite soit très similaire pour 35 et 55 cm, les résultats établissent fortement qu'en moyenne le temps de

réalisation est plus petit lorsque les cibles sont plus proches de l'utilisateur (15 cm : 4 s IC [3,5 ; 4,6] ; 35 cm : 5 s IC [4,5 ; 5,4] ; 55 cm : 5,33 s IC [4,6 ; 6,19], cf. Figure 4.10).

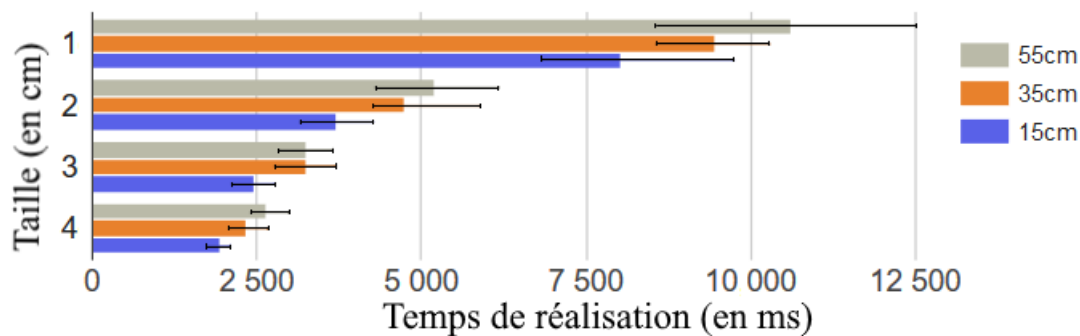


Figure 4.10 – Temps moyen de réalisation en ms (avec IC à 95%) en fonction de la taille de la cible et de la distance à la vitrine

Les résultats n'ont pas révélé d'effet de la référence spatiale ou de la position de la tête de l'utilisateur sur le temps de réalisation.

4.3 Sélection au premier contact : centroïde, offset et spread

Nous avons analysé les résultats de la sélection au premier contact en matière de centroïde, d'*offset* et de *spread*. Dans cette partie, nous utilisons un intervalle de confiance à 95% sans bootstrapping.

4.3.1 Analyse des centroïdes

L'affichage de toutes les positions des centroïdes sur le même référentiel donne une vue d'ensemble de la sélection au premier contact. Plus de 65% de tous les centroïdes sont situés à gauche des cibles. Près de 70% sont situés en haut et 45% dans la partie supérieure gauche (cf. Figure 4.11). Ce résultat indique que les utilisateurs avaient une tendance générale à

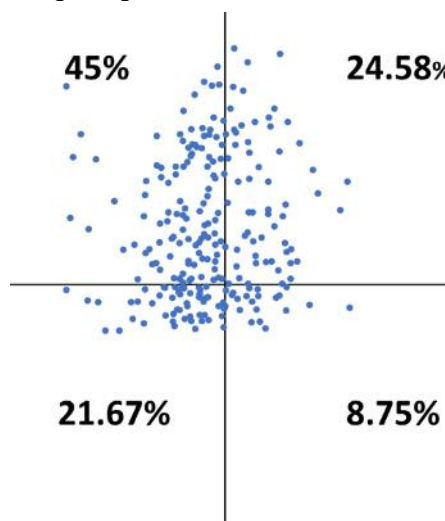


Figure 4.11 – Distribution de tous les centroïdes de sélection au premier contact autour des cibles. La plupart d'entre eux se trouvent dans la zone supérieure gauche

sélectionner trop haut et trop à gauche des cibles physiques. Nous affinons cette vue d'ensemble en examinant l'*offset*, l'*offsetX* et l'*offsetY* pour nos différentes conditions.

4.3.2 *Offset, offsetX et offsetY*

Nous n'avons trouvé aucune différence sur les valeurs d'*offset* en fonction de la taille de la cible, nous avons donc analysé les résultats pour toutes les tailles de cibles incluses. En ce qui concerne la distance de la cible, l'*offset* augmente avec la distance (cf. Figure 4.12): par rapport à l'*offset* calculé pour des cibles placées à 15 cm (0,66 cm IC [0,55 ; 0,76]), l'*offset* est 28,8% plus grand pour des cibles à 55 cm (0,85 cm IC [0,75 ; 0,95]) et 19,7% plus grand pour des cibles à 35 cm (0,79cm IC [0,68 ; 0,90]). Les résultats n'établissent pas de différence claire entre l'*offset* observé pour 35 et 55 cm.

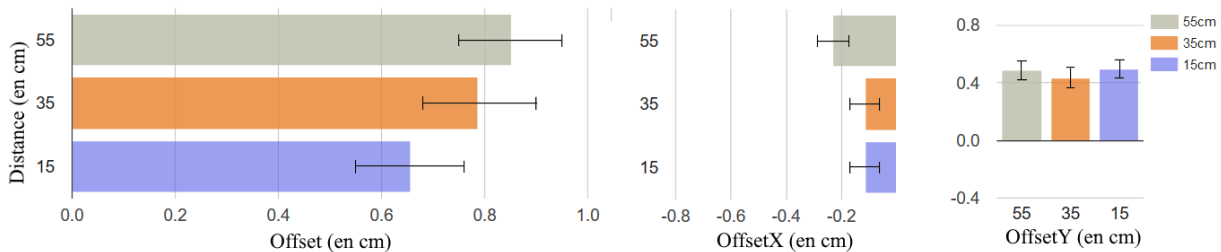


Figure 4.12 – Offset moyen (gauche), offsetX (centre) et offsetY (droite) en cm (avec IC 95 %) en fonction de la distance de la cible par rapport au verre

Nous affinons cette analyse en examinant l'*offsetX* et l'*offsetY*. Il est intéressant de noter que si nous trouvons la même tendance pour l'*offsetX* (c'est-à-dire que l'*offsetX* moyen est 29,8% plus grand pour les cibles à 55 cm et 16,4% plus grand pour les cibles à 35 cm que pour celles à 15 cm), les résultats ne révèlent aucune différence dans l'*offsetY* pour les cibles à différentes distances (cf. Figure 4.12). Nous pouvons conclure que le facteur distance a surtout un impact sur la précision de la sélection au premier contact sur l'axe horizontal.

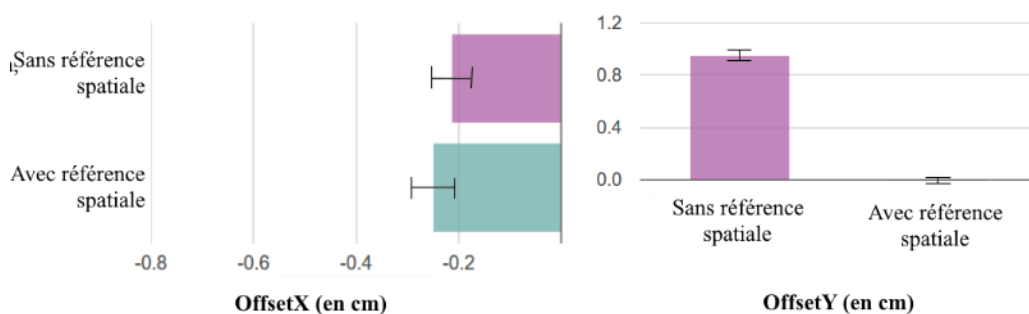


Figure 4.13 – OffsetX (gauche) et offsetY (droite) moyens en cm (avec 95% Cis) selon la condition de référence spatiale

En ce qui concerne l'effet de la référence spatiale sur la valeur de l'*offset*, nous avons observé que le décalage moyen est 125% plus grand pour les cibles sans référence spatiale que pour celles avec une référence spatiale : l'*offset* est de 1,06 cm (IC [1,02 ; 1,10]) sans référence spatiale contre 0,47 cm (IC [0,44 ; 0,49]) avec une référence spatiale. Lorsqu'on examine

l'impact de la position de la cible sur l'*offsetX* et l'*offsetY*, les résultats ne révèlent aucune différence sur l'*offsetX*. Cependant, l'*offsetY* est plus grand pour les cibles sans référence spatiale (0,95 cm IC [0,87 ; 1,04]) que pour celles avec une référence spatiale (0,01 cm IC [0,05 ; 0,04]), comme l'illustre la Figure 4.13.

Cette analyse peut être renforcée lorsque nous visualisons tous les centroïdes (cf. Figure 4.14) en fonction des deux positions de la cible. Nous pouvons voir la grande distribution verticale pour les cibles sans référence spatiale par rapport aux cibles avec référence.

Enfin, nous avons analysé l'*offset* en fonction de la position de la tête de l'utilisateur. Nous avons constaté que l'*offset* est presque 60% plus grand lorsque la tête de l'utilisateur se trouve en face de la vitrine (0,88 cm IC [0,82 ; 0,95]) qu'au-dessus (0,65 cm IC [0,61 ; 0,68]). Ce résultat est valable à la fois pour l'*offsetX* et l'*offsetY*. Ce résultat est surprenant car nous pourrions nous attendre à ce que les cibles dans la condition où l'utilisateur se trouve en face de la vitrine soient plus faciles à sélectionner, puisque le regard de l'utilisateur est aligné avec le doigt et la cible.



Figure 4.14 – Distribution des centroïdes en fonction de la référence spatiale : avec référence spatiale (gauche) ou sans référence spatiale (droite). Le point rouge correspond à la moyenne des centroïdes

4.3.3 Spread

Le *spread* représente la distribution des premiers points de contact autour du centroïde. Les résultats sont en accord avec nos conclusions précédentes. La distance de la cible a un impact sur le *spread*, qui est nettement plus élevé lorsque la distance augmente (15 cm : 2,95 cm IC [2,81 ; 3,10] ; 35 cm : 3,91 cm IC [3,76 ; 4,06] ; 55 cm : 4,23 cm IC [4,10 ; 4,36] cf. Figure 4.15).

En outre, les cibles avec une référence spatiale réduisent fortement le *spread* avec 3,24 cm (IC [3,15 ; 3,33]) par rapport à la cible sans référence spatiale (4,16 cm IC [4,02 ; 4,30] cf. Figure 4.15).

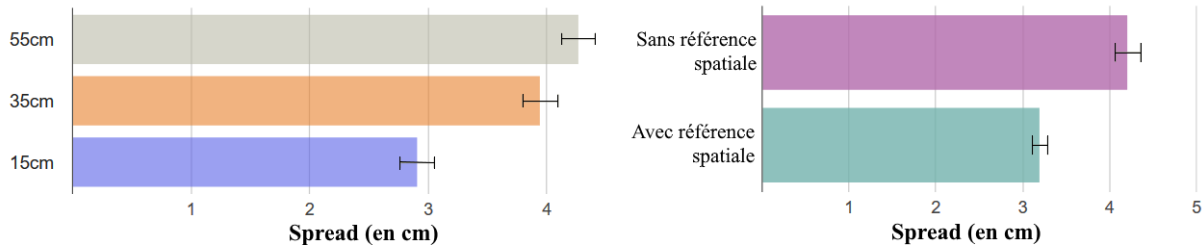


Figure 4.15 – Spread moyen en cm (avec 95% Cis) par distance cible (à gauche) et par référence spatiale (à droite)

La taille de la cible et la position de la tête de l'utilisateur ne semblent pas avoir d'impact sur le *spread*.

4.3.4 Résumé

Notre étude a révélé que la taille de la cible et sa distance par rapport à la vitrine ont un fort impact sur les performances de sélection. Les très petites cibles (1 cm) sont très difficiles à sélectionner et prennent beaucoup de temps à atteindre. L'utilisation de cibles plus grandes et leur placement plus près de la vitrine augmentent à la fois le taux de réussite et le temps de réalisation. Le fait de placer les cibles près d'une référence spatiale, par exemple près d'un bord de la vitrine, réduit l'*offset* et favorise les performances de sélection. Un résultat surprenant de cette étude est que le point de vue de l'utilisateur a un impact limité sur les performances de sélection.

Dans l'ensemble, nos résultats montrent que l'utilisation d'une vitrine comme surface tactile est une approche prometteuse, qui permet de sélectionner facilement des cibles situées entre 3 et 35 cm de la vitrine. Dans la section suivante, nous avons analysé les implications de ces résultats pour la conception d'un système d'interaction avec un référent physique utilisant le toucher sur une vitrine.

5 Discussion

Dans cette section, nous émettons un ensemble de recommandations de conception basées sur les résultats de notre étude ainsi que plusieurs perspectives d'évolutions et d'amélioration de notre approche TouchGlass.

5.1 Recommandations de conception

Nos recommandations ont pour but d'aider les concepteurs d'un système d'interaction avec un référent physique à sélectionner la taille, la distance, la position et la hauteur des cibles dans le contexte d'une sélection absolue sur une vitrine entourant un objet physique.

La taille de l'objet. Les concepteurs doivent privilégier les lancers de rayon orthogonaux à partir de la vitrine pour sélectionner des objets physiques placés derrière la vitrine et dont la taille est supérieure à 2 cm.

Distance de l'objet par rapport à la vitrine. Les objets de taille inférieure à 4 cm doivent être placés à moins de 35 cm de la vitrine. Seuls les objets d'une taille égale ou supérieure à 4 cm peuvent être placés plus loin, car leur taux de réussite moyen est toujours supérieur à 90 %.

Position des objets. Placer des objets près des bords de la vitrine (comme le bas) fournit une référence spatiale qui améliore la sélection. Une approche pourrait consister à placer les petits objets près d'une référence spatiale, et les plus gros objets, plus faciles à sélectionner, au milieu de la vitrine.

Hauteur de l'objet et de la vitrine. Les concepteurs d'interaction peuvent placer les objets à n'importe quelle hauteur, de la taille à la tête, puisque nos résultats n'ont pas révélé d'impact de la position de la tête de l'utilisateur par rapport à l'objet sur la précision de la sélection.

5.1.1 Amélioration de la sélection des cibles

Nos résultats soulignent que l'utilisation d'une vitrine comme surface tactile pour interagir avec des objets exposés est une approche prometteuse, qui peut être utilisée avec des cibles relativement petites placées jusqu'à 35 cm de la vitrine. Nous présentons ici différentes façons d'améliorer la sélection des cibles, ce qui pourrait être utile pour faciliter l'interaction avec des cibles encore plus petites et plus éloignées.

5.1.2 Taille du rayon

Nous avons choisi d'étudier la sélection d'un objet physique en lançant un rayon dont la taille (c'est-à-dire le diamètre) est de 0,2 mm. Étant donné que le décalage moyen par rapport à la cible lors de l'atterrissage varie de 0,6 à 0,8 cm en moyenne, l'utilisation d'un diamètre plus grand améliorera la sélection de la cible dans toutes les conditions. Nous avons refait l'analyse de nos données expérimentales avec deux tailles de rayon différentes (1 cm et 2 cm) : nous avons constaté que le taux de réussite sur les sélections au premier contact des cibles passe de 42% avec une taille de 0,2 mm à 60% avec une taille de 1 cm et à 73% avec une taille de 2 cm.

5.1.3 Augmenter la taille virtuelle de la cible

L'augmentation de la taille virtuelle (c'est-à-dire la zone qui peut être sélectionnée) des cibles est une approche bien connue (McGuffin & Balakrishnan, 2005, 2002) pour améliorer la performance de la sélection lorsque la taille réelle de la cible ne peut être modifiée. Notre analyse de l'*offset* et du *spread* montre que les premiers contacts avec la vitrine ont tendance à être situés au-dessus de la cible. L'extension de la taille virtuelle des cibles pourrait faciliter la sélection.

5.2 Perspectives

Au-delà de l'étude de l'impact d'une taille du rayon ou d'une taille virtuelle de cible différente, comme détaillé précédemment, une première perspective est de réaliser une étude longitudinale *in-situ* pour analyser le comportement des visiteurs avec un tel système dans un contexte public. Cette étude permettrait également d'évaluer comment et où présenter les informations liées au référent physique sélectionné à l'aide de la vitrine. Quand le référent physique est protégé par une vitrine, les données situées sont généralement affichées sur un écran séparé. Il serait alors pertinent d'étudier les différentes façons de présenter les données (sur un smartphone, sur un écran séparé ou à l'intérieur de la maquette). De plus, la démocratisation de certains dispositifs de réalité augmentée pourrait permettre de présenter les données numériques (cf. section 1 dans le chapitre État de l'art). Par exemple, après avoir sélectionné le référent physique, les données numériques pourraient s'afficher en l'air grâce à un casque de réalité augmentée.

Une deuxième perspective serait d'étendre notre approche pour permettre la sélection de cibles volumétriques. Cela pourrait être intéressant pour interagir avec des modèles architecturaux physiques, qui sont généralement placés derrière une vitrine. Par exemple, la maquette d'un terminal d'aéroport peut être utile aux passagers pour visualiser le chemin à suivre pour atteindre le terminal approprié. À cette fin, nous devons proposer et étudier des solutions pour étendre notre sélection actuelle en 2D à la 3D. Une solution consistera à adapter les techniques existantes de lancer de rayon pour les environnements immersifs (Grossman & Balakrishnan, 2006; Kopper et al., 2011; Vanacken et al., 2007).

6 Synthèse

Dans ce travail, nous avons présenté une étude du lancer de rayon depuis une vitrine pour sélectionner des objets physiques placés derrière elle. Cette étude a été motivée par une étude préliminaire qui nous a permis d'identifier la meilleure approche « *walk-up and use* » pour un

lancer de rayon à l'aide d'une vitrine interactive : un rayon orthogonal à la surface de la vitrine. Nous avons ensuite étudié ce type particulier de lancer de rayon dans une expérience de sélection d'objets physiques. L'objectif était d'évaluer l'utilisation d'une vitrine comme entrée tactile pour sélectionner des cibles physiques placées derrière la vitrine, en fonction de la taille de la cible, de sa distance, de la présence de références spatiales et de la position de la tête de l'utilisateur par rapport à la vitrine. Nos résultats expérimentaux révèlent que l'utilisation d'une vitrine comme surface tactile pour sélectionner des objets exposés derrière elle est une approche prometteuse qui peut être utilisée avec des cibles relativement petites placées jusqu'à 35 cm de distance de la vitrine. À partir de ces résultats, nous avons proposé plusieurs directives de conception pour l'interaction avec un objet physique derrière une vitrine.

Il résulte donc de ce chapitre, une technique de sélection d'un référent physique dans un contexte public. Nous avons déjà mentionné dans l'état de l'art que ces systèmes publics ne sont pas les seuls systèmes de visualisation de données situées. L'interaction avec ces systèmes peut dépendre des technologies utilisées ou des contextes d'utilisation. De plus cette étude se concentre uniquement sur la première phase d'une interaction avec un système support à la visualisation de données situées. En effet, interagir avec un système de visualisation de données situées nécessite une sélection sur le référent physique puis une manipulation ou au moins une exploration des données situées en lien avec la cible physique choisie. Au-delà des multiples solutions de conception de la phase de sélection, liées aux technologies utilisables et contextes, il convient donc de prolonger cette conception par une réflexion relative aux solutions d'exploration des données liées : le processus complet de développement de ces systèmes interactifs de visualisation de données situées est donc complexe, et leur évaluation en présence d'utilisateurs particulièrement difficile (variabilité des tâches, multiplicité des options, etc.). Une approche possible est de concevoir un modèle facilitant l'évaluation, la comparaison et la conception de ces systèmes. Nous présentons une extension d'un modèle prédictif dans le chapitre suivant.

Chapitre 5 KLM

Contenu du Chapitre

1 Notre extension KLM pour la réalité mixte.....	106
1.1 Nouveaux opérateurs pour l'interaction en réalité mixte	106
1.2 Adaptation d'un opérateur existant : Simple Reaction (Ms) et ses heuristiques	109
2 Études utilisateur : définition d'une unité de temps pour les nouveaux opérateurs	111
2.1 Participants et environnement	111
2.2 Collecte et analyse des données.....	112
2.3 Opérateur Button Click.....	112
2.4 Opérateur Raise Hand.....	113
2.5 Opérateur Air Tap.....	114
2.6 Coarse Head Point	115
2.7 Opérateur Precise Head Point.....	117
2.8 Résumé	119
3 Études de consolidation.....	119
3.1 Participants et environnement	120
3.2 Données recueillies et analyse des données.....	121
3.3 Étude de consolidation 1 : pointer hors du FoV	121
3.4 Étude de consolidation 2 : sélections multiples hors du FoV	123
3.5 Étude de consolidation 3 : sélections multiples dans le FoV	125
3.6 Résumé	126
4 Conclusions et Travaux Futurs.....	127

La visualisation de données situées a connu un nouvel essor avec l'avènement de la réalité mixte immersive basée sur les casques (ci-après dénommée réalité mixte). En effet, les casques de réalité mixte permettent de surmonter les limites des systèmes basés sur un dispositif tenu en main (p. ex. l'impossibilité de réaliser des gestes en l'air avec la main tenant le dispositif) et des systèmes basés sur de la vidéoprojection (p. ex. l'immobilité des dispositifs cf. chapitre État de l'art section 1.2). La visualisation des données situées se fait alors par le biais de l'écran présent sur le casque. De plus, ces casques permettent désormais d'enregistrer et modéliser en temps réel l'environnement dans lequel se trouve l'utilisateur et les objets présents. Il est alors aisé pour les concepteurs d'ancrer les données ou une interface dans leur environnement ou sur une maquette reconnue par le casque. De plus, ces casques sont équipés de dispositifs permettant de reconnaître des gestes en l'air ou des rotations de la tête pour sélectionner des éléments.

Par conséquent, le développement de systèmes de réalité mixte basés sur un casque nécessite de prendre en compte les effets de la disposition spatiale de l'interface pouvant être située en dehors du champ de vision de l'utilisateur (Field of View ou FoV), ainsi que sur l'interaction et en particulier les tâches de pointage et de validation. Un problème pour les chercheurs et les concepteurs est d'évaluer le temps nécessaire pour effectuer ces tâches de pointage et de validation. Le développement et le test d'applications de réalité mixte sur un casque sont encore techniquement fastidieux, et la combinaison habituelle de développement basée sur des prototypes et des études utilisateur peut se révéler inefficace dans le contexte de l'interaction en réalité mixte (Al-Megren et al., 2018). Pour faciliter l'exploration des possibilités de conception dans un tel contexte, nous présentons une extension d'un modèle prédictif bien connu : le Keystroke-Level Model (KLM) (Card et al., 1980). KLM permet aux concepteurs de prédire les temps d'interaction à partir d'interactions atomiques appelées opérateurs, et d'identifier les problèmes d'utilisabilité du point de vue des performances pendant la phase de conception plutôt que pendant la phase de développement (Bernhaupt et al., 2008). Bien que KLM ait fait l'objet de nombreuses extensions pour des domaines spécifiques (tels que les interactions sur smartphone cf. chapitre État de l'art section 3.3.1), peu d'attention a été accordée à la modélisation des interactions dans le contexte des environnements de réalité mixte (Al-Megren et al., 2018).

Pourtant, un tel contexte implique des comportements spécifiques tels que le déplacement de la tête pour pointer le bon élément d'information dans l'espace immersif, ou la détection de la main dans le champ de vision du casque pour effectuer une validation en l'air qui peut être détectée

et reconnue par le casque. L'originalité de notre travail est de se concentrer spécifiquement sur la conception d'une extension de KLM, qui sera utilisée pour modéliser l'interaction pour effectuer des tâches de réalité mixte telles que le pointage et la sélection d'objets numériques. À cette fin, nous proposons d'abord un ensemble constitué de nouveaux opérateurs et d'un opérateur existant dans la littérature pour modéliser les interactions atomiques nécessaires dans le contexte des environnements de réalité mixte. Pour les opérateurs nouvellement identifiés, nous mesurons expérimentalement leurs temps unitaires respectifs, c'est-à-dire le temps nécessaire pour effectuer l'action correspondante. Dans trois études finales, nous mesurons le temps nécessaire pour effectuer une tâche qui combine les opérateurs KLM nouvellement introduits, et nous comparons ce temps de réalisation mesuré au temps prédit calculé avec notre extension du modèle KLM. Ces études finales valident nos contributions puisque les différences entre les temps observés et prédits sont en moyenne inférieures à 5%, ce qui est conforme à l'approche KLM (Olson & Olson, 1990).

