



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*
Discipline ou spécialité : *Informatique*

Présentée et soutenue par *Nehla Ghouaiel*
Le 23 janvier 2014

Titre : *Ingénierie de la conception de systèmes de réalité augmentée mobile*
Applications de la réalité augmentée mobile au tourisme culturel

JURY

<i>Gaëlle Calvary</i>	<i>Professeur</i>	<i>Rapporteur</i>
<i>Guillaume Moreau</i>	<i>Professeur</i>	<i>Rapporteur</i>
<i>Samir Otmane</i>	<i>Professeur</i>	<i>Président</i>
<i>Jean-Marc Cieutat</i>	<i>Maître de Conférence (HDR)</i>	<i>Co-directeur</i>
<i>Jean-Pierre Jessel</i>	<i>Professeur</i>	<i>Co-directeur</i>
<i>Sébastien Bottecchia</i>	<i>Maître de Conférence</i>	<i>Examineur</i>
<i>David Oyarzun</i>	<i>Responsable de département</i>	<i>Examineur</i>

Ecole doctorale : *Mathématique Informatique TéléCommunication Toulouse (MITT)*
Unité de recherche : *Institut de Recherche en Informatique de Toulouse (IRIT)*
Directeur(s) de