Nos contributions sont 1) l'identification d'opérateurs KLM pertinents pour décrire les interactions avec la réalité mixte, 2) la définition de temps unitaires appropriés pour ces opérateurs à travers deux sessions d'études utilisateur et 3) la consolidation expérimentale de notre modèle.

1 Notre extension KLM pour la réalité mixte

Dans cette section, nous présentons notre extension KLM pour l'interaction en réalité mixte. En particulier, nous visons à modéliser le clic du bouton, le lever de la main, un geste en l'air de validation (ci-après appelé « Air Tap »), le pointage gros grain avec la tête et le pointage fin grain avec la tête dans les environnements de réalité mixte. Ces tâches sont des composantes fondamentales de toute interaction en réalité mixte ; leur définition et leur caractérisation sont donc essentielles. Nous présentons d'abord les opérateurs nouvellement introduits avant de décrire un opérateur hérité que nous réutilisons de la littérature.

1.1 Nouveaux opérateurs pour l'interaction en réalité mixte

Pour étudier la validation dans les environnements de réalité mixte, nous avons considéré deux techniques différentes couramment utilisées dans de tels environnements : appuyer sur un bouton physique et effectuer un geste de validation en l'air. Ces deux techniques peuvent se décomposer en trois actions atomiques représentées par trois opérateurs : 1) une validation en appuyant sur un bouton physique sur un dispositif dédié est représentée avec l'opérateur Button

Click, et 2) une validation gestuelle impliquant deux actions atomiques effectuées séquentiellement et représentées par les opérateurs, Raise Hand et Air Tap. Comme nous l'avons déjà mentionné, bouger la tête pour pointer vers le bon élément d'information dans un espace immersif est également l'une des actions récurrentes requises dans un tel environnement. Nous avons donc étudié deux formes de pointage dans l'environnement de réalité mixte : le pointage gros grain avec la tête (Coarse Head Point) et le pointage fin grain avec la tête (Precise Head Point). Nous détaillons ces cinq opérateurs dans les sous-sections suivantes.

1.1.1 *Button Click (Bc)*

L'exécution de gestes répétitifs en l'air peut être fastidieuse. Plusieurs casques de réalité mixte (par exemple HoloLens ou Magic LeapOne) sont donc livrés avec un dispositif portable dédié pour effectuer des actions de validation. Notre opérateur Button Click (**Bc**) modélise l'action du clic avec un tel dispositif dédié (cf. Figure 5.2).



Figure 5.2 – Illustration de l'opérateur Button Click (**Bc**)

1.1.2 *Raise Hand (Rh)*

La première étape pour effectuer un geste de la main en l'air est de lever la main afin que le casque puisse détecter la main avec ses technologies embarquées et reconnaître les gestes.



Figure 5.1 – Illustration de l'opérateur Raise Hand (**Rh**)

L'opérateur Raise Hand (**Rh**) modélise le geste nécessaire pour que l'utilisateur fasse passer sa main d'une position de repos à une position en l'air détectée par le casque (cf. Figure 5.1).

1.1.3 Air Tap (*At*)

Sélectionner un objet virtuel ou cliquer sur un bouton virtuel sont deux actions majeures dans les environnements de réalité mixte pour sélectionner un objet ou activer une commande. De nombreux casques offrent la possibilité de valider un pointage en effectuant un geste en l'air. Notre opérateur Air Tap (**At**) modélise cette validation gestuelle en l'air. Dans ce travail, nous avons décidé de modéliser le geste intégré à HoloLens, qui consiste à abaisser l'index et à le relever, en raison de sa robustesse et de sa simplicité d'exécution (cf. Figure 5.3).

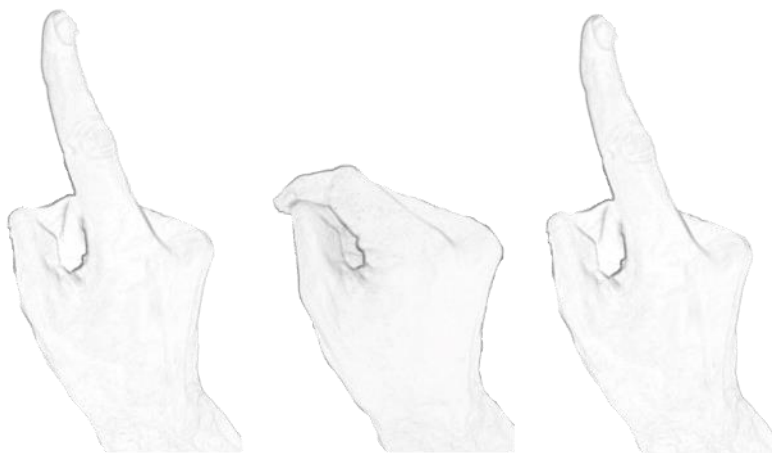


Figure 5.3 – Illustration de l'opérateur Air Tap (**At**)

1.1.4 Coarse Head Point (*Chp*)

Les dispositifs de réalité mixte ont un champ de vision (FoV) restreint par rapport au champ de vision humain. En général, les casques offrent un FoV de 30° pour HoloLens ou de 90° pour Meta2, alors que le FoV binoculaire humain est d'environ 120°. Par conséquent, avant de pouvoir lire ou sélectionner un élément, celui-ci doit être amené dans le champ de vision du



Figure 5.4 – Illustration de l'opérateur Coarse Head Point (**Chp**)

casque. C'est ce qui est représenté par l'opérateur Coarse Head Point (**Chp**) qui décrit donc le geste d'un utilisateur avec sa tête alors qu'il souhaite visualiser un objet en dehors du champ de vision du casque (cf. Figure 5.4). Cette action implique de larges mouvements de tête.

1.1.5 Precise Head Point (Php)

Une fois qu'un élément est visible dans le champ de vision, une sélection précise de cet élément, potentiellement parmi plusieurs éléments, doit être effectuée avant de pouvoir interagir avec lui. L'opérateur Precise Head Point (**Php**) modélise la tâche qui consiste pour un utilisateur à pointer un objet déjà visible dans le FoV du casque avec un curseur. Cet opérateur implique des mouvements précis de la tête pour amener le curseur d'un point de départ à un point d'arrivée (c'est-à-dire la cible). Le pointage fin grain avec la tête dans la plupart des casques est basé sur le champ de vision du casque (*Viewport-Based* (Chen et al., 2020)): un curseur est placé au centre du FoV et reste au centre lorsque les utilisateurs bougent la tête (cf. Figure 5.5).

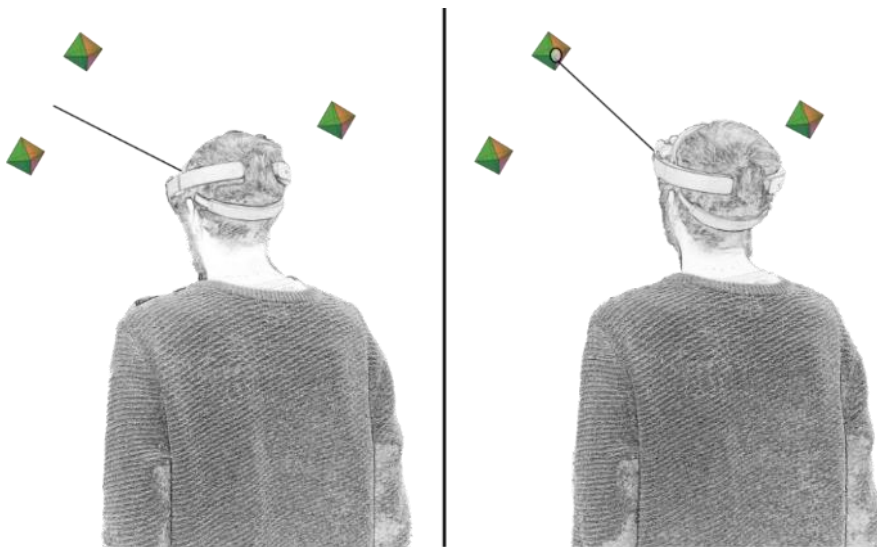


Figure 5.5 – Illustration de l'opérateur Precise Head Point (**Php**)

1.2 Adaptation d'un opérateur existant : Simple Reaction (Ms) et ses heuristiques

Initialement, Card et collab. (Card et al., 1980), ont modélisé un temps nécessaire entre 2 actions physiques atomiques différentes : l'opérateur Mental Act (**M**). Cet opérateur peut modéliser différentes tâches cognitives (c'est-à-dire la recherche visuelle, la planification d'actions futures, etc.). MacKenzie (MacKenzie, 2013) a affiné cet opérateur en le divisant en 5 opérateurs mentaux différents : Simple Reaction (**Ms**), Physical Matching (**Mp**), Name Matching (**Mn**), Class Matching (**Ml**) et Visual Search (**Mv**). Le temps unitaire de chaque opérateur a été mesuré expérimentalement. Nous avons choisi de nous appuyer sur eux dans notre extension de KLM,

en particulier sur l'opérateur de réaction (**Ms**) et son temps unitaire (277 ms) qui est le seul pertinent pour les tâches de pointage et de validation.

L'insertion des différents opérateurs mentaux dans une prédiction KLM est basée sur des heuristiques. Le KLM original proposé par Card et collab. (Card et al., 1980) proposent cinq heuristiques pour placer l'opérateur Mental Act (**M**) original (cf. chapitre État de l'art section 3.2). D'abord, on ajoute un opérateur **M** devant chaque opérateur physique (règle 0). Ensuite, on supprime les opérateurs **M** devant tous les opérateurs qui peuvent être entièrement anticipés (règle 1). Puis dans une unité cognitive, on supprime tous les opérateurs **M** sauf le premier (règle 2). Les deux règles suivantes concernent les interfaces de ligne de commande (règles 3 et 4) et ne sont pas pertinentes pour notre extension de KLM.

Kieras (Kieras, 2001) a proposé quelques directives pour améliorer l'utilisation de ces heuristiques originales. Parmi ces directives, les concepteurs devraient se concentrer sur le nombre d'opérateurs mentaux plutôt que sur leur emplacement. Cette directive spécifique est prise en compte dans les travaux existants sur les extensions de KLM (Erazo & Pino, 2018; Holleis et al., 2007) tandis que les autres sont spécifiques à un domaine et non pertinentes pour notre contexte. Sur la base de ces directives et des heuristiques originales, quatre règles guideront l'utilisation des opérateurs mentaux dans notre extension KLM :

- **Règle 0** : Placer des **Ms** devant chacun de nos opérateurs (Air Tap, Button Click, Raise Hand, Precise Head Point, Coarse Head Point).
- **Règle 1** : Si un opérateur qui suit un **Ms** est anticipé dans l'opérateur avant le **Ms**, supprimer le **Ms**. Par exemple, lorsque l'utilisateur doit pointer et sélectionner une cible, supprimer les **Ms** entre les opérateurs de pointage et de sélection.
- **Règle 2** : Si une suite de **Ms + Bc** ou de **Ms + At** appartient à une unité cognitive, supprimer tous les **Ms** sauf le premier. Par exemple, si l'utilisateur doit effectuer de multiples clics sur un bouton avec un dispositif spécifique **Bc** (ou un Air Tap **At**), supprimer tous les **Ms** entre chaque **Bc** (ou **At**).
- **Règle 3** : En cas de doute, insister davantage sur le nombre que sur l'emplacement des occurrences de l'opérateur **Ms**.

Ces cinq opérateurs (**At**, **Bc**, **Rh**, **Php** et **Chp**) et quatre règles constituent donc notre extension du modèle KLM. Dans la section suivante, nous définissons à l'aide d'études utilisateur l'unité de temps associée à chaque opérateur nouvellement introduit.

2 Études utilisateur : définition d'une unité de temps pour les nouveaux opérateurs

Après avoir présenté les nouveaux opérateurs de notre extension KLM, nous mesurons expérimentalement le temps unitaire de chacun d'entre eux, c'est-à-dire la durée prédite de l'action correspondante.

2.1 Participants et environnement

Nous avons réalisé deux sessions distinctes pour déterminer les temps unitaires des opérateurs : l'une pour l'opérateur Button Click, l'autre pour les opérateurs Air Tap, Raise Hand, Coarse Head Point et Precise Head Point. 12 participants ont pris part à la première session (3 femmes et 9 hommes âgés de 25 à 43 ans, moyenne 30,2). Parmi eux, 5 sont des doctorants ou des étudiants de premier ou deuxième cycle, 2 sont des post-doctorants, 3 sont des ingénieurs et 2 sont des enseignants-chercheurs dans notre laboratoire d'informatique. 12 participants ont pris part à la deuxième session (3 femmes et 9 hommes âgés de 23 à 32 ans, moyenne 27,4). Parmi eux, 10 sont des doctorants ou des étudiants de premier ou deuxième cycle et 2 sont des ingénieurs.

Pour ces deux sessions, nous avons utilisé un HoloLens casque v1 et son Clicker (c'est-à-dire le dispositif dédié à la sélection). Les expériences ont eu lieu dans un environnement contrôlé, sans lumière naturelle. Une vidéo illustrant chaque tâche a été présentée aux participants avant l'expérience. Étant donné que l'expérimentateur ne peut pas voir ce que les participants voient dans le casque, la vidéo permet de surmonter la difficulté d'expliquer précisément une tâche. Les participants pouvaient poser toutes leurs questions pendant la lecture de la vidéo. Après cette vidéo, les participants ont exécuté un tutoriel officiel de HoloLens intitulé « Learn Gestures », dans lequel ils ont appris à effectuer les gestes de base pour les interactions avec HoloLens : Raise Hand, Air Tap, et Precise Head Point.

Toutes ces études ont été réalisées de manière à ce que les mesures sanitaires imposées par le COVID-19 soient scrupuleusement appliquées.

2.2 Collecte et analyse des données

Toutes les mesures temporelles recueillies dans ces expériences sont capturées par le HoloLens et correspondent à des événements du système gérés par la Mixed Reality Toolkit (MRTK)¹⁶. Nous avons calculé des moyennes géométriques au lieu de moyennes arithmétiques, car elles conviennent mieux aux données ayant une distribution asymétrique positive comme le temps de réalisation d'une tâche (Ens et al., 2016), avec des intervalles de confiance (IC) de 95 %.

2.3 Opérateur Button Click

2.3.1 Tâche

Pour mesurer le temps nécessaire pour cliquer avec un dispositif dédié, nous avons demandé aux participants d'effectuer quatre clics consécutifs. Un retour visuel s'affichait entre chaque clic (**Bc**) pour indiquer le nombre de clics déjà effectués par les participants. Pour limiter les actions cognitives ou physiques des utilisateurs, les participants devaient regarder un minuteur affiché devant leurs yeux et attendre la fin d'un compte à rebours de 3s pour commencer l'essai. La même tâche était réalisée avec la main dominante et la main non dominante. Les participants ont d'abord suivi une session d'entraînement composée de cinq essais, avant de réaliser l'expérience qui comprenait 15 essais. Nous avons enregistré les temps de réalisation depuis la fin du compte à rebours jusqu'au premier clic et entre chaque clic. Nous avons recueilli 4 clics de bouton x 2 mains x 15 essais = 120 clics de bouton par participant.

2.3.2 Résultats

Dans une première analyse, nous n'avons trouvé aucune différence entre chaque main. Par conséquent, nous avons décidé de calculer la moyenne géométrique pour la main dominante et la main non dominante ensemble.

Le premier clic a été effectué en 481ms (IC [468 ; 494]). Ce premier clic a pris beaucoup plus de temps que les autres (2e clic : 205ms IC [199 ; 211] ; 3ème clic : 205ms IC [199 ; 210] ; 4ème clic : 213ms IC [209 ; 221]). Cela peut s'expliquer par le fait que, comme le souligne MacKenzie (MacKenzie, 2013), un temps de réaction est nécessaire afin que l'utilisateur réagisse après la fin du compte à rebours. Cette réaction (l'opérateur **Ms**) est en moyenne de 277ms. Si l'on soustrait ce temps de réaction au premier clic, le premier clic a pris 204ms IC [191 ; 218] (cf. Figure 5.6) ce qui est en accord avec le temps nécessaire pour effectuer les trois

¹⁶ <https://microsoft.github.io/MixedRealityToolkit-Unity/Documentation/GettingStartedWithTheMRTK.html>

autres clics. Selon nos heuristiques (règle 2 cf. section 1.2), il n'est pas nécessaire de soustraire le temps unitaire de l'opérateur Simple Reaction des autres clics.

Après avoir ajusté le temps mesuré pour le premier clic, nous avons calculé la moyenne géométrique de chaque clic pour la main dominante et non dominante. Cette moyenne correspond à l'unité de temps de l'opérateur Button Click (**Bc**) et est égale à 207 ms IC [202 ; 213].

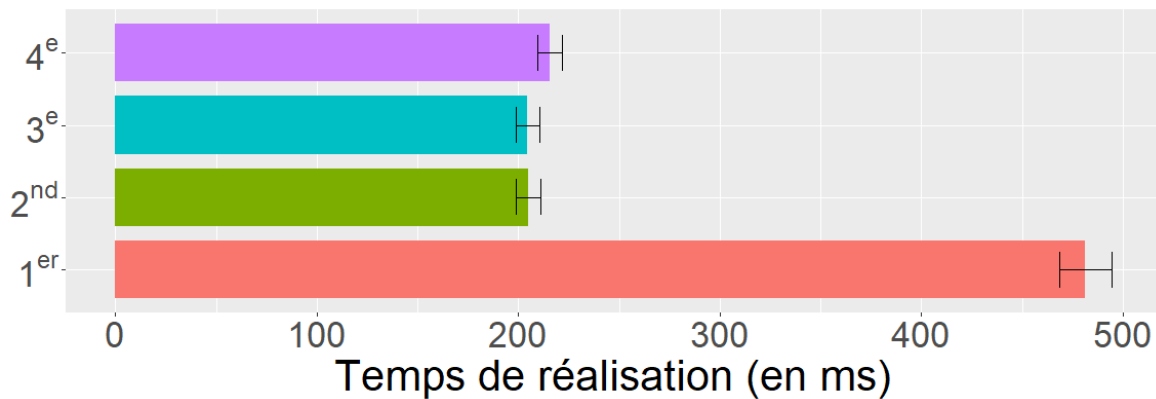


Figure 5.6 – Temps de réalisation des 4 Button Click (en ms) avec les IC à 95%

2.4 Opérateur Raise Hand

2.4.1 Tâche

Pour mesurer le temps nécessaire pour lever la main (**Rh**), avant d'effectuer le premier Air Tap (**At**), nous avons demandé aux participants d'amener leur main dominante dans le champ de vision du HoloLens. Avant de commencer chaque essai, les participants devaient placer leur main le long de leur cuisse. Pour limiter les actions cognitives ou physiques des utilisateurs, les participants devaient regarder un minuteur affiché devant leurs yeux et attendre le compte à rebours de 3 secondes pour commencer l'essai. En effet, la réalisation d'actions (cognitives ou physiques) autres que l'action correspondante à l'opérateur **Rh** pourrait entraîner un biais dans l'unité de temps mesurée. Les participants ont effectué 5 essais d'entraînement, suivis de 15 essais d'étude. Nous avons enregistré le temps correspondant à l'opérateur **Rh** (c'est-à-dire entre la fin du compte à rebours et la détection de la main par l'HoloLens). Nous avons recueilli 15 gestes Raise Hand par participant.

2.4.2 Résultats

Le temps moyen entre la fin du compte à rebours et le moment où le HoloLens détecte la main est de 899 ms (IC [861 ; 938]). Cependant, comme souligné par MacKenzie (MacKenzie, 2013), et déjà détaillé dans la section précédente, l'utilisateur prend un temps de réaction (**Ms**) pour réagir après la fin du compte à rebours. Par conséquent, nous soustrayons le temps de réaction

de 277 ms du temps moyen mesuré : le temps unitaire correspondant à l'opérateur de la main levée est de 623ms IC [586 ; 662].

2.5 Opérateur Air Tap

2.5.1 Tâche

En ce qui concerne l'opérateur Air Tap (At), les participants devaient effectuer quatre gestes Air Tap successifs après l'action Raise Hand. Un retour visuel s'affichait entre chaque geste pour indiquer le nombre de gestes Air Tap déjà effectués par le participant. Le temps enregistré pour un Air Tap correspond au temps nécessaire pour abaisser l'index et le lever à nouveau. Les participants ont effectué 5 essais d'entraînement, suivis de 15 essais d'étude. Nous avons recueilli 4 Air Tap x 15 essais = 60 Air Tap par participant

2.5.2 Résultats

Nous avons étudié le temps de réalisation en considérant chacun des quatre gestes Air Tap indépendamment. Comme l'illustre la Figure 5.7, nous observons que le premier geste Air Tap prend très nettement moins de temps que les trois suivants (300 ms, IC [269 ; 334]). Nous pensons que cette différence est due au fait que les participants ont commencé le premier Air Tap au moment où le geste correspondant à l'opérateur Raise Hand était effectué. Nous avons donc décidé d'exclure le premier geste Air Tap de notre calcul.

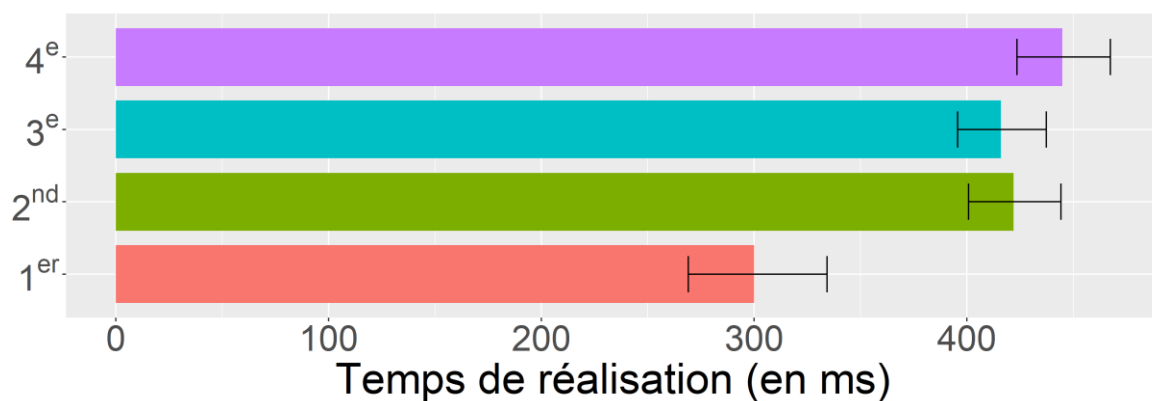


Figure 5.7 – Temps de réalisation des 4 Air Tap (en ms) avec les IC à 95%

Par conséquent, le temps moyen du Air Tap est de 427ms IC [408 ; 448]. À la différence des opérateurs **Bc** ou **Rh**, l'action d'effectuer les trois derniers Air Tap consécutifs ne nécessite pas de temps de réaction puisqu'ils faisaient suite à une action déjà engagée (règle 2 cf. section 1.2), il n'est donc pas nécessaire de soustraire le temps unitaire correspondant à l'opérateur Simple Reaction (**Ms**). Par conséquent, le temps unitaire caractérisant notre opérateur **At** est de 427ms IC [408 ; 448].

2.6 Opérateur Coarse Head Point

2.6.1 Tâche et protocole

Amener le FoV autour d'une cible pour la visualiser dans un environnement de réalité mixte consiste à déplacer (tourner) la tête jusqu'à ce que la cible entre dans le FoV du casque. Le temps nécessaire pour amener le FoV sur la cible dépend de la distance angulaire entre la cible et la position initiale du centre du FoV.

Positionnement des cibles. D'après Tilley et Dreyfuss (Tilley & Dreyfuss Associates, 1994), la rotation maximale confortable de la tête est de 60° (sans rotation des épaules et sans inclinaison du corps). Nous avons donc défini trois distances angulaires de la cible à l'extérieur du FoV du HoloLens (45° , 60° et 75°). De plus, comme souligné dans (Tilley & Dreyfuss Associates, 1994), il est important de distinguer les mouvements verticaux des mouvements horizontaux de la tête. Dans notre étude, nous avons donc positionné les cibles dans 8 directions correspondant aux combinaisons possibles de mouvements horizontaux et verticaux de la tête : deux verticaux (nord, sud), deux horizontaux (est, ouest), et quatre diagonaux (nord-est, nord-ouest, sud-est et sud-ouest, cf. Figure 5.8).

Enfin, et conformément aux directives du HoloLens¹⁷, des cibles de 7,5 cm ont été placées sur une sphère de 2 m de rayon centrée sur les yeux de l'utilisateur (soit une taille angulaire de 2,14 degrés). Chaque cible est placée dans l'une des huit directions possibles. En combinant les 3 distances angulaires et les 8 directions, on obtient 24 positions possibles de la cible autour du champ de vision.

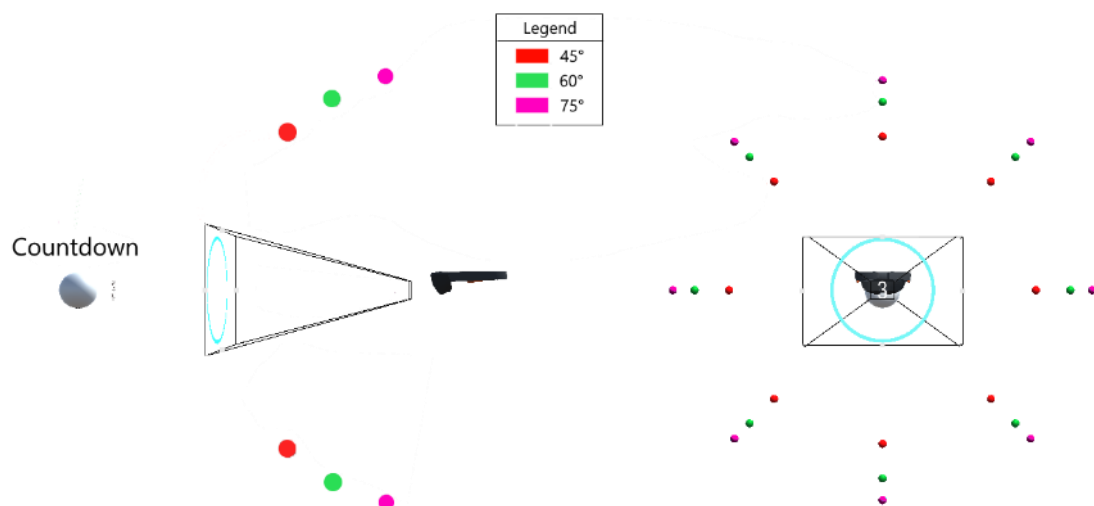


Figure 5.8 – Les différentes conditions de l'étude de l'opérateur Coarse Head Point (**Chp**) avec (en rouge) les cibles à 45° , (en vert) les cibles à 60° , (en rose) les cibles à 75° , (en blanc) le compteur permettant de débuter la tâche et (en bleu) le cercle inscrit dans le champ de vision de l'HoloLens

¹⁷ <https://docs.microsoft.com/en-us/windows/mixed-reality/gaze-and-commit-head>

Tâche. Pour commencer un essai, les participants devaient attendre un compte à rebours de 3 secondes, placé sur un objet affiché devant leurs yeux, en regardant droit devant eux. Ensuite, une flèche indiquait la direction à suivre pour trouver la cible, et donc amener le FoV autour de la cible.

Plus précisément, les participants devaient placer la cible à l'intérieur d'un cercle affiché au centre du FoV et inscrit dans le FoV, c'est-à-dire dont le diamètre correspondait exactement à la hauteur du FoV. Ce cercle inscrit dans le FoV permet d'éliminer tout effet dû à la différence entre la largeur et la hauteur du FoV. Le rayon de ce cercle (inscrit dans le FoV) a une distance angulaire de $8,75^\circ$, ce qui correspond à la moitié de la hauteur du FoV du HoloLens ($17,5^\circ$) et se situe dans la plage que peut percevoir l'œil humain (c'est-à-dire entre 5° et 10° (Dreyfuss Associates & Wilcox, 2001; Kress, 2019)). L'essai se terminait automatiquement lorsque la cible était restée dans ce cercle pendant 500 ms. Les cibles sont affichées l'une après l'autre.

Design expérimental. Cette partie de l'étude suit un plan intra-sujet 3×8 avec comme facteurs la distance angulaire (45° , 60° , 75°) et la direction (2 horizontales, 2 verticales, 4 diagonales). Chaque bloc correspondait à une distance angulaire et comprenait huit cibles (une dans chaque direction) présentées au hasard au participant. Les blocs étaient contrebalancés entre les participants. Chaque bloc a été répété trois fois, le premier servant d'entraînement. Au total, nous avons recueilli 3 distances angulaires \times 8 directions \times 2 répétitions = 48 essais par participant.

Mesures. Le temps de réalisation a été mesuré entre le moment où le curseur du HoloLens quittait l'objet sur lequel le compte à rebours était affiché et le moment où la cible entrait dans le cercle de vision (c'est-à-dire que nous avons supprimé les 500 ms de validation de l'essai).

2.6.2 Résultats

Premièrement, nous avons observé que la distance angulaire a un impact clair sur le temps nécessaire pour amener le cercle inscrit dans le FoV autour de la cible : le temps augmente clairement avec la distance angulaire (45° : 561ms IC [518 ; 607] ; 60° : 852ms IC [796 ; 913] ; 75° : 1044ms IC [969 ; 1125]). La direction a également un impact sur ce temps de réalisation (986ms IC [907 ; 1072] pour les mouvements verticaux, 613ms IC [560 ; 671] pour les mouvements horizontaux et 810ms IC [757 ; 866] pour les mouvements diagonaux). La Figure 5.9 illustre ces résultats et montre également que ces résultats restent vrais lorsque l'on considère ces deux facteurs simultanément.

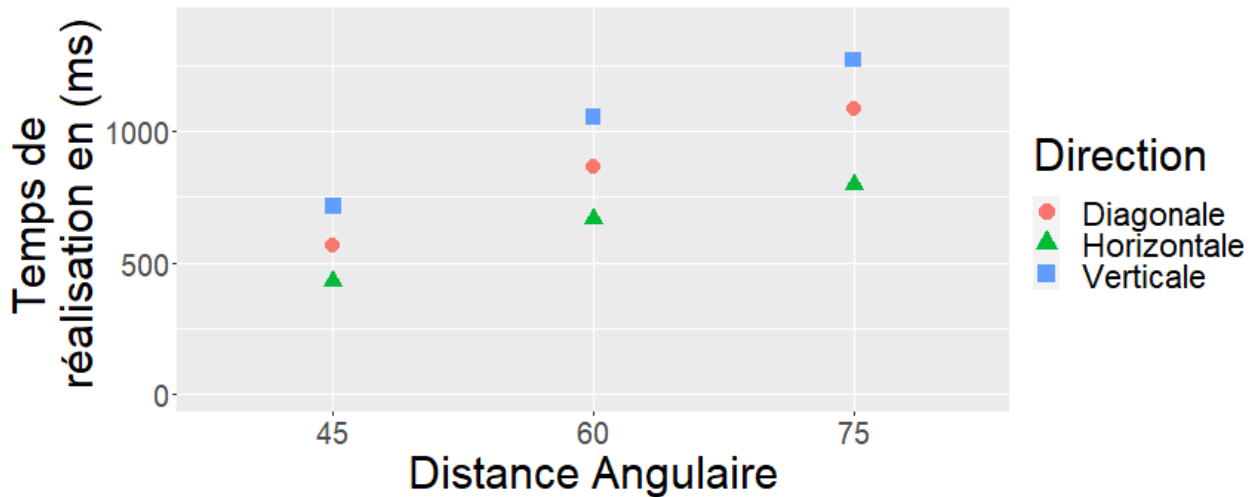


Figure 5.9 – Les temps de réalisation pour la tâche de *Coarse Head Point* avec (carrés bleus) la direction verticale, (triangles verts) la direction horizontale et (cercles rouges) la direction diagonale

À partir de ces résultats, nous avons extrait 9 moyennes géométriques, résumées dans le Tableau 3 et qui peuvent être utilisées comme unité de temps pour l'opérateur Coarse Head Point. Ces valeurs dépendent de la direction (d) et de la distance angulaire (a). Le temps de réaction potentiel de l'utilisateur (opérateur Simple Reaction **Ms**) n'est pas inclus dans ces valeurs, car la mesure du temps de réalisation commence lorsque le curseur du regard quitte l'objet sur lequel le décompte est affiché, après que l'utilisateur a déjà réagi.

Tableau 3 – Temps de réalisation de la tâche *Coarse Head Point* (en ms) avec les IC à 95% selon la direction (d) et la distance angulaire (a) : **Chp(d,a)**

	Horizontale	Diagonale	Verticale
45	431 IC [365 ; 510]	567 IC[510 ; 629]	715 IC[617 ; 829]
60	668 IC [595 ; 752]	865 IC[783 ; 957]	1054 IC[936 ; 1187]
75	797 IC [695 ; 914]	1083 IC[970 ; 1209]	1270 IC[1121 ; 1441]

2.7 Opérateur Precise Head Point

2.7.1 Tâche

Cette étude vise à mesurer le temps nécessaire pour pointer du regard une cible déjà présente dans le champ de vision, une action correspondant à notre dernier opérateur : Precise Head Point (**Php**). Nous avons positionné huit cibles sur le cercle inscrit dans le FoV tel que défini pour la partie précédente de l'étude, selon les huit directions cardinales (cf. Figure 5.10).

Pour commencer un essai, les participants devaient attendre la fin d'un compte à rebours de 3s affiché droit devant eux, comme dans la partie précédente de l'étude. Ensuite, ils devaient

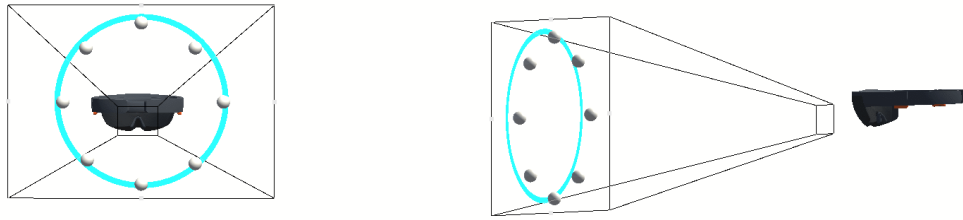


Figure 5.10 – Configuration pour la tâche Precise Head Point avec le champ de vision de l’HoloLens et les cibles (en blanc)

amener le curseur du HoloLens sur la cible et rester 500ms pour valider l’essai. Le temps de réalisation était mesuré entre le moment où le curseur du regard du HoloLens quittait la zone où le compte à rebours était affiché et le moment où il entrait dans la cible, évitant ainsi les 500ms de validation de l’essai.

2.7.2 Résultats

Contrairement à l’action correspondant à l’opérateur Coarse Head Point (**Chp**), les résultats concernant l’opérateur Precise Head Point (**Php**) ne mettent pas en évidence un impact de la direction sur les temps de réalisation requis pour pointer une cible déjà présente dans le FoV (Horizontal = 452ms, IC [369 ; 554] ; Diagonal = 439ms IC [378 ; 506] ; Vertical = 358ms IC [299 ; 429]) cf. Figure 5.11). Pour la même raison qu’avec l’opérateur Coarse Head Point, il n’est pas nécessaire de soustraire un temps de réaction. L’unité de temps retenu est alors la moyenne géométrique de tous les essais soit 419ms [408 ; 448].

En outre, nous avons recueilli le nombre de fois où le curseur est entré et sorti de la cible. Sur les 192 essais, 105 ont été réalisés en ne sortant pas de la cible (54%), 76 ont été réalisés en sortant une fois de la cible (39%) et seulement 11 (5,7%) ont nécessité plus de deux entrées sur la cible avant la sélection finale.

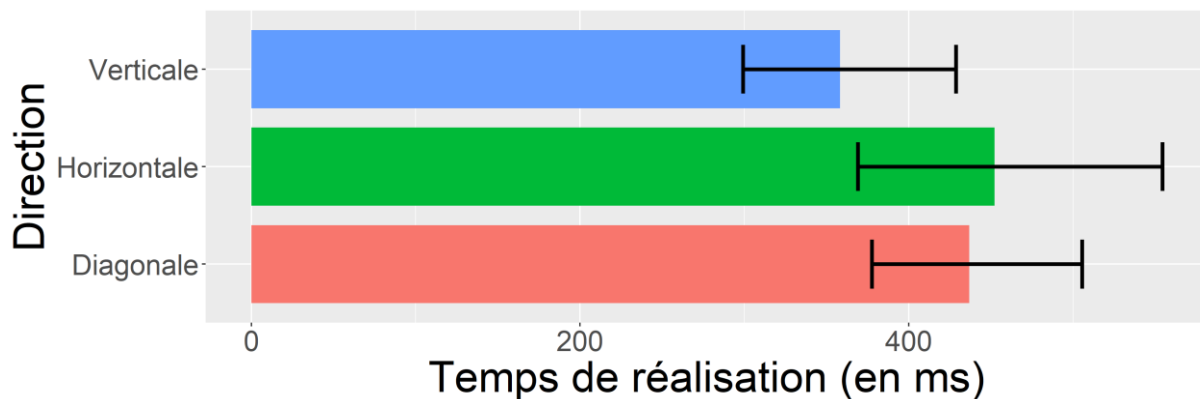


Figure 5.11 – Les temps de réalisation de la tâche de Precise Head Point avec les IC à 95%

2.8 Résumé

Au travers de ces différentes expériences, nous avons défini des valeurs unitaires de temps pour les nouveaux opérateurs de notre extension KLM pour l'interaction en réalité mixte : Button Click (207 ms), Raise Hand (623 ms), Air Tap (427 ms), Coarse Head Point (9 valeurs selon la distance angulaire et la direction), et Precise Head Point (419 ms). Tous les gestes et leurs temps unitaires associés sont résumés dans le Tableau 4.

Ces opérateurs et leurs temps unitaires associés doivent maintenant être consolidés dans des tâches écologiques. À cette fin, trois études utilisateur ont été conçues pour confronter notre modèle à différentes combinaisons d'opérateurs.

Tableau 4 – Notre Extension de KLM

Nom	Unité de Temps (en ms) et IC à 95%	Origine
Simple Reaction (Ms)	277	(MacKenzie, 2013)
Button Click (Bc)	207 IC [202 ; 213]	Nouvellement introduit
Raise Hand (Rh)	623 IC [586 ; 662]	Nouvellement introduit
Air Tap (At)	427 IC [408 ; 448]	Nouvellement introduit
Coarse Head Point (Chp)	De 431 à 1270 cf. Tableau 3	Nouvellement introduit
Precise Head Point (Php)	419 IC [369 ; 554]	Nouvellement introduit

3 Études de consolidation

Les trois études de consolidation correspondent à différentes combinaisons d'opérateurs nécessaires à l'exécution de différentes tâches courantes dans les applications de réalité mixte :

1. Une tâche de pointage où les cibles sont placées à l'extérieur du champ de vision initial
2. Une tâche de sélection multiple où les cibles sont placées à l'extérieur du champ de vision initial
3. Une tâche de sélection multiple où les cibles sont placées à l'intérieur du champ de vision initial.

Les opérateurs impliqués dans chaque tâche sont résumés dans le Tableau 5. Chacune des trois études de consolidation consiste à comparer un temps de réalisation de la tâche, mesuré expérimentalement, avec un temps prédit de la tâche basé sur l'utilisation des temps unitaires associés aux opérateurs de notre extension de KLM.

Tableau 5 – Les opérateurs impliqués dans les 3 études de consolidation

Opérateurs	Pointer hors du FoV	Sélections multiples hors du FoV	Sélections multiples dans le FoV
Button Click (Bc)		X	
Raise Hand (Rh)			X
Air Tap (At)			X
Coarse Head Point (Chp)	X	X	
Precise Head Point (Php)	X	X	X
Simple Reaction (Ms)	X	X	X

3.1 Participants et environnement

Ces trois études de consolidation se sont déroulées en deux sessions : la première concernait uniquement la tâche de pointage (étude 1) et la seconde concernait les deux tâches de sélections multiples (études 2 et 3).

Douze participants (3 femmes et 9 hommes âgés de 25 à 43 ans, moyenne 30,2) ont pris part à la première étude de consolidation. Parmi eux, 5 sont des doctorants ou des étudiants de premier ou deuxième cycle, 2 sont des post-doctorants, 3 sont des ingénieurs et 2 sont des enseignants-chercheurs dans notre laboratoire. Tous les participants de cette première étude étaient des participants de la première session d'études sur les opérateurs. Douze participants (3 femmes et 9 hommes âgés de 26 à 51 ans, moyenne 31,2) ont effectué les études de consolidation 2 et 3. Parmi eux, 8 sont des doctorants, 2 des post-doctorants et 2 des ingénieurs. 3 d'entre eux n'ont participé à aucune étude sur les opérateurs, 3 ont participé à la première session d'études sur les opérateurs, 4 ont participé à la deuxième session et 2 ont participé aux deux sessions.

Pour ces trois études, nous avons utilisé un HoloLens casque v1 et le Clicker (c'est-à-dire le dispositif dédié à la sélection). Les expériences se sont déroulées dans un environnement contrôlé, sans lumière naturelle. Une vidéo illustrant chaque tâche était présentée aux participants avant l'expérience. Au début de la session, les participants ont effectué un tutoriel officiel Microsoft HoloLens intitulé « Learn Gestures ».

Toutes ces études ont été réalisées de manière à ce que les mesures sanitaires imposées par le COVID-19 soient scrupuleusement appliquées. Tous les scripts sont disponibles en ligne¹⁸.

¹⁸ https://osf.io/pjyhk/?view_only=bb36af530d71431094c4c76eae7ed1b1

3.2 Données recueillies et analyse des données

Toutes les mesures temporelles recueillies dans ces expériences sont capturées par le HoloLens et correspondent à des événements système gérés par la Mixed Reality Toolkit (MRTK)¹⁹. Nous avons calculé les moyennes géométriques de chaque essai pour chaque participant (Ens et al., 2016). Comme pour les extensions précédentes de KLM, nous avons considéré que notre modèle était validé si la différence entre les temps prédits et les temps observés était inférieure au critère d'erreur de 20% d'Olson et Olson (Olson & Olson, 1990).

3.3 Étude de consolidation 1 : pointer hors du FoV

Une tâche très courante dans un environnement de réalité mixte consiste à pointer une cible qui se trouve initialement en dehors du champ de vision. Par exemple, dans une tâche de maintenance augmentée ou une opération chirurgicale en réalité mixte, l'utilisateur (p. ex. le technicien ou le chirurgien) peut pointer et sélectionner des commandes dans une application en réalité mixte affichée hors de son champ de vision. Avec un casque de réalité mixte le pointage, peut être effectué avec les mouvements de la tête. Nous avons décidé d'implémenter la validation avec un temps d'arrêt de 500 ms (p. ex. comme la technique de validation utilisée par Bailly et collab. (Bailly et al., 2019)). Cette tâche est donc composée d'une combinaison d'une tâche de pointage gros grain avec la tête, notre opérateur **Chp**, pour amener le FoV autour de la cible, et d'une tâche de pointage fin grain avec la tête, notre opérateur **Php**, pour sélectionner la cible qui est maintenant dans le FoV, pour laquelle nous avons défini expérimentalement les temps unitaires dans les études décrites dans la section 2.

3.3.1 Plan d'étude et tâche

La tâche s'est déroulée en deux étapes : 1) rechercher la cible dans la direction indiquée par une flèche et la placer à l'intérieur du cercle inscrit dans le champ de vision (opérateur **Chp**) et 2) pointer la cible avec le curseur du HoloLens (opérateur **Php**). Nous avons utilisé le même protocole que pour la tâche de Coarse Head Point (section 2.6) avec 24 cibles (huit cibles placées à des distances triangulaires). Au total, nous avons recueilli 3 distances angulaires x 8 directions x 2 répétitions = 48 essais par participant.

3.3.2 Modélisation de la tâche

L'exécution de la tâche implique deux opérateurs, **Chp** et **Php**. De plus, le modèle KLM propose des heuristiques pour placer des opérateurs Simple Reaction entre deux actions

¹⁹ <https://microsoft.github.io/MixedRealityToolkit-Unity/Documentation/GettingStartedWithTheMRTK.html>

physiques. Dans notre modélisation de la tâche et selon notre heuristique (règle 0), un opérateur Simple Reaction doit être ajouté avant les opérateurs **Chp** et **Php**. Or le temps de réalisation de la tâche n'est calculé qu'à partir du moment où l'utilisateur commence à faire son geste représenté par l'opérateur Coarse Head Point. Le premier temps de réaction (**Ms**) de l'utilisateur n'est donc pas contenu dans le temps de réalisation et donc dans notre modélisation (cf. section 2.6).

Dans l'étude précédente, nous avons trouvé que **Chp** dépend de la distance angulaire de la cible et de la direction du mouvement de la tête. Nous avons défini neuf temps unitaires différents pour **Chp(d,a)**, où d est la direction (Horizontale, Verticale ou Diagonale) et a est la distance angulaire (45°, 60° ou 75°). Ainsi, neuf temps prédits (pt), un pour chaque combinaison de direction et de distance angulaire, ont été calculés avec notre extension KLM (cf. section 2.6.2) : $pt(d,a) = \mathbf{Chp}(d,a) + \mathbf{Ms} + \mathbf{Php}$. Par exemple, pour une cible placée sur la diagonale et à une distance angulaire de 60°, le temps prédit est $pt(\text{diagonale}, 60^\circ) = 865 + 277 + 419 = 1561$ ms.

Tableau 6 – Temps de réalisation en ms (TR) avec les IC à 95 %, le temps prédit par notre modèle (TP) et la différence (en ms et en pourcentage) entre le temps moyen de réalisation et le temps prédit par le modèle

	Horizontal			Diagonal			Vertical		
	TR	TP	Diff	TR	TP	Diff	TR	TP	Diff
45	1155 [1040 ;1283]	1127	+28 (2,5%)	1299 [1218 ;1385]	1263	+36 (2,8%)	1491 [1295 ;1716]	1411	+80 (5,4%)
60	1229 [1109 ;1360]	1364	-135 (-10%)	1567 [1462 ;1680]	1561	+6 (0%)	1601 [1449 ;1768]	1750	-149 (-9,3%)
75	1440 [1303 ;1592]	1493	-53 (-3,7%)	1796 [1665 ;1938]	1779	+17 (1%)	1924 [1761 ;2102]	1966	-42 (2,1%)

3.3.3 Comparaison entre les résultats de temps mesurés expérimentalement et prédits.

Dans le Tableau 6, nous rapportons le temps moyen de réalisation de la tâche avec des IC à 95 % pour chacune des 3 directions et distances angulaires. Dans chaque cellule, nous calculons la différence (en ms et en pourcentage) entre le temps prédit et les moyennes des temps observés.

Pour les neuf prédictions, le pourcentage de différence varie de 0% (6ms de différence pour l'horizontale 60) à 10% (135ms de différence pour la diagonale 60) (cf. Tableau 6). En

moyenne, nous avons trouvé une erreur de 4%. Tous ces résultats consolident notre modèle pour la séquence des neuf temps unitaires **Chp**, **Ms** et **Php**.

3.4 Étude de consolidation 2 : sélections multiples hors du FoV

La sélection de différentes cibles est une autre tâche très courante dans les applications de réalité mixte. Par exemple, un modèle 3D complexe peut être affiché à l'intérieur du FoV et différentes commandes utiles pour modifier, mettre à jour ou manipuler le modèle sont disponibles sur une interface séparée en dehors du FoV initial. Dans un tel contexte, l'utilisateur doit sélectionner le modèle 3D ou une partie de ce modèle 3D, rechercher les commandes en regardant à l'extérieur de son FoV initial, et enfin sélectionner l'une d'entre elles avant d'effectuer à nouveau ces deux étapes afin de modifier ou d'annoter complètement le modèle 3D. Par rapport à la méthode consistant simplement à « pointer dans le champ de vision » (cf. section 3.3), cette étude inclut la validation finale à chaque étape réalisée à l'aide d'un Button Click, et augmente sa validité externe car elle considère une séquence d'actions (recherche d'une cible en dehors du champ de vision et pointage de la cible), au lieu d'une seule action isolée.

3.4.1 Conception et tâche de l'étude

Dans cette étude, la tâche reflète le scénario décrit ci-dessus. Concrètement, la tâche commençait par la sélection d'une première cible affichée au centre du champ de vision (cf. Figure 5.12 à gauche), puis la recherche d'une cible placée à 45° dans une direction horizontale (Coarse Head Point **Chp**(horizontal,45)), la sélection de la cible et le retour à la première cible pour la sélectionner (cf. Figure 5.12, à droite), et ainsi de suite. La sélection se faisait par l'utilisation séquentielle d'un Coarse Head Point **Chp**(horizontal,45), d'un Simple Reaction (**Ms**), d'un Precise Head Point (**Php**) et d'une étape de validation basée sur un clic de bouton (**Bc**).

Pendant la tâche, une flèche placée au-dessus des cibles indiquait la direction de la cible suivante. Comme dans les études précédentes, les participants devaient regarder un minuteur

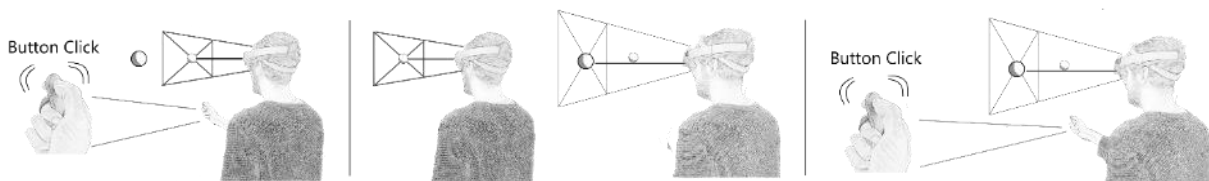


Figure 5.12 – Configuration d'un essai pour la tâche de sélection en dehors du champ de vision avec (à gauche) l'utilisateur sélectionne (**Bc**) la première cible placée au centre du champ visuel initial, (au centre) l'utilisateur tourne la tête de 45° vers la gauche (**Chp**) pour pointer la deuxième cible (**Php**) et (à droite) l'utilisateur sélectionne la cible (**Bc**) avec le dispositif dédié

affiché sur la première cible devant leurs yeux et attendre le compte à rebours de 3s pour commencer l'essai.

Les cibles mesuraient 7,5 cm (taille angulaire = 2,14 degrés) et étaient placées à 2 mètres de l'utilisateur (c'est à dire sur une sphère de 2m de rayon dont l'utilisateur était le centre). Les cibles que l'utilisateur devait sélectionner en dehors du champ de vision initial étaient placées de façon aléatoire à gauche ou à droite de ce champ (c'est-à-dire direction ouest ou est). L'essai se terminait lorsque la dernière cible était sélectionnée. Les participants ont effectué 6 essais pour l'entraînement, suivis de 18 essais pour l'étude. Comme nous avons déjà établi la validité du temps unitaire défini pour la séquence des opérateurs Precise Head Point et Coarse Head Point dans n'importe quelle direction et distance angulaire, nous avons décidé de ne considérer qu'une seule direction et distance angulaire pour l'opérateur Coarse Head Point dans cette étude.

Cette tâche de sélection de cinq cibles nécessite de nombreuses répétitions de Coarse Head Point (**Chp**(horizontal,45)), Precise Head Point (**Php**) et Button Click (**Bc**). La première sélection ne comprend qu'un clic de bouton (**Bc**) et les quatre suivantes comprennent (dans cet ordre) : 1) **Chp** 2) **Php** et 3) **Bc**. Selon la règle 0 de l'heuristique de la section 1.2, un opérateur de réaction (**Ms**) est ajouté avant chaque opérateur. Ensuite, l'utilisation de la règle 1 induit la suppression de tous les **Ms** qui peuvent être anticipés. En effet, les participants anticipent à la fois le clic du bouton lors du pointage fin grain avec la tête et le pointage gros grain avec la tête pour chercher la cible suivante après la sélection d'une cible (Button Click). Ainsi, chaque **Ms** entre **Php** et **Bc** et entre le **Bc** et **Chp**, doit être retiré de notre modélisation. La première sélection est modélisée à l'aide d'un opérateur de réaction (**Ms**) et d'une validation du clic du bouton (**Bc**). Ensuite, les quatre sélections suivantes sont modélisées avec un **Chp**(horizontal, 45), un **Ms**, un **Php**, et un **Bc**.

Ainsi, le temps prédit est : $1 \times (\mathbf{Ms} + \mathbf{Bc}) + 4 \times (\mathbf{Chp}(\text{horizontal}, 45) + \mathbf{Ms} + \mathbf{Php} + \mathbf{Bc}) = 1 \times (277 + 207) + 4 \times (431 + 277 + 419 + 207) = 5\,820\text{ms}$.

3.4.2 Comparaison entre les temps observés et prédits

Comme détaillé ci-dessus, le temps prédit pour l'ensemble de l'essai est de 5 820ms. La moyenne géométrique de tous les temps de réalisation observés est de 5 783ms (IC [5 641 ; 5928]). Ainsi, la différence entre les temps observés et le temps prédit est de 37 ms (c'est-à-dire moins de 1% d'erreur entre les temps prédits et observés). Ce résultat valide fortement les différents éléments de notre modèle dans le contexte d'une tâche impliquant une sélection multiple de cibles placées en dehors du FoV.

3.5 Étude de consolidation 3 : sélections multiples dans le FoV

La plupart des applications de réalité mixte affichent plusieurs objets en même temps. Ainsi, la sélection de différents objets dans le FoV est une troisième tâche assez courante dans ces applications. Par exemple, dans le tutoriel « Learn Gestures » intégré à l'HoloLens, il faut sélectionner des hologrammes pour apprendre le geste Air Tap. Dans cette troisième étude de consolidation, nous visons à valider notre modèle lorsque l'interaction implique une tâche de sélection multiple de cibles affichées exclusivement à l'intérieur du FoV. Une autre différence avec l'étude de consolidation 2 est que nous explorons l'utilisation du Raise Hand et Air Tap au lieu du Button Click : les temps unitaires supplémentaires établis dans cet article sont donc confrontés à des situations interactives écologiques.

3.5.1 Conception de l'étude et tâche

La tâche consistait à sélectionner cinq cibles successives placées à une distance angulaire de $8,75^\circ$ les unes des autres et de gauche à droite. Pour éviter tout effet d'apprentissage, un essai est composé de deux cibles placées au nord-est de la cible précédente et de deux cibles placées au sud-est (cf. Figure 5.13)

Pour sélectionner une cible, les participants devaient apporter le curseur du regard à l'intérieur de la cible (**Php**) et effectuer un Air Tap (**At**). Au début de l'essai, seule la première cible était visible. La cible suivante apparaît à droite de la cible actuelle dès que cette dernière est sélectionnée. Lorsque l'utilisateur sélectionne la dernière cible, le compte à rebours réapparaît pour lancer un nouvel essai. Pendant le compte à rebours, nous avons demandé aux participants de baisser leur bras afin d'éviter la fatigue.

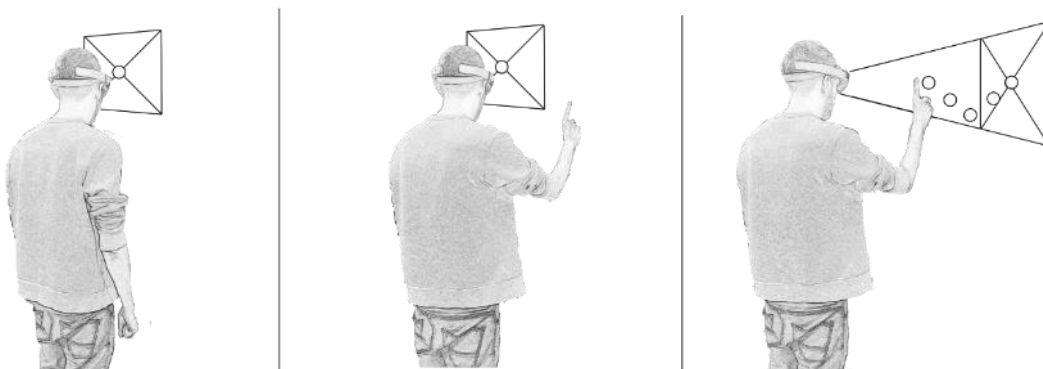


Figure 5.13 – Configuration d'un essai de sélection à l'intérieur du champ visuel avec (à gauche) le participant pointe la première cible avec l'HoloLens, (au centre) le participant lève sa main et son doigt pour effectuer la validation Air Tap et (à droite) la sélection de la cinquième (dernière) cible et (à droite) la sélection de la cinquième (la dernière) cible. L'essai comprend deux cibles placées au sud-est (cible numéro deux et cible numéro trois) et deux cibles placées au sud-est (cible numéro trois)

Les cibles de 7,5 cm (taille angulaire = 2,14 degrés) étaient affichées à 2 mètres de l'utilisateur. Les participants ont effectué 6 essais pour l'entraînement et ont répété trois fois les 6 essais pour l'étude. Ainsi, $6 \times 3 = 18$ temps de réalisation observés ont été recueillis par participant.

Nous avons recueilli les temps observés depuis le premier pointage (c'est-à-dire lorsque le curseur du regard se dirige vers la cible) jusqu'à la sélection de la dernière cible.

3.5.2 Modélisation des tâches

Nous modélisons la première sélection avec un Raise Hand (**Rh**) et un Air Tap (**At**) car nous enregistrons le moment où le curseur entre dans la première cible. Les quatre sélections suivantes ont été modélisées avec un Precise Head Point (**Php**) et une validation à l'aide d'un Air Tap (**At**). Il n'y a pas de Raise Hand parce que la main reste dans le FoV entre deux cibles. Ensuite, conformément à notre règle 0 (cf. section 1.2), nous ajoutons un opérateur **Ms** devant chaque opérateur **Rh**, **At** et **Php**. Selon notre heuristique (règle 1), nous supprimons les **Ms** entre les opérateurs **Rh** et **At** (première sélection) et entre les opérateurs **Php** et **At** (quatre dernières sélections), car l'action d'Air Tap (**At**) peut être entièrement anticipée pendant les gestes Raise Hand (**Rh**) et Precise Head Point (**Php**).

Ainsi, le temps prédit pour cette tâche est : $1 \times (\mathbf{Rh} + \mathbf{At}) + 4 \times (\mathbf{Ms} + \mathbf{Php} + \mathbf{At}) = 1 \times (623 + 427) + 4 \times (277 + 419 + 427) = 5\,542\text{ms}$.

3.5.3 Comparaison entre les temps observés et prédits

Comme détaillé ci-dessus, le temps prédit pour l'ensemble de l'essai est de 5 542ms. La moyenne géométrique de tous les temps observés est de 5 271 ms (IC [5101 ; 5446]). Ainsi, la différence entre le temps moyen observé et le temps prédit est de 271 ms (soit une différence de 4,8 % entre le temps prédit et le temps moyen observé). Ce résultat conforte notre combinaison de différents opérateurs pour pointer et sélectionner des objets affichés dans le champ de vision.

3.6 Résumé

Nous avons consolidé notre modèle à travers trois tâches écologiques différentes : 1) une tâche de pointage à l'extérieur du champ de vision, 2) une tâche de sélection à l'extérieur du champ de vision et 3) une tâche de sélection à l'intérieur du champ de vision. Tous les temps observés varient de 1 à 4,8% des temps prédits calculés avec notre modèle. Ainsi, notre modèle semble être une approche intéressante pour modéliser les temps prédits des tâches fondamentales dans les applications de réalité mixte.

4 Conclusions et Travaux Futurs

Dans ce chapitre, nous avons présenté une nouvelle extension de KLM (Card et al., 1980) pour les applications de réalité mixte modélisant les tâches de pointage et de validation en l'air ou à l'aide d'un dispositif dédié. Tout d'abord, nous avons identifié et décrit les nouveaux opérateurs nécessaires à la modélisation des tâches de pointage et de validation en réalité mixte : Precise Head Point (**Php**), Coarse Head Point (**Chp**), Raise Hand (**Rh**), Air Tap (**At**) et Button Click (**Bc**). Grâce à un ensemble d'études utilisateur, nous avons défini expérimentalement un ou plusieurs temps unitaires pour chaque opérateur. L'utilisation de KLM est basée sur l'identification et la combinaison d'opérateurs correspondant à des étapes d'interaction atomiques. Chaque opérateur est associé à un temps unitaire correspondant au temps nécessaire pour réaliser cette interaction atomique. Dans le cadre de nos études, l'utilisation du tutoriel intégré à HoloLens pour apprendre les gestes dure environ 10 minutes et les différentes phases d'entraînement précédant chaque tâche durent entre un tiers et un quart du temps total des expériences. Ce tutoriel et cet entraînement ont donc permis aux participants de se familiariser avec les gestes requis et de les maîtriser. De plus, les gestes effectués (lever la main, tourner la tête, baisser le doigt et appuyer sur un bouton d'un dispositif dédié) sont des gestes mécaniques simples qui ne nécessitent pas d'expertise particulière. Par conséquent, il semble raisonnable de considérer que les temps établis dans ces travaux sont indépendants du degré d'expertise de l'utilisateur avec les dispositifs de réalité mixte.

Grâce à trois études d'utilisateurs supplémentaires, nous avons consolidé notre modèle en utilisant des tâches réalistes. Tout d'abord, nous avons réalisé une étude pour utiliser les temps des unités **Php**, **Ms** et toutes les unités de temps de **Chp** dans une tâche de pointage où les cibles étaient placées en dehors du FoV. Ensuite, nous avons comparé notre modèle à une tâche de sélection multiple avec des cibles placées à l'extérieur du champ de vision, impliquant les opérateurs **Chp**, **Php**, **Ms** et **Bc**. Enfin, dans une dernière étude, nous avons comparé notre temps prédit avec une tâche de sélections multiples avec des cibles déjà visibles dans le FoV impliquant les opérateurs **Php**, **Rh**, **Ms** et **At**. La différence entre le temps prédit et le temps observé est conforme au critère d'erreur de 20% de (Olson & Olson, 1990), ce qui valide notre extension KLM.

Pour soutenir davantage l'exploration de cet espace de conception, plusieurs extensions peuvent être envisagées. Premièrement, les dispositifs de réalité mixte permettent à l'utilisateur d'effectuer d'autres formes d'interactions que celles considérées dans ce travail (par exemple, la

manipulation d'objets numériques 3D). Ces interactions nécessitent donc des opérateurs KLM supplémentaires qui doivent être modélisés de manière appropriée. Cependant, ces interactions ne peuvent être réalisées qu'après une tâche de recherche, de pointage et de sélection d'un objet numérique. C'est pourquoi nous nous sommes concentrés dans cette première étude sur les opérateurs fondamentaux pour sélectionner des éléments numériques. Des extensions futures impliquant de nouveaux opérateurs (par exemple Pinch, Grab et Move d'objets 3D, etc.) peuvent être envisagées.

Enfin, les travaux précédents sur les extensions KLM ont modélisé les interactions basées sur les smartphones (cf. chapitre État de l'art section 3.3.1). La combinaison d'écrans tactiles avec des casques pourrait être utile pour surmonter les problèmes d'interaction en l'air en réalité mixte, tels que la fatigue (Hansberger et al., 2017; Hincapié-Ramos, Guo, Moghadasian, et al., 2014; Jang et al., 2017), la précision limitée (Velloso et al., 2015) et la « découvrabilité » des gestes (Baudel & Beaudouin-Lafon, 1993). À cette fin, il sera également intéressant d'étendre notre modèle KLM avec des opérateurs pour les interactions basées sur les smartphones présentés dans l'État de l'art dans le contexte de la réalité mixte.

Chapitre 6 Conclusion et Perspectives

Contenu du Chapitre

1 Résumés des contributions	130
1.1 Fonctionnelle	130
1.2 Conceptuelle	131
2 Perspectives	132
2.1 Moyen Terme	132
2.2 Long Terme	135
2.3 Nouvelles solutions pour neOCampus.....	138
3 Projet de recherche	139
3.1 Support d’affichage	140
3.2 Objet d’intérêt.....	141
3.3 Instrument d’interaction	141

Les travaux présentés dans ce manuscrit s’inscrivent dans le cadre du projet neOCampus promouvant un campus durable. Ce projet doit permettre aux acteurs du campus (étudiants, enseignants-chercheurs, personnel administratif, etc.) de prendre conscience des différentes dépenses énergétiques du campus. Pour ce faire, l’ensemble de nos travaux s’inspirent du domaine novateur de la représentation de données situées (Willett et al., 2017). Au travers d’un état de l’art détaillé, nous avons exploré les différentes technologies permettant de concevoir des applications de représentations de données situées en détaillant leurs spécificités. Nous avons concentré cet état de l’art sur les dispositifs de données situées quand le référent physique est un objet. Cet état de l’art a fait émerger deux problématiques principales :

- **Une problématique fonctionnelle.** Dans le cadre d'une visualisation de données situées dans un contexte public, quelle technique d'interaction permet le mieux à des usagers de sélectionner un objet physique ? (Chapitre TouchGlass)
- **Une problématique conceptuelle.** La conception et l'évaluation d'un système de réalité mixte pour la visualisation interactive de données situées peuvent être couteuses en ressources humaines ou temps. Un modèle prédictif peut-il prévoir le temps d'interaction d'un système de réalité mixte, en particulier les tâches de pointage et de validation ? (Chapitre KLM)

Dans ce chapitre, nous résumons les différentes contributions que nous avons pu apporter dans ce contexte. Dans un deuxième temps, nous présentons les différentes perspectives envisagées à moyen et long terme à ces travaux.

1 Résumés des contributions

1.1 Fonctionnelle

Dans le chapitre TouchGlass, nous présentons la conception et l'évaluation d'une technique de sélection d'un objet physique au travers d'une vitrine sans avoir de retour visuel pour faciliter la sélection. Les travaux présentés dans ce chapitre proposent d'utiliser les contraintes environnementales d'un système de visualisation de données situées pour explorer une nouvelle technique d'interaction. En effet, il n'est pas rare que les systèmes de visualisation de données situées soit présentés dans des environnements publics. Or les objets physiques présentés publiquement doivent être protégés de la saleté ou des potentielles dégradations. Ces objets physiques sont alors généralement protégés par une vitrine. De plus, certains objets comme les œuvres d'art ne permettent pas l'affichage d'information directement sur l'objet, car la lumière pourrait les endommager.

Les utilisateurs de systèmes de visualisation de données situées dans un espace public n'ont généralement pas le temps d'apprendre des interactions complexes. Dans une première étude, nous étudions quelle technique de sélection basée sur la vitrine serait la plus « *walk-up and use* » pour des utilisateurs passant devant notre dispositif. Les résultats de cette étude montrent que 75% des utilisateurs utilisent un lancer de rayon orthogonal à la vitrine. Cette approche n'est cependant pas étudiée dans la littérature.

Dans une deuxième étude, nous proposons d'étudier cette technique de sélection selon 4 facteurs : la taille de l'objet physique à sélectionner, la distance de cet objet à la vitrine, la présence d'une référence spatiale et la position de la tête de l'utilisateur par rapport à la vitrine. Cette technique de sélection est évaluée selon le temps de réalisation de la tâche. Nous étudions aussi le premier contact avec la vitrine. Autrement dit, est-ce que les utilisateurs sont en capacité de sélectionner des cibles lors du premier contact avec la vitrine. Nous étudions ce premier contact selon quatre métriques : l'*offset*, l'*offsetX*, l'*offsetY* et le *spread*. Les résultats de l'étude montrent que les plus petites cibles (1cm) sont complexes à sélectionner. Cependant, la présence d'une référence spatiale réduit l'*offset* et le *spread* facilitant ainsi la sélection. Nous proposons plusieurs conseils de conception concernant l'utilisation d'une vitrine interactive pour sélectionner des objets placés à l'intérieur.

1.2 Conceptuelle

Dans le chapitre KLM, nous présentons un modèle de prédiction du temps de réalisation de tâche basé sur KLM et étendu au cas particulier des interactions en réalité mixte. La démocratisation de nouveaux casques de réalité mixte (par ex. HoloLens ou MagicLeap One) a permis d'ouvrir de nombreuses perspectives pour la visualisation de données situées. Ces casques sont dotés de capteurs permettant de cartographier et d'enregistrer l'environnement en temps réel. Il est donc possible d'ancrer des données numériques sur des objets détectés par le casque. Cependant, la conception et l'évaluation des applications de réalité mixte peuvent être coûteuses en temps et en ressources (humaines ou matérielles). L'utilisation d'un modèle peut prévenir des problèmes d'utilisabilité en amont du développement. Ce modèle a pour but de prédire le temps de réalisation de tâches courantes des applications de réalité mixte : le pointage et la validation. Nous proposons une extension au modèle KLM en : 1) ajoutant de nouveaux opérateurs pour décrire les actions de l'utilisateur, 2) définissant les unités de temps relatives à ces nouveaux opérateurs et 3) consolidant ce modèle à l'aide de trois études utilisateur.

Dans la première section de ce chapitre, nous définissons cinq opérateurs utiles pour sélectionner (pointer + valider) des objets numériques présents en réalité mixte. Nous proposons deux opérateurs permettant d'effectuer une action de pointage : l'opérateur Coarse Head Point et l'opérateur Precise Head Point. Ces deux opérateurs permettent de modéliser l'action de déplacer le champ de vision de l'utilisateur pour visualiser des cibles et l'action de pointer des cibles quand elles sont à l'intérieur du champ de vision. Nous proposons ensuite trois opérateurs pour valider ce pointage : Button Click, Raise Hand et Air Tap. Le premier modélise l'appui sur un bouton présent sur un dispositif dédié (par ex. le Clicker de l'HoloLens). Les deux autres

permettent de modéliser une validation gestuelle consistant à lever la main pour qu'elle soit reconnue dans le champ de vision du casque puis à effectuer un geste de validation.

Nous avons ensuite défini le temps de réalisation de chaque action correspondant aux nouveaux opérateurs (c'est-à-dire l'unité de temps relative à chaque opérateur). Au travers de deux sessions d'études utilisateur, nous avons attribué une ou plusieurs unités de temps à chaque opérateur (cf. Tableau 4). Enfin, nous avons consolidé ce modèle en intégrant ces opérateurs dans des tâches écologiques : 1) le pointage d'objets en dehors du champ de vision, 2) la sélection d'objets en dehors du champ de vision et 3) la sélection d'objets dans le champ de vision. Pour chacune de ces tâches, nous avons comparé la moyenne des temps réalisés par les utilisateurs au temps prédit par notre modèle. Notre modèle peut prédire des tâches de pointage et de validation avec une différence d'au maximum 5% ce qui est conforme au critère de validation d'Olson et Olson (Olson & Olson, 1990).

Ces contributions ont permis : 1) d'étudier et de concevoir une nouvelle technique d'interaction pour des référents physiques protégés par une vitrine et 2) de décrire et concevoir un nouveau modèle de prédiction du temps de réalisation de tâches avec des données situées pour la réalité mixte. De plus, durant cette thèse, ces différentes contributions ont donné lieu à la conception et au développement de solutions interactives avancées concrètes, déclinées dans le contexte du projet neOCampus dont notamment deux scénarios décrits dans le chapitre Contexte et Outils et deux solutions interactives avec des données situées relatives à la gestion du bruit dans la bibliothèque universitaire. Ces travaux sont détaillés en Annexe.

Au-delà de ces contributions et illustrations, ces travaux ont aussi mis en évidence plusieurs perspectives à prendre en compte à moyen et à long terme.

2 Perspectives

2.1 Moyen Terme

2.1.1 *TouchGlass*

Dans le chapitre *TouchGlass*, nous proposons une étude de sélection où les référents physiques sont des sphères présentées sur un panneau. La sélection de ces sphères se fait dans un espace en 2 dimensions. Certains référents physiques placés à l'intérieur d'une vitrine peuvent en occulter d'autres. Il serait alors intéressant de définir un espace de conception des techniques d'interaction 3D pour des objets placés derrière une vitrine. Cet espace pourrait être composé

de nombreuses propriétés comme : le nombre de doigts, le nombre de mains, le type de correspondance entre la vitrine et la profondeur (c'est-à-dire la 3^e dimension). Ensuite, les différentes techniques issues de cet espace de conception peuvent être évaluées à l'aide d'études utilisateur.

De plus, la conception de TouchGlass a pour but de sélectionner de nombreux types de référents physiques différents. Dans notre outil logiciel (chapitre Contexte et Outils section 2.5) nous présentons différents types de maquettes de bâtiment servant de référent physique. Parmi elles, nous proposons une maquette *3D fermée* (c'est-à-dire, l'intérieur du bâtiment est caché par les murs et les façades du bâtiment). Or, la sélection d'un élément précis présent à l'intérieur de cette maquette peut être nécessaire pour visualiser des données de granularité fine (par ex. les valeurs collectées par un capteur à l'intérieur d'une salle du bâtiment). Par exemple, le domaine de la *X Ray Vision* (Sandor et al., 2010), permet de visualiser des données sur des surfaces occultant les objets placés derrière elles. Cependant, cette visualisation dégrade la perception de l'objet ou des données si celles-ci sont affichées sur l'objet. Il serait alors intéressant de définir de nouvelles techniques de *X Ray Vision* limitant la dégradation de la perception du référent physique puis d'étendre TouchGlass à la sélection d'éléments se trouvant à l'intérieur du référent. Par exemple, l'affichage de l'intérieur du bâtiment pourrait être déporté à proximité du bâtiment soit en 2D avec de la vidéoprojection soit en 3D en utilisant un casque de réalité mixte. Pour renforcer alors le lien entre la représentation du référent, le plan de son intérieur et des données numériques, il deviendrait alors crucial de réfléchir à la distribution des modalités d'interaction sur un continuum allant du tout physique (rendu physicalisé et interaction au contact de la maquette) au totalement virtuel (rendu en réalité mixte et interaction gestuelle) en passant par une combinaison hybride et les retours d'informations appropriés.

Enfin, TouchGlass a été conçu sans afficher de retour visuel permettant d'aider l'utilisateur à sélectionner un objet physique. En effet, certains objets physiques peuvent se dégrader s'ils sont illuminés. Néanmoins, il pourrait être intéressant d'étudier des retours visuels déportés à côté de l'objet physique voire sur la vitrine. L'apparition d'écrans tactiles transparents (p. ex. les derniers écrans du constructeur ProDisplay²⁰) pourrait permettre de détecter l'interaction en entrée et d'afficher des indices visuels pour sélectionner des éléments présents à l'intérieur de la vitrine.

²⁰ <https://prodisplay.com/products/interactive-transparent-touch-screen/>

2.1.2 KLM

Dans le chapitre KLM, nous avons présenté un nouveau modèle, basé sur KLM, de prédiction du temps de réalisation de tâches de pointage et de validation en réalité mixte. Ce modèle est composé de cinq nouveaux opérateurs et de leur unité de temps associée. Nous avons décrit les unités de temps de nos opérateurs avec une ou plusieurs valeurs. Or, comme évoqué dans l'État de l'art section 3.3.3, il est possible de définir certaines unités de temps par des équations paramétriques. Ces équations rendent le modèle plus complexe à utiliser car le moindre changement d'un paramètre fait évoluer la prédiction du modèle. Néanmoins, à l'instar de l'extension proposée par (A. Lee et al., 2015), il est possible de proposer deux types de valeurs pour un opérateur permettant donc une plus grande flexibilité d'utilisation au modèle. Un concepteur expert développant une application pour un système critique pourrait donc vouloir utiliser les équations présentes dans le modèle pour évaluer très précisément son application alors que le concepteur novice développant une application pour le grand public pourrait utiliser les unités de temps déjà présentes dans notre modèle. Par exemple, décrire le mouvement permettant de déplacer le champ de vision pour afficher le contenu numérique (c'est-à-dire notre opérateur Coarse Head Point) pourrait être décrit par une équation paramétrique. Cette équation pourrait avoir comme paramètre la distance angulaire verticale (dav) et la distance angulaire horizontale (dah) à parcourir par l'utilisateur. Pour chaque valeur de dav et de dah , cette équation prédirait le temps de réalisation du mouvement effectué par l'utilisateur.

Ensuite, la création de notre modèle avait pour but de prédire le temps de réalisation des tâches de pointage et de validation. Cependant, d'autres tâches peuvent être réalisées en complément du pointage et de la validation. Par exemple, le casque HoloLens permet de reconnaître des gestes pour attraper puis déplacer des objets numériques dans l'espace. Notre modèle peut être étendu pour décrire de nouveaux gestes ou interactions pouvant être utilisés dans des applications de réalité mixte. Il sera nécessaire de décrire ces nouveaux opérateurs, d'évaluer leur unité de temps et de valider leur utilisation dans des tâches écologiques.

Enfin, les systèmes interactifs tendent à devenir de plus en plus ubiquitaires et plusieurs dispositifs interactifs peuvent s'associer pour définir de nouveaux paradigmes d'interaction. Par exemple, Zhu et Grossman (Zhu & Grossman, 2020) proposent un cadre conceptuel pour la conception de systèmes interactifs combinant un casque de réalité mixte et le smartphone. Les deux dispositifs servant à la fois de dispositif d'affichage (par ex. le casque affiche des objets numériques en 3D et le téléphone affiche de longs textes détaillant l'objet) et de dispositif d'interaction en entrée (par ex. le casque permet de pointer un élément avec sa tête et le

téléphone permet de valider la sélection ou d'appliquer une commande présente dans un menu). Notre modèle KLM pourrait alors être étendu pour intégrer des opérateurs basés sur le smartphone (par ex. comme indiqué dans l'État de l'art section 3.3.3) et des opérateurs modélisant le cout de transition entre les deux dispositifs (p. ex. regarder un élément dans le casque puis regarder un élément sur le téléphone). Ces opérateurs seraient ensuite évalués dans des tâches écologiques.

2.1.3 Outils

Dans le chapitre Contexte et Outils, nous avons présenté un outil logiciel et un outil conceptuel permettant de prendre en main les solutions technologiques et la littérature nécessaire à nos deux contributions. L'outil logiciel définit un plan des opérations à mettre en place pour construire une maquette interactive dotée de plusieurs types de visualisation de données situées. Cependant, certaines opérations sont encore fastidieuses. Par exemple, la correspondance entre les briques, groupes de briques et les données doit se faire sur l'interface graphique du module complémentaire développé avec Unity. Cette étape pourrait se faire directement à la construction de la maquette physique. De plus, la calibration des vidéoprojecteurs n'est pas encore automatisée. L'utilisateur doit alors indiquer dans Unity le positionnement exact des vidéoprojecteurs et les adapter aux lentilles des vidéoprojecteurs. Ces étapes doivent être améliorées pour tendre vers une automatisation de ce processus.

L'outil conceptuel, au travers d'un espace de caractérisation, a permis de décrire les systèmes de visualisation de données situées selon 5 propriétés : la localisation des données, la distance utilisateur-référent, la superposition moteur-visuel, la relation référent-médiateur et la localisation du *feedback*. Dans le chapitre Contexte et Outils, nous avons établi le pouvoir descriptif et comparatif de notre espace. Cependant, le pouvoir génératif nécessaire à la validation complète d'un espace de conception n'a pu être évalué et devra donc être mis en place dans un futur proche. De plus, les pouvoirs descriptifs et comparatifs ont été illustrés à l'aide d'un corpus de 49 articles issus de la conférence CHI. Une revue systématique de la littérature devrait être menée pour maintenir à jour ce corpus d'articles.

2.2 Long Terme

Les travaux présentés dans ce manuscrit s'inspirent du domaine des représentations de données situées récemment conceptualisé par Willett et collab. (Willett et al., 2017). Ce domaine prend forme dans le monde physique grâce au référent et dans le monde numérique grâce aux données relatives au référent. Les techniques d'interaction adaptées à ce domaine restent encore sous-

exploitées. En effet, contrairement à de nombreux travaux ayant utilisé le monde physique et ses objets comme dispositif d'interaction en entrée (on parle alors d'interaction tangible), l'interaction avec un référent physique utilise l'objet physique comme cible de l'interaction plutôt que comme sa source. Cette problématique est au croisement de deux domaines : l'interaction en réalité virtuelle (voire augmentée) qui étudie de nombreuses interactions avec des objets 3D numériques et l'interaction tangible qui étudie l'interaction avec des objets physiques selon leur forme, leur taille, etc. De nombreuses pistes de réflexion doivent être étudiées pour concevoir des systèmes de visualisation de données situés performants et évaluer l'efficacité de ces systèmes sur les utilisateurs.

La conception de nouvelles techniques facilitant la sélection d'objets 3D numériques à différentes profondeurs a déjà fait l'objet de nombreux travaux (Argelaguet & Andujar, 2013). Cependant certaines cibles sont parfois occultées (en partie ou entièrement) par d'autres cibles ce qui rend leur sélection complexe, voire impossible. Les stratégies utilisées pour faciliter la sélection de ces cibles proviennent généralement du caractère numérique de l'interaction ou des cibles représentées en 3D. Les stratégies mises en place telles que l'agrandissement des cibles (Guillon et al., 2015), la forme de l'outil de sélection (Hincapié-Ramos et al., 2015) ou la réorganisation des cibles dans l'espace (Grossman & Balakrishnan, 2006), etc. n'ont pas d'équivalent dans l'environnement physique. Il serait pertinent soit d'adapter ces stratégies pour les appliquer à des paramètres inhérents aux cibles physiques (c'est-à-dire l'immobilité de la cible, la taille, la forme, le point de vue de l'utilisateur qui peut se déplacer) soit de concevoir de nouvelles stratégies.

Dans chapitre TouchGlass, nous avons étudié la déviation entre la sélection perçue par l'utilisateur et la sélection réelle au premier contact sur une vitrine avec un rayon orthogonal. Cette étude est une piste intéressante pour comprendre comment l'utilisateur perçoit sa manière de sélectionner un objet physique et ainsi de concevoir des techniques facilitant la sélection en fonction de la sélection perçue par l'utilisateur. Le geste de désignation avec l'index est maîtrisé très tôt chez l'enfant (Butterworth, 2003). Ce geste de désignation avec l'index a été implémenté dans les premiers lanceurs de rayon avec la main pour sélectionner un objet à distance (Mine, 1995). Une première étude a permis d'évaluer et corriger l'écart entre le pointage perçu et le pointage réel de cibles 2D projetées sur un écran avec lancer de rayon à partir de l'index (Mayer et al., 2018). Dans cette étude, aucun retour visuel n'était apporté aux utilisateurs pour ne pas perturber leur geste de désignation (c'est-à-dire lancer de rayon partant de leur index).

Cependant, cette étude se limite à des cibles 2D placées en face des utilisateurs. Il serait intéressant d'étudier l'écart entre le pointage perçu et le pointage réel pour des cibles 3D physiques (p. ex. des référents physiques placés dans un environnement dense) et numériques. Ces cibles pourraient alors prendre différentes formes, se trouver à différentes distances et à différentes positions ou encore être présentes dans différentes densités. L'étude de chacun de ces facteurs permettrait de comprendre les stratégies mises en place par l'humain pour désigner un objet avec son doigt et pourrait permettre de mieux comprendre comment l'utilisateur pointe avec son index et ainsi adapter les techniques de pointage en fonction des différents paramètres étudiés. De plus, cette étude pourrait être répliquée en utilisant des dispositifs différents (contrôleur, stylet, etc.). Ces résultats seraient une avancée significative vers la vision ubiquitaire de Weiser (Weiser, 1994) où les interactions avec les systèmes doivent être le plus naturel possible.

Dans le chapitre KLM, nous avons montré qu'une approche basée modèle permettait d'évaluer et d'aider la conception d'applications de réalité mixte avec un casque. Néanmoins, ces modèles s'inspirent de la compréhension que la communauté de l'IHM ou de la psychologie a de l'utilisateur dans ces environnements. Or ces environnements deviennent de plus en plus ubiquitaires, et plusieurs dispositifs d'interaction peuvent se combiner et s'échanger des informations. Il est alors intéressant de comprendre comment l'utilisateur navigue entre ces dispositifs, dans quels buts et selon quelles stratégies. Par exemple, l'utilisation combinée d'un casque de réalité mixte et d'un téléphone (Zhu & Grossman, 2020) permet de tirer parti des avantages de chacun des dispositifs (par ex. la possibilité de visualiser du contenu 3D pour le casque et du contenu à haute résolution pour le smartphone). Cependant, changer de dispositif a potentiellement un coût cognitif dégradant ainsi le flux d'interaction de l'utilisateur. Tout d'abord, l'œil humain doit s'adapter à différentes profondeurs (c'est-à-dire le conflit accommodation-vergence (Kramida, 2016)) ce qui peut fatiguer l'utilisateur et réduire ses performances (p. ex. temps de réalisation de tâche). Cette compréhension est fondamentale pour comprendre comment l'utilisateur évolue dans un environnement ubiquitaire composé de plusieurs écrans ou outils interactifs ; d'autant plus si ces environnements proposent des représentations de données situées (c'est-à-dire fortement ancrées dans l'environnement). Plusieurs métriques pourraient décrire les différents effets de ce conflit.

Par exemple, une étude pourrait permettre de quantifier le temps de changement d'écran. Une étude précédente tend à indiquer que le changement d'écrans placés à plusieurs distances

focales entre un HoloLens et une projection sur un carton (Eiberger et al., 2019) influence le temps d'interaction. Néanmoins cette étude se concentre uniquement sur une tâche de recherche visuelle, une seule distance et un positionnement fixe des écrans.

Des recherches antérieures (Guitton & Volle, 1987; Sidenmark & Gellersen, 2020) ont permis de modéliser les mouvements de la tête et des yeux pour déplacer son regard d'une cible à une autre. Cette modélisation pourrait être ainsi comparée à un temps de réalisation collecté lors d'une expérimentation contrôlée. Durant cette étude, l'utilisateur devrait réaliser plusieurs tâches visuelles (p. ex. recherche d'un élément de couleur, recherche d'une forme, lecture d'un texte ou comptage d'éléments) dans le casque puis sur le smartphone et vice-versa. L'étude suivrait un plan intra-sujet avec comme facteurs potentiels : la présence d'un casque ou l'absence de casque (c'est-à-dire notre groupe contrôle), la distance angulaire entre les dispositifs, les différences de profondeur, la résolution des dispositifs ou la tâche visuelle. Ainsi la différence entre le temps modélisé et le temps effectué par les utilisateurs nous permettra d'évaluer le coût de transition entre des dispositifs.

De plus, il est maintenant reconnu que le port d'un casque de réalité virtuelle (Y. Wang et al., 2019) ou les écrans 3D (Bang et al., 2014) entraîne de la fatigue visuelle. Concernant la réalité augmentée, à notre connaissance, une seule étude a permis de quantifier la fatigue visuelle en fonction de la distance de vergence (Tian et al., 2018). Il sera primordial de faire en sorte de limiter la fatigue dans un premier temps avant de peut-être l'évaluer à l'aide d'une technologie d'*eye-tracking* (Y. Wang et al., 2019). De telles études seront incontournables pour élaborer les opérateurs KLM correspondants et permettre leur utilisation dans la modélisation de solutions encore plus complexes que celles traitées dans ce manuscrit.

2.3 Nouvelles solutions pour neOCampus

Le projet neOCampus a pour but d'impliquer tous les acteurs du campus pour réduire l'empreinte écologique de celui-ci ou améliorer la qualité de vie de ces usagers. C'est dans ce but que nous avons proposé TouchGlass un dispositif interactif relativement bon marché, solide et facile d'utilisation pour les passants (cf. Annexe). De plus, nous avons proposé un modèle de prédiction du temps basé sur KLM pour l'interaction en réalité mixte. La combinaison des deux projets pourrait donc permettre de créer de nouvelles applications pour neOCampus.

Par exemple, la consommation énergétique du campus principal de l'université Toulouse 3 est gérée par une personne dédiée (c'est-à-dire une économiste des flux) à cette mission. Or le campus contient un grand nombre de bâtiments et 68 structures de recherche. Chaque bâtiment ou

chaque salle produit des données en temps réel. Actuellement, l'économe des flux utilise de nombreux tableurs pour naviguer dans ce large ensemble de données, les comparer et révéler des anomalies. Concevoir une maquette interactive du campus permettrait à la personne en charge de naviguer spatialement et ainsi comparer plusieurs bâtiments spatialement proches sans avoir à effectuer des tâches complexes sur un tableur.

Le domaine en plein essor de l'*Immersive Analytics* propose de représenter des données complexes (réseau, topologie ou données multidimensionnelles) dans un environnement immersif, généralement à l'aide de casque de réalité mixte (Marriott et al., 2018). La combinaison de la maquette physique et d'un casque de réalité mixte permettrait alors de représenter des données complexes proches de leur référent physique. L'économe des flux lancerait une application dans son casque. Une fois l'application lancée, le casque scannerait la maquette afin de reconnaître chaque bâtiment. L'économe des flux pourrait à l'aide d'une interaction phygitale trier les données, les filtrer, les comparer ou changer le type de représentation.

Pour concevoir l'application et définir l'emplacement optimal de l'interaction phygitale, les concepteurs pourraient alors utiliser le modèle KLM défini dans le chapitre KLM. L'utilisation de notre modèle KLM permettrait de concevoir et d'évaluer de nouvelles techniques d'interaction en amont du développement. Cette application pourrait être comparée à l'utilisation du tableur actuel et ainsi quantifier les gains potentiels en matière de temps de réalisation de tâche.

La combinaison de nos deux contributions permettrait d'allier une interaction avec un support de contact potentiellement moins fatigante qu'une interaction gestuelle (cf. chapitre État de l'art section 2.3) et de concevoir rapidement une application en réalité mixte optimisant le temps d'interaction avec les données numériques.

3 Projet de recherche

Les travaux réalisés durant cette thèse ont permis de mettre en exergue le besoin d'interactivité des visualisations de données situées. Plus largement, l'apport du monde physique dans les systèmes interactifs dédiés à l'analyse de données situées est encore trop limité. Le monde physique et les objets qui le composent sont plus souvent utilisés comme outil d'interaction en entrée que comme un objet sur lequel on veut interagir. Beaudouin-Lafon (Beaudouin-Lafon, 2000) propose alors le concept d'interaction instrumentale où l'objet physique peut être vu

comme un **instrument d'interaction** quand l'objet physique est utilisé comme outil d'interaction en entrée et comme **objet d'intérêt** quand l'objet physique est la cible de l'interaction. Beaudouin-Lafon prend l'exemple du crayon comme **instrument** qui permet de dessiner sur une feuille alors **objet d'intérêt**, si la mine de ce crayon casse, le crayon devient **objet d'intérêt** d'un taille-crayon devenu **instrument d'interaction**. Le crayon est à la fois **instrument d'interaction** et **objet d'intérêt** dans une même tâche qui serait dessiner. Cette dualité offre une grande flexibilité d'interaction. Les objets composant le monde physique comme **instruments d'interaction** relèvent du domaine de l'interaction tangible. Si nos travaux se sont concentrés sur l'étude du référent physique comme **objet d'intérêt**. Il n'en reste pas moins que les objets physiques comme **objet d'intérêt** ne sont que peu étudiés dans la littérature. Mais nous pensons que ceci peut aussi être lié à la complexité de leur utilisation. Par défaut, ils ne communiquent pas d'informations numériques, leur forme ou leur taille ne sont pas modifiables. De plus, Beaudouin-Lafon (Beaudouin-Lafon, 2000) indique que l'action sur l'**objet d'intérêt** ou avec l'**instrument d'interaction** doit nécessiter une réponse (p. ex. visuelle ou kinesthésique). L'objet physique doit donc être aussi pensé comme **support d'affichage**.

Pour dépasser nos premières contributions sur le domaine de l'interaction phygitale et ainsi couvrir plus largement les différents rôles possibles de l'objet physique dans un environnement interactif il nous semble incontournable de concevoir des études ou des théories pour : 1) comprendre comment l'utilisateur peut interagir avec ces objets de tailles, de couleurs, de formes et dotés de capacités technologiques différentes, 2) définir quelles sont les propriétés des objets physiques qui favorisent leur utilisation comme **instrument d'interaction**, **objet d'intérêt** ou **support d'affichage**, 3) identifier et caractériser des mécanismes permettant de faire évoluer leur état parmi ces trois rôles en agissant sur leurs propriétés. Le projet TouchGlass présenté en Chapitre 4, est un premier pas dans cette direction. En effet, ce projet a permis de comprendre les facteurs influençant la sélection d'objets physiques (c'est-à-dire des **objets d'intérêt**).

3.1 Support d'affichage

Différents travaux proposent de nouveaux dispositifs d'affichages en deux (Serrano et al., 2017) ou en trois dimensions (p. ex. Shape-Changing Display (Qamar et al., 2018; Rasmussen et al., 2012) ou Data Physicalization (Jansen et al., 2015)) pouvant se modifier selon leur contenu ou les besoins des utilisateurs. Ces dispositifs peuvent prendre de nombreuses formes ou tailles différentes. Néanmoins, comprendre comment adapter le contenu aux propriétés physiques de

l'objet reste un défi majeur (Alexander et al., 2018). Dans le cas précis de la visualisation de données (c'est-à-dire l'interprétation du contenu est basée sur des variables visuelles telles que la couleur, la forme, la taille ou la distance), et plus particulièrement la visualisation située, cette adaptation est indispensable pour que l'information ne soit pas mal interprétée par l'utilisateur. Ce travail d'adaptation a déjà pu être étudié pour évaluer des visualisations de données sur une smartwatch (Blascheck et al., 2019). Il serait donc intéressant d'étendre ce travail à la visualisation de données sur objets physiques pour comprendre quelles propriétés de l'objet physique ont un effet sur la compréhension des données représentées. De plus, cette compréhension permettrait d'ouvrir de nombreuses perspectives d'interaction pour créer des techniques d'interaction innovantes ou adapter celles provenant d'autres domaines (p. ex. la réalité virtuelle (Hurter et al., 2019)).

3.2 Objet d'intérêt

Durant cette thèse, nos travaux se sont concentrés sur l'objet physique comme **objet d'intérêt**. Nous avons en effet proposé : 1) un outil logiciel où une maquette physique pouvait être sélectionnée et 2) une technique de sélection d'objets physiques au travers d'une vitrine. À terme, sélectionner ces objets doit permettre aux utilisateurs de mieux comprendre les données relatives à ces objets physiques. Il est alors important de pouvoir manipuler cet objet physique ou ces propriétés. Par exemple, la maquette physique peut être utilisée comme un outil de gestion d'un campus (cf. Annexe). L'utilisateur voudrait alors définir des seuils d'alertes. Si les données relatives à un bâtiment dépassent un certain seuil, une alarme est déclenchée. L'utilisateur veut ensuite copier le paramétrage de cette alerte et l'affecter à d'autres bâtiments. Comment effectuer alors une action de « copier-coller » sur un objet physique ? Plus largement, il serait intéressant d'étudier l'application de commandes ou la manipulation de widgets à des objets physiques utilisés comme **objet d'intérêt**.

3.3 Instrument d'interaction

L'interaction avec des objets physiques comme **instrument d'interaction** relève du domaine de l'interaction tangible. Généralement, les travaux proposés par la communauté se concentrent sur la création de nouveaux objets physiques permettant de détecter des manipulations à 6 degrés de liberté, des interactions tactiles, etc. Dans le cadre d'une visualisation de données situées proches d'un référent physique, il serait intéressant d'évaluer un dispositif permettant de sélectionner des éléments au contact et à distance du référent puis de manipuler, filtrer et trier les données numériques associées aux éléments sélectionnés. Bien qu'elle permette

d'évaluer les avantages et les inconvénients d'une interaction tangible, cette approche soulève plusieurs problèmes (p. ex. la portabilité de l'interaction (Drogemuller et al., 2021)). Les interactions opportunistes proposent d'utiliser des objets du quotidien présents dans l'environnement pour interagir avec une application (Henderson & Feiner, 2008). À notre connaissance, aucun travail ne propose d'utiliser des objets du quotidien pour interagir à la fois avec un objet 3D physique (le référent) et des données numériques.

Bibliographie

- Al-Megren, S., Khabti, J., & Al-Khalifa, H. S. (2018). A Systematic Review of Modifications and Validation Methods for the Extension of the Keystroke-Level Model. *Advances in Human-Computer Interaction, 2018*, 1–26. <https://doi.org/10.1155/2018/7528278>
- Alexander, J., Roudaut, A., Steimle, J., Hornbæk, K., Bruns Alonso, M., Follmer, S., & Merritt, T. (2018). Grand Challenges in Shape-Changing Interface Research. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3173574.3173873>
- Alonso, L., Zhang, Y. R., Grignard, A., Noyman, A., Sakai, Y., ElKatsha, M., Doorley, R., & Larson, K. (2018). CityScope: A Data-Driven Interactive Simulation Tool for Urban Design. Use Case Volpe. In A. J. Morales, C. Gershenson, D. Braha, A. A. Minai, & Y. Bar-Yam (Eds.), *Unifying Themes in Complex Systems IX* (pp. 253–261). Springer International Publishing.
- American Psychological Association. (2020). *Publication manual of the American Psychological Association (7th ed.)*. American Psychological Association. <https://doi.org/10.1037/0000165-000>
- Andujar, C., & Argelaguet, F. (2007, March). Virtual Pads: Decoupling Motor Space and Visual Space for Flexible Manipulation of 2D Windows within VEs. *2007 IEEE Symposium on 3D User Interfaces*. <https://doi.org/10.1109/3DUI.2007.340781>
- Argelaguet, F., & Andujar, C. (2013). A survey of 3D object selection techniques for virtual environments. *Computers & Graphics, 37*(3), 121–136. <https://doi.org/10.1016/j.cag.2012.12.003>
- Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355–385. <https://doi.org/10.1.1.30.4999>
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications, 21*(6), 34–47. <https://doi.org/10.1109/38.963459>
- Bach, B., Sicat, R., Pfister, H., & Quigley, A. (2017). *Drawing into the AR-CANVAS: Designing Embedded Visualizations for Augmented Reality*.

https://aviz.fr/~bbach/immersive2017/papers/IA_1539-paper.pdf

- Bailly, C., Leitner, F., & Nigay, L. (2019). Head-Controlled Menu in Mixed Reality with a HMD. In D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, & P. Zaphiris (Eds.), *Human-Computer Interaction -- INTERACT 2019* (pp. 395–415). Springer International Publishing.
- Ballendat, T., Marquardt, N., & Greenberg, S. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10*, 121. <https://doi.org/10.1145/1936652.1936676>
- Bandyopadhyay, D., Raskar, R., & Fuchs, H. (2001). Dynamic shader lamps: painting on movable objects. *Proceedings IEEE and ACM International Symposium on Augmented Reality*, 207–216. <https://doi.org/10.1109/ISAR.2001.970539>
- Bang, J., Heo, H., Choi, J.-S., & Park, K. (2014). Assessment of Eye Fatigue Caused by 3D Displays Based on Multimodal Measurements. *Sensors*, 14(9), 16467–16485. <https://doi.org/10.3390/s140916467>
- Baudel, T., & Beaudouin-Lafon, M. (1993). Charade: remote control of objects using free-hand gestures. *Communications of the ACM*, 36(7), 28–35. <https://doi.org/10.1145/159544.159562>
- Beaudouin-Lafon, M. (2004). Designing interaction, not interfaces. *Proceedings of the Working Conference on Advanced Visual Interfaces - AVI '04*, 15. <https://doi.org/10.1145/989863.989865>
- Beaudouin-Lafon, M. (2000). Instrumental interaction: an interaction model for designing post-WIMP user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '00*, 446–453. <https://doi.org/10.1145/332040.332473>
- Beheshti, E., Kim, D., Ecanow, G., & Horn, M. S. (2017). Looking Inside the Wires: Understanding Museum Visitor Learning with an Augmented Circuit Exhibit. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1583–1594. <https://doi.org/10.1145/3025453.3025479>
- Bernhaupt, R., Navarre, D., Palanque, P., & Winckler, M. (2008). *Model-Based Evaluation: A New Way to Support Usability Evaluation of Multimodal Interactive Applications* (pp. 96–

- 119). https://doi.org/10.1007/978-1-84628-941-5_5
- Blascheck, T., Besancon, L., Bezerianos, A., Lee, B., & Isenberg, P. (2019). Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches. *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 630–640. <https://doi.org/10.1109/TVCG.2018.2865142>
- Böhmer, M., Gehring, S., Löchtefeld, M., Ostkamp, M., & Bauer, G. (2011). The mighty un-touchables: creating playful engagement on media façades. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*, 605. <https://doi.org/10.1145/2037373.2037468>
- Brasier, E., Chapuis, O., Ferey, N., Vezien, J., & Appert, C. (2020). ARPads: Mid-air Indirect Input for Augmented Reality. *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 332–343. <https://doi.org/10.1109/ISMAR50242.2020.00060>
- Büschel, W., Mitschick, A., & Dachsel, R. (2018). Here and Now: Reality-Based Information Retrieval. *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval - CHIIR '18*, 171–180. <https://doi.org/10.1145/3176349.3176384>
- Butterworth, G. (2003). Pointing Is the Royal Road to Language for Babies. In S. Kita (Ed.), *Pointing: Where Language, Culture, and Cognition Meet*. (pp. 9–33). Lawrence Erlbaum Associates Publishers.
- Byrne, M. D. (2002). Cognitive Architecture. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 97–117). L. Erlbaum Associates Inc.
- Cabric, F., Dubois, E., Irani, P., & Serrano, M. (2019). TouchGlass: Raycasting from a Glass Surface to Point at Physical Objects in Public Exhibits. *IFIP Conference on Human-Computer Interaction*, 249–269. https://doi.org/10.1007/978-3-030-29387-1_15
- Caggianese, G., Colonnese, V., & Gallo, L. (2019). Situated Visualization in Augmented Reality: Exploring Information Seeking Strategies. *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 390–395. <https://doi.org/10.1109/SITIS.2019.00069>
- Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7), 396–410.

<https://doi.org/10.1145/358886.358895>

- Casiez, G. (2012). Du mouvement à l'interaction et au geste: études, techniques, outils et périphériques. *Habilitation à Diriger Des Recherches*.
- Chen, Y., Katsuragawa, K., & Lank, E. (2020). Understanding Viewport- and World-based Pointing with Everyday Smart Devices in Immersive Augmented Reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3313831.3376592>
- Choo, K. T. W., Balan, R. K., & Lee, Y. (2019). Examining Augmented Virtuality Impairment Simulation for Mobile App Accessibility Design. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–11. <https://doi.org/10.1145/3290605.3300605>
- Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Pardo, T. A., & Scholl, H. J. (2012). Understanding Smart Cities: An Integrative Framework. *2012 45th Hawaii International Conference on System Sciences*, 2289–2297. <https://doi.org/10.1109/HICSS.2012.615>
- Chubb, G. P. (1981). SAINT, A Digital Simulation Language for the Study of Manned Systems. In *Manned Systems Design* (pp. 153–179). Springer US. https://doi.org/10.1007/978-1-4613-3306-7_6
- Corsten, C., Cherek, C., Karrer, T., & Borchers, J. (2015). HaptiCase: Back-of-Device Tactile Landmarks for Eyes-Free Absolute Indirect Touch. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2171–2180. <https://doi.org/10.1145/2702123.2702277>
- Danyluk, K., Jenny, B., & Willett, W. (2019). Look-From Camera Control for 3D Terrain Maps. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3290605.3300594>
- Debarba, H. G., Grandi, J. G., Maciel, A., Nedel, L., & Boulic, R. (2013). Disambiguation canvas: A precise selection technique for virtual environments. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8119 LNCS(PART 3), 388–405. https://doi.org/10.1007/978-3-642-40477-1_24

- Desai, R., Anderson, F., Matejka, J., Coros, S., McCann, J., Fitzmaurice, G., & Grossman, T. (2019). Geppetto: Enabling Semantic Design of Expressive Robot Behaviors. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3290605.3300599>
- Dreyfuss Associates, H., & Wilcox, S. B. (2001). *The Measure of Man and Woman: Human Factors in Design (Updated Edition)* (A. Tilley (ed.)). Wiley Online Library.
- Drogemuller, A., Walsh, J., Smith, R. T., Adcock, M., & Thomas, B. H. (2021). Turning everyday objects into passive tangible controllers. *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 1–4. <https://doi.org/10.1145/3430524.3442460>
- Dubois, E., Gray, P., & Nigay, L. (2002). *ASUR++: A Design Notation for Mobile Mixed Systems* (pp. 123–139). https://doi.org/10.1007/3-540-45756-9_11
- Dupuy, G. (1992). *L'informatisation des villes*. PUF. <https://halshs.archives-ouvertes.fr/halshs-00438870>
- Eiberger, A., Kristensson, P. O., Mayr, S., Kranz, M., & Grubert, J. (2019). Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display. *Symposium on Spatial User Interaction*, 1–9. <https://doi.org/10.1145/3357251.3357588>
- El Batran, K., & Dunlop, M. D. (2014). Enhancing KLM (keystroke-level model) to fit touch screen mobile devices. *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services - MobileHCI '14*, 283–286. <https://doi.org/10.1145/2628363.2628385>
- ElSayed, N., Thomas, B. H., Marriott, K., Piantadosi, J., & Smith, R. T. (2016). Situated Analytics: Demonstrating immersive analytical tools with Augmented Reality. *Journal of Visual Languages & Computing*, 36, 13–23. <https://doi.org/10.1016/j.jvlc.2016.07.006>
- ElSayed, N., Thomas, B., Marriott, K., Piantadosi, J., & Smith, R. (2015). Situated Analytics. *2015 Big Data Visual Analytics (BDVA)*, 1–8. <https://doi.org/10.1109/BDVA.2015.7314302>
- Engelke, U., Marendy, P., Susanto, F., Williams, R., Mahbub, S., Nguyen, H., & De Souza, P. (2016). A Visual Analytics Framework to Study Honey Bee Behaviour. *2016 IEEE 18th*

- International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 1504–1511. <https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0214>
- Engelke, U., Rogers, C., Klump, J., & Lau, I. (2019). HypAR: Situated Mineralogy Exploration in Augmented Reality. *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, 1–5. <https://doi.org/10.1145/3359997.3365715>
- Ens, B., Ahlström, D., & Irani, P. (2016). Moving Ahead with Peephole Pointing: Modelling Object Selection with Head-Worn Display Field of View. *Proceedings of the 2016 Symposium on Spatial User Interaction*, 107–110. <https://doi.org/10.1145/2983310.2985756>
- Erazo, O., & Pino, J. A. (2018). Predicting user performance time for hand gesture interfaces. *International Journal of Industrial Ergonomics*, 65, 122–138. <https://doi.org/10.1016/j.ergon.2017.07.010>
- Erazo, O., & Pino, J. A. (2015). Predicting Task Execution Time on Natural User Interfaces based on Touchless Hand Gestures. *Proceedings of the 20th International Conference on Intelligent User Interfaces*, 97–109. <https://doi.org/10.1145/2678025.2701394>
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391. <https://doi.org/10.1037/h0055392>
- Forlines, C., & Balakrishnan, R. (2008). Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*, 1563. <https://doi.org/10.1145/1357054.1357299>
- Forsberg, A., Herndon, K., & Zeleznik, R. (1996). Aperture based selection for immersive virtual environments. *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology - UIST '96*, 95–96. <https://doi.org/10.1145/237091.237105>
- Freeman, E., Williamson, J., Subramanian, S., & Brewster, S. (2018). Point-and-Shake: Selecting from Levitating Object Displays. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–10. <https://doi.org/10.1145/3173574.3173592>

- Fruchard, B., Lecolinet, E., & Chapuis, O. (2017). MarkPad: Augmenting Touchpads for Command Selection. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5630–5642. <https://doi.org/10.1145/3025453.3025486>
- Gallo, L., De Pietro, G., & Marra, I. (2008). 3D Interaction with Volumetric Medical Data: experiencing the Wiimote. *Proceedings of the First International Conference on Ambient Media and Systems*. <https://doi.org/10.4108/ICST.AMBISYS2008.2880>
- Gehring, S., Löchtfeld, M., Daiber, F., Böhmer, M., & Krüger, A. (2012). Using intelligent natural user interfaces to support sales conversations. *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces - IUI '12*, 97. <https://doi.org/10.1145/2166966.2166985>
- Gilliot, J., Casiez, G., & Roussel, N. (2014). Impact of form factors and input conditions on absolute indirect-touch pointing tasks. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 723–732. <https://doi.org/10.1145/2556288.2556997>
- Green, J., Pridmore, T., & Benford, S. (2014). Exploring attractions and exhibits with interactive flashlights. *Personal and Ubiquitous Computing*, 18(1), 239–251. <https://doi.org/10.1007/s00779-013-0661-3>
- Green, P. (1999). Estimating Compliance with the 15-Second Rule for Driver-Interface Usability and Safety. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 43(18), 987–991. <https://doi.org/10.1177/154193129904301809>
- Grignard, A., Macià, N., Alonso, L., Noyman, A., Zhang, Y., & Larson, K. (2018). CityScope Andorra: A Multi-Level Interactive and Tangible Agent-Based Visualization. *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 1939–1940.
- Grignard, A., Nguyen-Huu, T., Gaudou, B., Nguyen-Ngoc, D., Brugiere, A., Dang-Huu, T., Nghi, H. Q., Khanh, N. T., & Larson, K. (2020). CityScope Hanoi: interactive simulation for water management in the Bac Hung Hai irrigation system. *2020 12th International Conference on Knowledge and Systems Engineering (KSE)*, 153–158. <https://doi.org/10.1109/KSE50997.2020.9287831>
- Grossman, T., & Balakrishnan, R. (2006). The design and evaluation of selection techniques for 3D volumetric displays. *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology - UIST '06*, 3.

<https://doi.org/10.1145/1166253.1166257>

- Grossman, T., Wigdor, D., & Balakrishnan, R. (2005). Multi-finger gestural interaction with 3D volumetric displays. *ACM Transactions on Graphics*, 24(3), 931–931. <https://doi.org/10.1145/1073204.1073287>
- Guarese, R., Becker, J., Fensterseifer, H., Walter, M., Freitas, C., Nedel, L., & Maciel, A. (2020). Augmented Situated Visualization for Spatial and Context-Aware Decision-Making. *Proceedings of the International Conference on Advanced Visual Interfaces*, 1–5. <https://doi.org/10.1145/3399715.3399838>
- Gugenheimer, J., Dobbstein, D., Winkler, C., Haas, G., & Rukzio, E. (2016). FaceTouch: Enabling Touch Interaction in Display Fixed UIs for Mobile Virtual Reality. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 49–60. <https://doi.org/10.1145/2984511.2984576>
- Guillon, M., Leitner, F., & Nigay, L. (2015). Investigating Visual Feedforward for Target Expansion Techniques. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2777–2786. <https://doi.org/10.1145/2702123.2702375>
- Guitton, D., & Volle, M. (1987). Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *Journal of Neurophysiology*, 58(3), 427–459. <https://doi.org/10.1152/jn.1987.58.3.427>
- Hall, A. D., Cunningham, J. B., Roache, R. P., & Cox, J. W. (1988). Factors affecting performance using touch-entry systems: Tactual recognition fields and system accuracy. *Journal of Applied Psychology*, 73(4), 711–720. <https://doi.org/10.1037/0021-9010.73.4.711>
- Hall, E. T. (1966). The hidden dimension, 1st ed. In *The hidden dimension, 1st ed.*
- Hansberger, J. T., Peng, C., Mathis, S. L., Areyur Shanthakumar, V., Meacham, S. C., Cao, L., & Blakely, V. R. (2017). Dispelling the Gorilla Arm Syndrome: The Viability of Prolonged Gesture Interactions. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10280, pp. 505–520). https://doi.org/10.1007/978-3-319-57987-0_41
- Harmon, B. A., Petrasova, A., Petras, V., Mitsova, H., & Meentemeyer, R. K. (2016). Tangible Landscape: Cognitively Grasping the Flow of Water. *ISPRS - International Archives of*

- the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B2*, 647–653.
<https://doi.org/10.5194/isprsarchives-XLI-B2-647-2016>
- Henderson, S. J., & Feiner, S. (2008). Opportunistic controls: leveraging natural affordances as tangible user interfaces for augmented reality. *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology - VRST '08*, 211.
<https://doi.org/10.1145/1450579.1450625>
- Hincapié-Ramos, J. D., Guo, X., & Irani, P. (2014). Designing Interactive Transparent Exhibition Cases. *7th International Workshop on Personalized Access to Cultural Heritage: The Future of Experiencing Cultural Heritage*, 16–19.
- Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., & Irani, P. (2014). Consumed endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1063–1072.
<https://doi.org/10.1145/2556288.2557130>
- Hincapié-Ramos, J. D., Ozacar, K., Irani, P. P., & Kitamura, Y. (2015). GyroWand: IMU-based Raycasting for Augmented Reality Head-Mounted Displays. *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, 89–98. <https://doi.org/10.1145/2788940.2788947>
- Hincapié-Ramos, J. D., Ozacar, K., Irani, P. P., & Kitamura, Y. (2016). GyroWand: An Approach to IMU-Based Raycasting for Augmented Reality. *IEEE Computer Graphics and Applications*, 36(2), 90–96. <https://doi.org/10.1109/MCG.2016.21>
- Hincapié-Ramos, J. D., Roscher, S., Büschel, W., Kister, U., Dachzelt, R., & Irani, P. (2014). cAR: Contact Augmented Reality with Transparent-Display Mobile Devices. *Proceedings of The International Symposium on Pervasive Displays*, 80–85.
<https://doi.org/10.1145/2611009.2611014>
- Hinckley, K., & Wigdor, D. (2002). Input Technologies and Techniques. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 151–168). L. Erlbaum Associates Inc.
<http://dl.acm.org/citation.cfm?id=772072.772085>
- Hoang, T., Reinoso, M., Joukhadar, Z., Vetere, F., & Kelly, D. (2017). Augmented Studio: Projection Mapping on Moving Body for Physiotherapy Education. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1419–1430.
<https://doi.org/10.1145/3025453.3025860>

- Hoggenmueller, M., & Wiethoff, A. (2016). LightBricks: A Physical Prototyping Toolkit for Do-it-yourself Media Architecture. *Proceedings of the 3rd Conference on Media Architecture Biennale*, 1–4. <https://doi.org/10.1145/2946803.2946811>
- Holleis, P., Otto, F., Hussmann, H., & Schmidt, A. (2007). Keystroke-level model for advanced mobile phone interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, 1505–1514. <https://doi.org/10.1145/1240624.1240851>
- Holleis, P., Scherr, M., & Broll, G. (2011). A Revised Mobile KLM for Interaction with Multiple NFC-Tags. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction -- INTERACT 2011* (pp. 204–221). Springer Berlin Heidelberg.
- Holst, A. (2021). *Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030*. <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/>
- Holz, C., & Baudisch, P. (2010). The generalized perceived input point model and how to double touch accuracy by extracting fingerprints. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 581. <https://doi.org/10.1145/1753326.1753413>
- Hurter, C., Riche, N. H., Drucker, S. M., Cordeil, M., Alligier, R., & Vuillemot, R. (2019). FiberClay: Sculpting Three Dimensional Trajectories to Reveal Structural Insights. *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 704–714. <https://doi.org/10.1109/TVCG.2018.2865191>
- Jagacinski, R. J., & Monk, D. L. (1985). Fitts' Law in Two Dimensions with Hand and Head Movements. *Journal of Motor Behavior*, 17(1), 77–95. <https://doi.org/10.1080/00222895.1985.10735338>
- Jang, S., Stuerzlinger, W., Ambike, S., & Ramani, K. (2017). Modeling Cumulative Arm Fatigue in Mid-Air Interaction based on Perceived Exertion and Kinetics of Arm Motion. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3328–3339. <https://doi.org/10.1145/3025453.3025523>
- Jansen, Y., Dragicevic, P., & Fekete, J.-D. (2013). Evaluating the efficiency of physical visualizations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2593–2602. <https://doi.org/10.1145/2470654.2481359>

- Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., Subramanian, S., & Hornbæk, K. (2015). Opportunities and Challenges for Data Physicalization. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3227–3236. <https://doi.org/10.1145/2702123.2702180>
- Jansen, Y., Schjerlund, J., & Hornbæk, K. (2019). Effects of Locomotion and Visual Overview on Spatial Memory when Interacting with Wall Displays. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3290605.3300521>
- Jin, S., Yin, J., Fu, C., Zhang, X., & Liu, T. (2020). Modeling Angle-Based Pointing Tasks in Augmented Reality Interfaces. *IEEE Access*, 8, 192597–192607. <https://doi.org/10.1109/ACCESS.2020.3031957>
- Jönsson, L., Broms, L., & Katzeff, C. (2010). Watt-Lite: energy statistics made tangible. *Proceedings of the 8th ACM Conference on Designing Interactive Systems - DIS '10*, 240. <https://doi.org/10.1145/1858171.1858214>
- Kawai, N., Sato, T., & Yokoya, N. (2016). Diminished Reality Based on Image Inpainting Considering Background Geometry. *IEEE Transactions on Visualization and Computer Graphics*, 22(3), 1236–1247. <https://doi.org/10.1109/TVCG.2015.2462368>
- Keddiseh, E., Serrano, M., & Dubois, E. (2021). KeyTch: Combining the Keyboard with a Touchscreen for Rapid Command Selection on Toolbars. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3411764.3445288>
- Khamis, M., Buschek, D., Thieron, T., Alt, F., & Bulling, A. (2018). EyePACT: Eye-Based Parallax Correction on Touch-Enabled Interactive Displays. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4), 1–18. <https://doi.org/10.1145/3161168>
- Kharlamov, D., Woodard, B., Tahai, L., & Pietroszek, K. (2016). TickTockRay: smartwatch-based 3D pointing for smartphone-based virtual reality. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 365–366. <https://doi.org/10.1145/2993369.2996311>
- Kieras, D. (2001). *Using the Keystroke-Level Model to Estimate Execution Times*.

- Kieras, D. (2009). Model-based Evaluation. In A. Sears & J. A. Jacko (Eds.), *Human-Computer Interaction: Development Process* (p. 18). CRC Press. <https://doi.org/10.1201/9781420088892>
- Kim, H.-J., Kim, J.-W., & Nam, T.-J. (2016). miniStudio: Designers' Tool for Prototyping Ubicomp Space with Interactive Miniature. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 213–224. <https://doi.org/10.1145/2858036.2858180>
- Kim, J., Bennett, G., Fu, S., & Kruse, J. (2019). Extended reality practice in art and design creative education. *SIGGRAPH Asia 2019 Courses*, 1–99. <https://doi.org/10.1145/3355047.3359407>
- Kirshenbaum, N., Hutchison, J., Theriot, R., Kobayashi, D., & Leigh, J. (2020). Data in Context: Engaging Audiences with 3D Physical Geo-Visualization. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–9. <https://doi.org/10.1145/3334480.3382968>
- Knierim, P., Kiss, F., Rauh, M., & Schmidt, A. (2020). Tangibility is Overrated: Comparing Learning Experiences of Physical Setups and their Virtual Equivalent in Augmented Reality. *19th International Conference on Mobile and Ubiquitous Multimedia*, 299–305. <https://doi.org/10.1145/3428361.3428379>
- Kopper, R., Bacim, F., & Bowman, D. A. (2011). Rapid and accurate 3D selection by progressive refinement. *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, 67–74. <https://doi.org/10.1109/3DUI.2011.5759219>
- Kopper, R., Bowman, D. A., Silva, M. G., & McMahan, R. P. (2010). A human motor behavior model for distal pointing tasks. *International Journal of Human-Computer Studies*, 68(10), 603–615. <https://doi.org/10.1016/j.ijhcs.2010.05.001>
- Kramida, G. (2016). Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(7), 1912–1931. <https://doi.org/10.1109/TVCG.2015.2473855>
- Kress, B. C. (2019). Digital optical elements and technologies (EDO19): applications to AR/VR/MR. In B. C. Kress & P. Schelkens (Eds.), *Digital Optical Technologies 2019* (p. 76). SPIE. <https://doi.org/10.1117/12.2544404>

- Kuhlman, C., Doherty, D., Nurbekova, M., Deva, G., Phyo, Z., Schoenhagen, P.-H., VanValkenburg, M., Rundensteiner, E., & Harrison, L. (2019). Evaluating Preference Collection Methods for Interactive Ranking Analytics. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–11. <https://doi.org/10.1145/3290605.3300742>
- Larson, K., Pastor, A. L. alberto, Babio Fernandez, G., Zhang, Y., Noyman, A., Grignard, A., ElKatsha, M., Sakai, Y., & Doorley, R. (2013). *CityScope Project*. <https://www.media.mit.edu/projects/cityscope/overview/>
- Laughery, K. R. (1989). Micro SAINT — A Tool for Modeling Human Performance in Systems. In *Applications of Human Performance Models to System Design* (pp. 219–230). Springer US. https://doi.org/10.1007/978-1-4757-9244-7_15
- Lavalle, A., Teruel, M. A., Maté, A., & Trujillo, J. (2020). Improving sustainability of smart cities through visualization techniques for Big Data from iot devices. *Sustainability (Switzerland)*, 12(14). <https://doi.org/10.3390/su12145595>
- Lee, A., Song, K., Ryu, H. B., Kim, J., & Kwon, G. (2015). Fingerstroke time estimates for touchscreen-based mobile gaming interaction. *Human Movement Science*, 44, 211–224. <https://doi.org/10.1016/j.humov.2015.09.003>
- Lee, J. H., & Bae, S.-H. (2013). Binocular cursor: enabling selection on transparent displays troubled by binocular parallax. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3169–3172. <https://doi.org/10.1145/2470654.2466433>
- Lee, S. C., Yoon, S. H., & Ji, Y. G. (2019). Modeling task completion time of in-vehicle information systems while driving with keystroke level modeling. *International Journal of Industrial Ergonomics*, 72, 252–260. <https://doi.org/10.1016/j.ergon.2019.06.001>
- Leithinger, D., Follmer, S., Olwal, A., Luescher, S., Hogge, A., Lee, J., & Ishii, H. (2013). Sublimate: state-changing virtual and physical rendering to augment interaction with shape displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1441–1450. <https://doi.org/10.1145/2470654.2466191>
- Leithinger, D., & Ishii, H. (2010). Relief: a scalable actuated shape display. *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '10*, 221. <https://doi.org/10.1145/1709886.1709928>

- Lewis, C., Polson, P. G., Wharton, C., & Rieman, J. (1990). Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Empowering People - CHI '90*, 235–242. <https://doi.org/10.1145/97243.97279>
- Liang, J., & Green, M. (1994). JDCAD: A highly interactive 3D modeling system. *Computers & Graphics*, 18(4), 499–506. [https://doi.org/10.1016/0097-8493\(94\)90062-0](https://doi.org/10.1016/0097-8493(94)90062-0)
- Lindlbauer, D., & Wilson, A. D. (2018). Remixed Reality: Manipulating Space and Time in Augmented Reality. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3173574.3173703>
- Lobo, M. J., & Christophe, S. (2020). Opportunities and challenges for augmented reality situated geographical visualization. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-4–2020, 163–170. <https://doi.org/10.5194/isprs-annals-V-4-2020-163-2020>
- Looser, J., Billinghamurst, M., Grasset, R., & Cockburn, A. (2007). An evaluation of virtual lenses for object selection in augmented reality. *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia - GRAPHITE '07*, 203. <https://doi.org/10.1145/1321261.1321297>
- Luo, S.-J., Yue, Y., Huang, C.-K., Chung, Y.-H., Imai, S., Nishita, T., & Chen, B.-Y. (2015). Legolization: Optimizing LEGO Designs. *ACM Transactions on Graphics*, 34(6), 1–12. <https://doi.org/10.1145/2816795.2818091>
- MacKenzie, I. S. (2013). Human-computer Interaction : An Empirical Research Perspective. In I. S. MacKenzie (Ed.), *Human-computer Interaction* (Morgan Kaufmann, pp. i–iii). Elsevier. <https://doi.org/10.1016/B978-0-12-405865-1.00009-1>
- Maeda, A., Hara, K., Kobayashi, M., & Abe, M. (2011). “ClearPlate” for capturing printed information: A Scanner and Viewfinder in One Optical Unit. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*, 2737. <https://doi.org/10.1145/1978942.1979346>
- Mäkelä, V., Sharma, S., Hakulinen, J., Heimonen, T., & Turunen, M. (2017). Challenges in Public Display Deployments: A Taxonomy of External Factors. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3426–3475. <https://doi.org/10.1145/3025453.3025798>

- Manes, D., & Green, P. (1997). *Evaluation of a Driver Interface: Effects of Control Type (Knob Versus Buttons) and Menu Structure (Depth Versus Breadth)*. <https://deepblue.lib.umich.edu/bitstream/handle/2027.42/1216/90529.0001.001.pdf?sequence=2>
- Manes, D., Green, P., & Hunter, D. (1997). *Prediction of Destination Entry and Retrieval Times Using Keystroke-Level Models*.
- Marquardt, N., & Greenberg, S. (2012). Informing the design of proxemic interactions. *IEEE Pervasive Computing*, 11(2), 14–23. <https://doi.org/10.1109/MPRV.2012.15>
- Marriott, K., Chen, J., Hlawatsch, M., Itoh, T., Nacenta, M. A., Reina, G., & Stuerzlinger, W. (2018). *Immersive Analytics: Time to Reconsider the Value of 3D for Information Visualisation* (pp. 25–55). https://doi.org/10.1007/978-3-030-01388-2_2
- Marshall, P., Morris, R., Rogers, Y., Kreitmayer, S., & Davies, M. (2011). : an in-the-wild study of how groups approach a walk-up-and-use tabletop interface. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*, 3033. <https://doi.org/10.1145/1978942.1979392>
- Matulic, F., & Vogel, D. (2018). Multiray: Multi-Finger Raycasting for Large Displays. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3173574.3173819>
- Mayer, S., Schwind, V., Schweigert, R., & Henze, N. (2018). The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3173574.3174227>
- McGuffin, M. J., & Balakrishnan, R. (2005). Fitts' law and expanding targets. *ACM Transactions on Computer-Human Interaction*, 12(4), 388–422. <https://doi.org/10.1145/1121112.1121115>
- McGuffin, M. J., & Balakrishnan, R. (2002). Acquisition of expanding targets. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Changing Our World, Changing Ourselves - CHI '02*, 57. <https://doi.org/10.1145/503376.503388>
- Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Keith Smith, J. E. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements.

- Psychological Review*, 95(3), 340–370. <https://doi.org/10.1037/0033-295X.95.3.340>
- Migge, B., & Kunz, A. (2010). User Model for Predictive Calibration Control on Interactive Screens. *2010 International Conference on Cyberworlds*, 32–37. <https://doi.org/10.1109/CW.2010.18>
- Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: a class of displays on the reality-virtuality continuum. In H. Das (Ed.), *Proceedings of SPIE - The International Society for Optical Engineering* (Vol. 2351, Issue Telemanipulator and Telepresence Technologies, pp. 282–292). <https://doi.org/10.1117/12.197321>
- Millar, G. C., Tabrizian, P., Petrasova, A., Petras, V., Harmon, B., Mitsova, H., & Meetenmeyer, R. K. (2018). Tangible Landscape: A Hands-on Method for Teaching Terrain Analysis. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3173574.3173954>
- Mine, M. R. (1995). *Virtual Environment Interaction Techniques*.
- Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., & Long, A. C. (2002). Interacting at a Distance: Measuring the Performance of Laser Pointers and Other Devices. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Changing Our World, Changing Ourselves - CHI '02*, 4, 33. <https://doi.org/10.1145/503376.503383>
- Mylavarapu, P., Yalcin, A., Gregg, X., & Elmqvist, N. (2019). Ranked-List Visualization: A Graphical Perception Study. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3290605.3300422>
- Myung, R. (2004). Keystroke-level analysis of Korean text entry methods on mobile phones. *International Journal of Human-Computer Studies*, 60(5–6), 545–563. <https://doi.org/10.1016/j.ijhcs.2003.10.002>
- Nowakowski, C., Utsui, Y., & Green, P. (2000). *Navigation System Destination Entry: The Effects of Driver Workload and Input Devices, And Implications for SAE Recommended Practice*. <https://deepblue.lib.umich.edu/bitstream/handle/2027.42/1323/93813.0001.001.pdf?sequence=2>
- O’Dea, S. (2020). *Number of smartphone users worldwide from 2016 to 2021*.

<https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/>

- Olson, J. R., & Olson, G. M. (1990). The Growth of Cognitive Modeling in Human-Computer Interaction Since GOMS. *Human-Computer Interaction*, 5(2-3), 221-265. <https://doi.org/10.1080/07370024.1990.9667155>
- Olwal, A., Benko, H., & Feiner, S. (2003). SenseShapes: using statistical geometry for object selection in a multimodal augmented reality. *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, 300-301. <https://doi.org/10.1109/ISMAR.2003.1240730>
- Perelman, G., Serrano, M., Raynal, M., Picard, C., Derras, M., & Dubois, E. (2015). The Roly-Poly Mouse: Designing a Rolling Input Device Unifying 2D and 3D Interaction. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 327-336. <https://doi.org/10.1145/2702123.2702244>
- Pham, D.-M., & Stuerzlinger, W. (2019). Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. *25th ACM Symposium on Virtual Reality Software and Technology*, 1-11. <https://doi.org/10.1145/3359996.3364264>
- Piccolo, L. S. G., Alani, H., De Liddo, A., & Baranauskas, C. (2014). Motivating online engagement and debates on energy consumption. *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 109-118. <https://doi.org/10.1145/2615569.2615695>
- Pietroszek, K., & Lank, E. (2012). Clicking blindly: using spatial correspondence to select targets in multi-device environments. *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '12*, 331. <https://doi.org/10.1145/2371574.2371625>
- Piper, B., Ratti, C., & Ishii, H. (2002). Illuminating clay: a 3-D tangible interface for landscape analysis. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Changing Our World, Changing Ourselves - CHI '02*, 355. <https://doi.org/10.1145/503376.503439>
- Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G., & Billingham, M. (2017). Exploring enhancements for remote mixed reality collaboration. *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on - SA '17*, 1-5.

<https://doi.org/10.1145/3132787.3139200>

- Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H., & Billingham, M. (2018). Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3173574.3173620>
- Priestnall, G., & Cheverst, K. (2019). Understanding visitor interaction with a projection augmented relief model display: insights from an in-the-wild study in the English Lake District. *Personal and Ubiquitous Computing*. <https://doi.org/10.1007/s00779-019-01320-2>
- Qamar, I. P. S., Groh, R., Holman, D., & Roudaut, A. (2018). HCI meets Material Science: A Literature Review of Morphing Materials for the Design of Shape-Changing Interfaces. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–23. <https://doi.org/10.1145/3173574.3173948>
- Rasmussen, M. K., Pedersen, E. W., Petersen, M. G., & Hornbæk, K. (2012). Shape-changing interfaces: a review of the design space and open research questions. *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12*, 735. <https://doi.org/10.1145/2207676.2207781>
- Rausch, T., Hummer, W., Stippel, C., Vasilijevic, S., Elvezio, C., Dustdar, S., & Krösl, K. (2021). *Towards a Platform for Smart City-Scale Cognitive Assistance Applications*.
- Rice, A. D., & Lartigue, J. W. (2014). Touch-level model (TLM): evolving KLM-GOMS for touchscreen and mobile devices. *Proceedings of the 2014 ACM Southeast Regional Conference on - ACM SE '14*, 1–6. <https://doi.org/10.1145/2638404.2638532>
- Ridel, B., Reuter, P., Laviolle, J., Mellado, N., Couture, N., & Granier, X. (2014). The Revealing Flashlight: Interactive Spatial Augmented Reality for Detail Exploration of Cultural Heritage Artifacts. *Journal on Computing and Cultural Heritage*, 7(2), 1–18. <https://doi.org/10.1145/2611376>
- Roberts, J., Banerjee, A., Hong, A., McGee, S., Horn, M., & Matcuk, M. (2018). Digital Exhibit Labels in Museums: Promoting Visitor Engagement with Cultural Artifacts. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3173574.3174197>

- Roo, J. S., Gervais, R., Frey, J., & Hachet, M. (2017). Inner Garden: Connecting Inner States to a Mixed Reality Sandbox for Mindfulness. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1459–1470. <https://doi.org/10.1145/3025453.3025743>
- Roudaut, A., Pohl, H., & Baudisch, P. (2011). Touch input on curved surfaces. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*, 1011. <https://doi.org/10.1145/1978942.1979094>
- Rühmann, L. M., Prilla, M., & Brown, G. (2018). Cooperative Mixed Reality: An Analysis Tool. *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, 107–111. <https://doi.org/10.1145/3148330.3154510>
- Saidi, H., Serrano, M., Irani, P., Hurter, C., & Dubois, E. (2019). *On-Body Tangible Interaction: Using the Body to Support Tangible Manipulations for Immersive Environments* (pp. 471–492). https://doi.org/10.1007/978-3-030-29390-1_26
- Sandor, C., Cunningham, A., Dey, A., & Mattila, V.-V. (2010). An Augmented Reality X-Ray system based on visual saliency. *2010 IEEE International Symposium on Mixed and Augmented Reality*, 27–36. <https://doi.org/10.1109/ISMAR.2010.5643547>
- Scheible, J., & Ojala, T. (2009). MobiSpray: Mobile Phone as Virtual Spray Can for Painting BIG Anytime Anywhere on Anything. *Leonardo*, 42(4), 332–341. <https://doi.org/10.1162/leon.2009.42.4.332>
- Schmidt, D., Block, F., & Gellersen, H. (2009). A comparison of direct and indirect multi-touch input for large surfaces. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5726 LNCS(PART 1), 582–594. https://doi.org/10.1007/978-3-642-03655-2_65
- Schmidt, G., Baillot, Y., Brown, D. G., Tomlin, E. B., & Swan, J. E. (2006). Toward Disambiguating Multiple Selections for Frustum-Based Pointing. *3D User Interfaces (3DUI'06)*, 87–94. <https://doi.org/10.1109/VR.2006.133>
- Schmidt, S., Steinicke, F., Irlitti, A., & Thomas, B. H. (2018). Floor-Projected Guidance Cues for Collaborative Exploration of Spatial Augmented Reality Setups. *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, 279–289. <https://doi.org/10.1145/3279778.3279806>

- Schneegaß, S., Pfleging, B., Kern, D., & Schmidt, A. (2011). Support for modeling interaction with automotive user interfaces. *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '11*, 71. <https://doi.org/10.1145/2381416.2381428>
- Sears, A., & Shneiderman, B. (1991). High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, 34(4), 593–613. [https://doi.org/10.1016/0020-7373\(91\)90037-8](https://doi.org/10.1016/0020-7373(91)90037-8)
- Seonho Lee, Yeseul Lim, Chun, J., Lee, S., Lim, Y., & Chun, J. (2013). 3D interaction in Augmented Reality with stereo-vision technique. *2013 15th International Conference on Advanced Communications Technology (ICACT)*, 401–405. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6488216
- Serrano, M., Roudaut, A., & Irani, P. (2017). Visual Composition of Graphical Elements on Non-Rectangular Displays. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 4405–4416. <https://doi.org/10.1145/3025453.3025677>
- Sharma, A., Liu, L., & Maes, P. (2013). Classified: an augmented ruler based on a transparent display for real-time interactions with paper. *Proceedings of the Adjunct Publication of the 26th Annual ACM Symposium on User Interface Software and Technology - UIST '13 Adjunct*, 21–22. <https://doi.org/10.1145/2508468.2514937>
- Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *The Craft of Information Visualization*. Elsevier. <https://doi.org/10.1016/B978-155860915-0/50046-9>
- Sidenmark, L., & Gellersen, H. (2020). Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Transactions on Computer-Human Interaction*, 27(1), 1–40. <https://doi.org/10.1145/3361218>
- Siek, K. A., Rogers, Y., & Connelly, K. H. (2005). Fat finger worries: How older and younger users physically interact with PDAs. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3585 LNCS, 267–280. https://doi.org/10.1007/11555261_24
- Simon, F., Roudaut, A., Irani, P., & Serrano, M. (2019). Finding Information on Non-Rectangular Interfaces. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–8. <https://doi.org/10.1145/3290605.3300332>

- Song, C. G., Kwak, N. J., & Jeong, D. H. (2000). Developing an efficient technique of selection and manipulation in immersive V.E. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology - VRST '00*, 142. <https://doi.org/10.1145/502390.502417>
- Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is Mixed Reality? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3290605.3300767>
- Spindler, M., & Dachsel, R. (2009). PaperLens: advanced magic lens interaction above the tabletop. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '09*, 1. <https://doi.org/10.1145/1731903.1731948>
- Statista Research Department. (2021). *Forecast unit shipments of wrist-worn wearables worldwide from 2019 to 2024*. <https://www.statista.com/statistics/296565/wearables-worldwide-shipments/>
- Steed, A. (2006). Towards a General Model for Selection in Virtual Environments. *3D User Interfaces (3DUI'06)*, 103–110. <https://doi.org/10.1109/VR.2006.134>
- Steinicke, F., Ropinski, T., & Hinrichs, K. (2006). Object Selection In Virtual Environments Using An Improved Virtual Pointer Metaphor. In *Computer Vision and Graphics* (pp. 320–326). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-4179-9_46
- Teather, R. J., & Stuerzlinger, W. (2013). Pointing at 3d target projections with one-eyed and stereo cursors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 159–168. <https://doi.org/10.1145/2470654.2470677>
- Teather, R. J., & Stuerzlinger, W. (2011). Pointing at 3D targets in a stereo head-tracked virtual environment. *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, 87–94. <https://doi.org/10.1109/3DUI.2011.5759222>
- Teather, R. J., Stuerzlinger, W., & Pavlovych, A. (2014). Fishtank fitts: a desktop VR testbed for evaluating 3D pointing techniques. *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, 519–522. <https://doi.org/10.1145/2559206.2574810>
- Thomas, B. H., Welch, G. F., Dragicevic, P., Elmqvist, N., Irani, P., Jansen, Y., Schmalstieg, D., Tabard, A., ElSayed, N. A. M., Smith, R. T., & Willett, W. (2018). *Situated Analytics* (pp. 185–220). https://doi.org/10.1007/978-3-030-01388-2_7
- Tian, M., Hernandez, J., Le, R., & Baldassi, S. (2018). Eye fatigue in Augmented Reality at

- different vergence distances. *Journal of Vision*, 18(10), 999. <https://doi.org/10.1167/18.10.999>
- Tilley, A., & Dreyfuss Associates, H. (1994). The Measure of Man and Woman – Human Factors in Design. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 2(2), 37–39. <https://doi.org/10.1177/106480469400200211>
- Vanacken, L., Grossman, T., & Coninx, K. (2007). Exploring the Effects of Environment Density and Target Visibility on Object Selection in 3D Virtual Environments. *2007 IEEE Symposium on 3D User Interfaces*. <https://doi.org/10.1109/3DUI.2007.340783>
- Velloso, E., Turner, J., Alexander, J., Bulling, A., & Gellersen, H. (2015). *An Empirical Investigation of Gaze Selection in Mid-Air Gestural 3D Manipulation* (pp. 315–330). https://doi.org/10.1007/978-3-319-22668-2_25
- Villanueva, A., Zhu, Z., Liu, Z., Pepler, K., Redick, T., & Ramani, K. (2020). Meta-AR-App: An Authoring Platform for Collaborative Augmented Reality in STEM Classrooms. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3313831.3376146>
- Vogel, D., & Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, 657–666. <https://doi.org/10.1145/1240624.1240727>
- Walker, W., Sung, H., Ong, C. K., & Casalegno, F. (2017). Exploring spatial meaning with a tangible map. *Proceedings of the 6th ACM International Symposium on Pervasive Displays*, 1–8. <https://doi.org/10.1145/3078810.3078826>
- Wang, F., & Ren, X. (2009). Empirical evaluation for finger input properties in multi-touch interaction. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 1063. <https://doi.org/10.1145/1518701.1518864>
- Wang, Y., Zhai, G., Chen, S., Min, X., Gao, Z., & Song, X. (2019). Assessment of eye fatigue caused by head-mounted displays using eye-tracking. *BioMedical Engineering OnLine*, 18(1), 111. <https://doi.org/10.1186/s12938-019-0731-5>
- Weiser, M. (1994). Ubiquitous computing. *Proceedings of the 22nd Annual ACM Computer Science Conference on Scaling up : Meeting the Challenge of Complexity in Real-World Computing Applications - CSC '94*, 418. <https://doi.org/10.1145/197530.197680>

- White, S., Feiner, S., & Kopylec, J. (2006). Virtual Vouchers: Prototyping a Mobile Augmented Reality User Interface for Botanical Species Identification. *3D User Interfaces (3DUI'06)*, 119–126. <https://doi.org/10.1109/VR.2006.145>
- White, Sean, & Feiner, S. (2009). SiteLens: situated visualization techniques for urban site visits. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 1117. <https://doi.org/10.1145/1518701.1518871>
- Wigdor, D., Jiang, H., Forlines, C., Borkin, M., & Shen, C. (2009). WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 1237. <https://doi.org/10.1145/1518701.1518886>
- Willett, W., Jansen, Y., & Dragicevic, P. (2017). Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 461–470. <https://doi.org/10.1109/TVCG.2016.2598608>
- Wyss, H. P., Blach, R., & Bues, M. (2006). iSith - Intersection-based Spatial Interaction for Two Hands. *3D User Interfaces (3DUI'06)*, 59–61. <https://doi.org/10.1109/VR.2006.93>
- Zheng, M., & Campbell, A. G. (2019). Location-Based Augmented Reality In-situ Visualization Applied for Agricultural Fieldwork Navigation. *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 93–97. <https://doi.org/10.1109/ISMAR-Adjunct.2019.00039>
- Zhou, Y., Xiao, S., Tang, N., Wei, Z., & Chen, X. (2016). Pmomo: Projection Mapping on Movable 3D Object. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 781–790. <https://doi.org/10.1145/2858036.2858329>
- Zhu, F., & Grossman, T. (2020). BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3313831.3376233>

Annexe : Transfert à neOCampus

Durant cette thèse, en plus des travaux présentés dans ce manuscrit, plusieurs projets ont été menés pour apporter de nouvelles solutions interactives pour la visualisation de données situées relatives au projet neOCampus. Dans le chapitre Contexte et Outils, plusieurs scénarios décrivant la gestion de la consommation énergétique ont été implémentés. Ces scénarios ont permis d'éprouver notre outil en fonction des différents données et types de maquettes présentes dans neOCampus. Nous rappellerons leur objectif et leur implémentation dans cette annexe. En plus du développement de ces scénarios, d'autres projets ont été menés. Ces projets avaient pour but de concevoir une application de gestion des capteurs installés dans des salles de classe et une application permettant d'améliorer la qualité de vie et de travail du personnel des usagers de la bibliothèque universitaire. Ce dernier projet se décompose en deux parties. La première pour indiquer aux étudiants la quantité de bruit qu'ils peuvent faire en travaillant. La deuxième pour permettre aux gestionnaires de la bibliothèque universitaire de gérer la bibliothèque (niveau de bruit, rappel à l'ordre, etc.) de leur poste de travail.

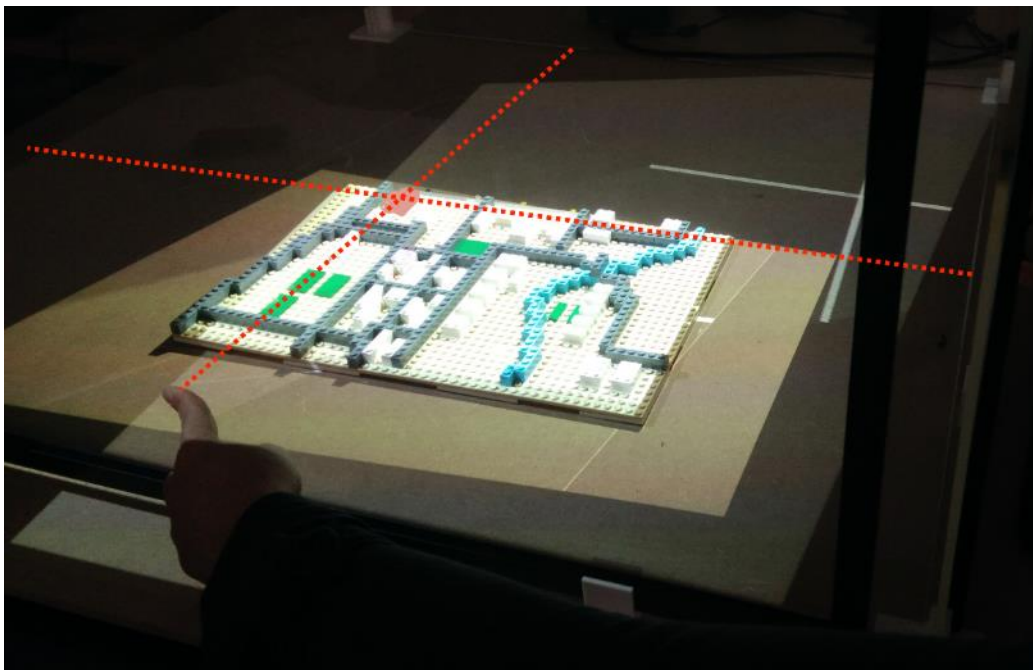


Figure 1 – La techniques d'interaction basée sur deux lancers de rayon orthogonaux

1 Gestion de la consommation énergétique

1.1 Au niveau du campus

Dans le cadre du projet neOCampus, il est important de pouvoir mesurer l’empreinte écologique du campus et de ses composantes (salles de classe, laboratoires, etc.). Pour ce faire, chaque bâtiment est équipé de capteurs permettant de mesurer la consommation instantanée du bâtiment. Ces données sont stockées sur un serveur et peuvent être affichées dans le campus à l’attention des étudiants ou du personnel. Afin de faciliter la compréhension et de prendre conscience de l’empreinte écologique du campus, les différents usagers doivent accéder aux données d’un ou plusieurs bâtiments. L’utilisation d’une maquette physique permet alors d’accéder aux données de tous les bâtiments sans avoir à s’y rendre physiquement.

Nous avons alors conçu une maquette physique en Lego du campus. Chaque bâtiment était représenté par une ou plusieurs briques Lego. La maquette était rendue interactive à l’aide de notre outil logiciel (cf. chapitre Contexte et Outils). Pour sélectionner un bâtiment, nous avons développé une technique inspirée du rayon orthogonal étudié dans le chapitre TouchGlass (cf. Figure 1) mais basée sur deux rayons orthogonaux et. L’intersection des deux rayons permettait de sélectionner un bâtiment.

Les données liées à ce modèle étaient toutes en relation avec la consommation énergétique (par ex. électricité, eau, gaz). Elles étaient modélisées pour chaque bâtiment. Si l’utilisateur

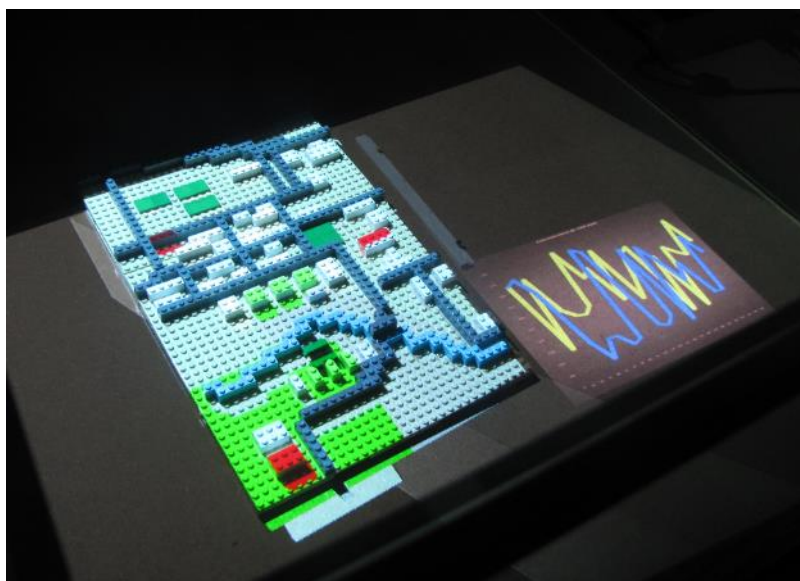


Figure 2 – la maquette 2,5D avec une comparaison de deux bâtiments

sélectionne un ou plusieurs bâtiments, les données relatives à ce ou ces bâtiments s'affichaient (cf. Figure 2)

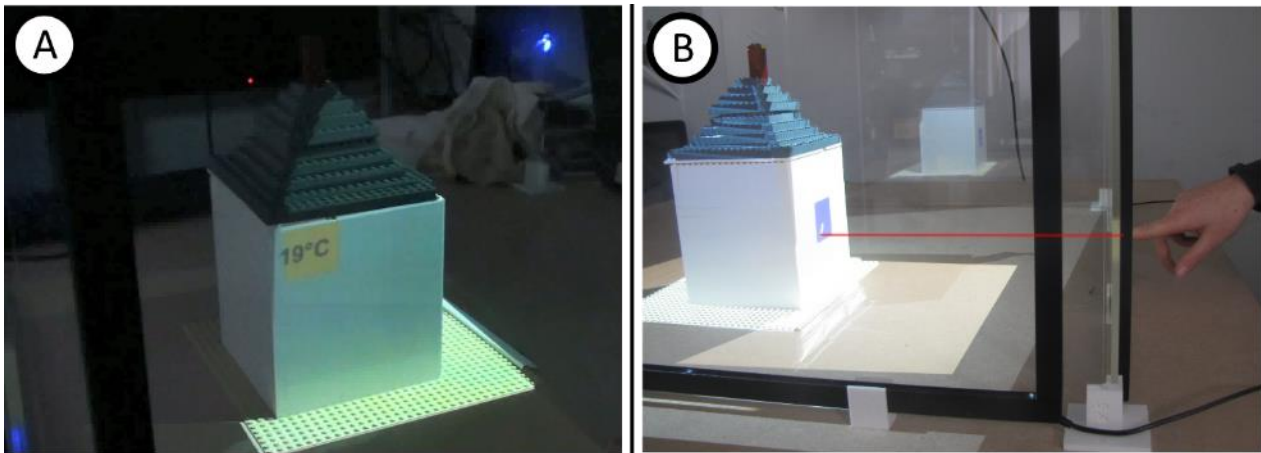


Figure 3 – Illustration de la maquette interactive du laboratoire avec (A) la maquette et (B) l'interaction utilisée pour sélectionner un élément de la maquette interactive

1.2 Au niveau d'un laboratoire de recherche

En 2020, l'Institut de Recherche en Informatique de Toulouse (IRIT) participait à un concours national (CUBE 2020²¹) ayant pour but de réduire la consommation énergétique du bâtiment. Pour faciliter la lecture de la consommation énergétique et inciter le personnel à la réduire, une maquette d'une partie du bâtiment a été construite. La maquette physique était liée à son pendant numérique à l'aide de notre outil (cf. chapitre Contexte et Outils). Trois étages étaient modélisés et plusieurs briques représentaient plusieurs équipes du laboratoire. Les dépenses énergétiques de chaque équipe étaient affichées directement sur la maquette (cf. Figure 3 A). Pour sélectionner un élément, la technique développée était similaire à TouchGlass. Un rayon perpendiculaire à la vitrine était projeté pour sélectionner la première pièce Lego qui était croisée par le rayon (cf. Figure 3 B).

2 Observation des données de capteurs installés dans des salles de classe

En plus de la réduction de l'empreinte énergétique du campus, les différents acteurs du projet neOCampus souhaitent aussi améliorer la qualité de vie du personnel et des étudiants. Pour ce faire, plusieurs salles de cours ont été équipées de capteurs permettant de collecter un ensemble de données (par ex. la température, l'humidité, le nombre de personnes, etc.). Ces données

²¹ <https://www.ifpeb.fr/nos-concours/cube-2020/>

peuvent être, par la suite, traitées pour comprendre l'activité dans une salle de cours ou pour adapter en temps réel les conditions de vie à l'intérieur de la salle (par ex. si la luminosité est trop forte, les volets se ferment pour garder des conditions d'apprentissage optimales). Afin de pouvoir observer ces données sans devoir se rendre directement dans les salles, nous avons proposé de construire une maquette physique de ces salles (cf. Figure 4 A). Les capteurs étaient imprimés en 3D et la maquette était augmentée à l'aide d'un vidéoprojecteur placé au-dessus de la maquette (cf. Figure 4 B).

L'utilisateur pouvait alors sélectionner un des capteurs pour afficher les données relatives au capteur sélectionné. Plusieurs techniques de sélection ont été conçues en s'inspirant d'une souris à 6 degrés de liberté (la Roly-Poly Mouse (Perelman et al., 2015)) et en se concentrant sur deux propriétés : le regroupement spatial des capteurs en îlot (c'est-à-dire un groupe de capteurs) et le type de navigation (spatiale à l'aide d'un curseur ou logique en parcourant un graphe de capteurs). Toutes ces techniques proposent différentes combinaisons de rotation, d'inclinaison et de translation pour sélectionner un capteur de la maquette.

Une fois un capteur sélectionné, les données étaient affichées sous forme de texte ou de graphes (cf. Figure 4 B). Cette maquette permettait de naviguer rapidement dans l'ensemble des capteurs (et donc des données) de cette salle.

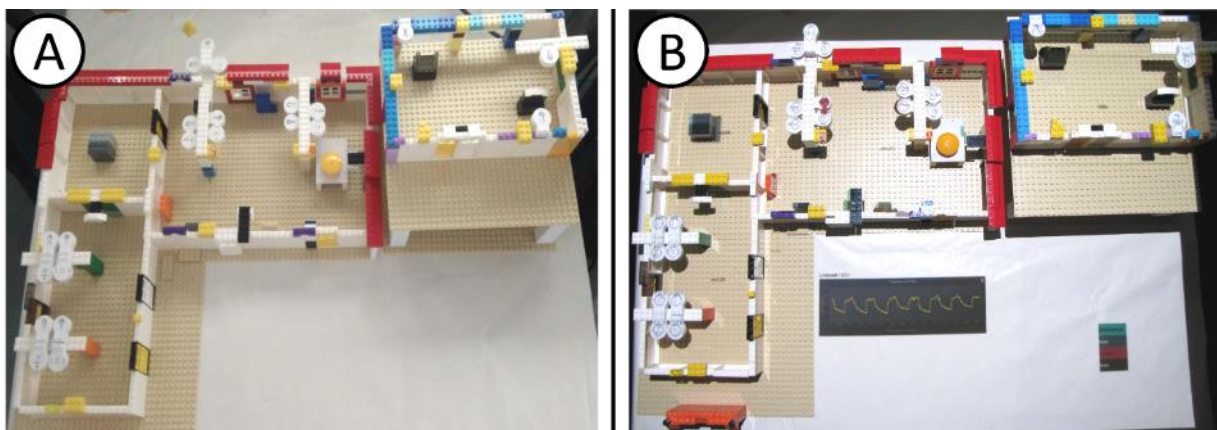


Figure 4 – La maquette physique créer pour observer des données de capteurs installés dans des salles de classe

3 Bibliothèque universitaire

La bibliothèque universitaire est un lieu de travail privilégié des étudiants. Maintenir un environnement calme est alors indispensable. À ce jour, à la bibliothèque universitaire de Rangueil, la gestion du niveau de bruit est effectuée par le personnel de la bibliothèque

universitaire. Le personnel intervient s'il détecte du bruit pendant une de leur « ronde » ou si un étudiant vient leur faire remarquer.

Le projet BuBruit initié par neOCampus propose de concevoir des outils technologiques pour améliorer la qualité de vie des étudiants au sein de la bibliothèque universitaire. Pour ce faire, nous avons proposé deux solutions pour : 1) faire prendre conscience aux étudiants du bruit qu'ils émettent et 2) aider le personnel à gérer le bruit dans la bibliothèque.

3.1 Faire prendre conscience aux étudiants du bruit qu'ils émettent

Ce projet a été réalisé en collaboration avec quatre étudiants du master IHM de l'Université Toulouse III – Paul Sabatier. L'idée de ce projet est de concevoir un système interactif permettant aux étudiants de prendre conscience du bruit qu'ils font et du désagrément que cela peut causer. Le système développé devait permettre de changer le comportement à court terme des utilisateurs. Pour ce faire, les étudiants ont mené des entretiens au sein de la bibliothèque universitaire et ont fait passer des questionnaires à plus d'une centaine d'étudiants. L'analyse de ces entretiens et de ces questionnaires révélait que de nombreux étudiants semblaient être dérangés par le bruit et que certains étudiants n'étaient pas conscients du bruit qu'ils faisaient en travaillant. L'outil développé devait donc permettre de signaler aux étudiants bruyants que leur comportement dérangeait possiblement les étudiants proches d'eux. Les étudiants ont alors conçu un système persuasif ambient permettant de représenter le niveau de bruit à court et moyen terme causé par les étudiants travaillant au sein de la bibliothèque.

Ce système permettait de détecter le niveau de bruit d'une table et d'afficher plusieurs variables telles que le niveau de bruit instantané, l'historique du niveau de bruit et le nombre de personnes installées sur les tables environnantes. Ce système était conçu pour s'afficher sur toutes les tables de la bibliothèque universitaire pour que chaque usager de la bibliothèque puisse prendre conscience du bruit qu'il émettait. Les étudiants ont proposé d'utiliser une métaphore d'un bassin aquatique pour représenter les données. Les poissons représentaient les étudiants en train de travailler. S'il n'y a pas de bruit, les poissons se déplaçaient normalement. Dès qu'un bruit intervenait, une onde apparaissait (à l'instar d'un caillou qui tomberait dans l'eau). Cette onde perturbait le mouvement des poissons qui devenaient alors plus rapides et s'éloignaient de la zone où le bruit était apparu. Le bassin aquatique était séparé en 9 zones. La zone centrale

représentait la table où était projetée la visualisation et les 8 autres zones représentaient les tables adjacentes à celle du milieu. Une illustration de ce système se trouve en Figure 5.

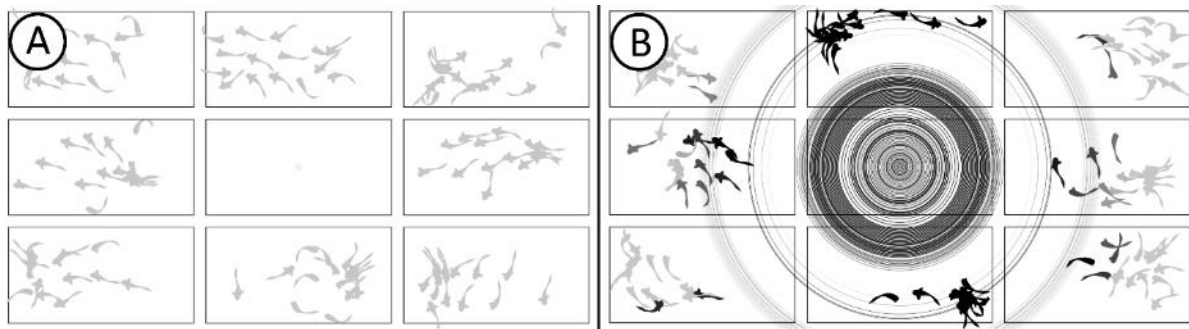


Figure 5 – Le système persuasif de réduction du bruit avec (A) quand le système ne détecte aucun bruit et (B) quand le système détecte un bruit

3.2 Gestion du bruit dans la bibliothèque par le personnel

En complément du projet précédent, un outil de gestion de la bibliothèque par le personnel a commencé à être développé par un étudiant de dernière année d'école d'ingénieurs que nous avons encadré. À l'instar du projet précédent, les besoins de l'outil ont été extraits à la suite de plusieurs entretiens avec le personnel de la bibliothèque. L'analyse de ces entretiens a permis de faire émerger plusieurs exigences auxquelles le système devait répondre :

- Pouvoir accéder aux détails des données de bruit (D'où provient le produit ? Depuis combien de temps le bruit est-il présent ? Quel type de son en est la source ?)
- La navigation dans l'espace d'information (Visualisation d'historique du bruit, analyse d'une période de bruit, etc.)
- L'activation et désactivation manuelle de la visualisation indiquant le niveau de bruit

Cette solution utilisait sur une maquette physique qui devait venir supporter les trois exigences présentées ci-dessus. La maquette physique avait l'avantage d'être : 1) un dispositif à part qui pouvait se trouver dans l'environnement du personnel (par ex. sur la table d'accueil) et qui ne requérait pas une attention constante de la part du personnel et 2) un objet physique qui permettait au personnel de localiser plus facilement la source du bruit et d'analyser les données relatives à ce bruit. Malheureusement, ce projet ayant dû se faire dans une des périodes de confinement, nous avons pu uniquement concevoir plusieurs preuves de concept pour, à la fois, sélectionner un élément de la maquette (par ex. une table) et sélectionner un item dans un menu.

Pour la sélection d'un item dans un menu, nous avons voulu explorer l'effet de la disposition et de l'ancrage du menu dans le monde physique. En effet, afficher un menu proche d'un référent physique n'est pas encore un domaine très étudié. Nous avons alors concentré nos travaux sur

la conception de nombreux types de menu selon les propriétés suivantes : la forme du menu de premier niveau, la forme du menu de second niveau et le niveau de physicalité du menu. Cette dernière propriété permettait d'explorer la physicalité du menu en fonction du point d'intérêt (par ex. une table dans la bibliothèque). Les quatre types de physicalité explorés étaient : Virtuel, Réalité augmentée sur smartphone, Réalité augmentée ancrée sur un objet tangible extérieur à la maquette, Réalité augmentée ancrée à côté de l'objet d'intérêt (c'est-à-dire sur la maquette) (cf. Figure 6).

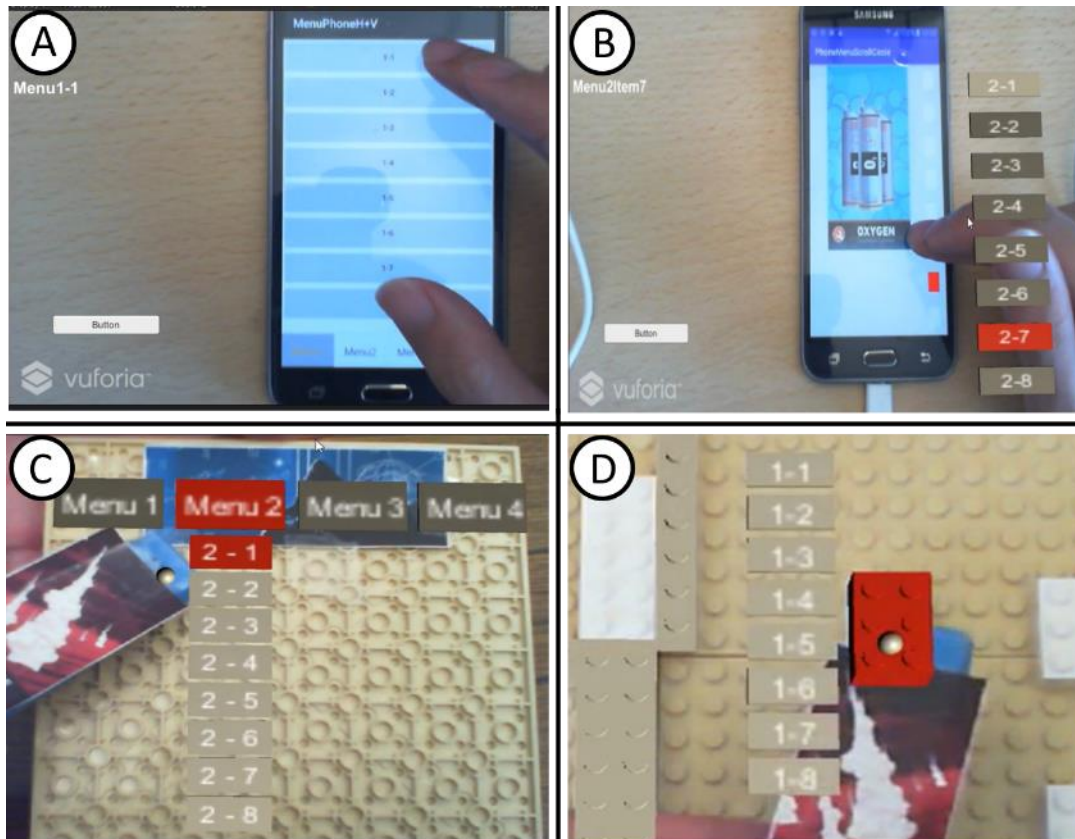


Figure 6 – Les quatre types de physicalité proposés dans ce projet avec (A) Virtuel, (B) Réalité augmentée sur smartphone, (C) Réalité augmentée ancrée sur un objet tangible et (D) Réalité augmentée ancrée à côté de l'objet d'intérêt

Résumé de thèse

Durant cette dernière décennie, la quantité de données numériques n'a cessé d'augmenter. Ces données provenant de l'Internet des Objets ou de smartphones permettent aux utilisateurs de comprendre et d'analyser leur santé, leur mode de vie ou encore leur consommation énergétique, et aux différentes institutions (villes, campus, promoteurs immobiliers) de gérer le flux de population, le trafic routier ou les pertes énergétiques dans les bâtiments administrés. Le domaine de la « ville intelligente » en tire particulièrement profit. Il est donc capital de mettre à disposition des visualisations interactives de ces données pour pouvoir faire émerger des connaissances et faciliter la prise de décision.

Une approche récente propose de représenter ces données proches de l'objet qui les capte ou les produit. Cependant, ce domaine est encore assez récent et il reste de nombreux défis à explorer. Un des défis majeurs identifié par Thomas et collab. (Thomas et al., 2018) est le développement et la conception de techniques d'interaction pour faciliter l'analyse de données situées.

L'objectif de nos travaux de thèse est de concevoir et d'évaluer des solutions interactives tirant profit du référent physique pour interagir avec des données numériques situées en se focalisant sur une tâche interactive fondamentale en IHM : la sélection. Dans ce manuscrit nous abordons cette tâche selon une approche fonctionnelle et une approche conceptuelle.

Abstract

Over the past decade, the amount of digital data has been increasing. This data from the Internet of Things or smartphones allows users to understand and analyse their health, lifestyle or energy consumption, and different institutions (cities, campuses, real estate developers) to manage the flow of people, road traffic or energy losses in the buildings they manage. One particular topic that benefits from this is the “smart city”. It is therefore essential to make interactive visualisations of this data available in order to generate knowledge and facilitate decision-making.

A recent approach proposes to represent these data close to the object that captures or produces them. However, this field is still quite new and there are still many challenges to explore. One of the major challenges identified by Thomas et al. (Thomas et al., 2018) is the development and design of interaction techniques to facilitate the analysis of situated data.

The objective of our thesis work is to design and evaluate interactive solutions leveraging the physical referent to interact with situated digital data by focusing on a fundamental interactive task in HCI: selection. In this manuscript we study this task from both a practical and a conceptual approach.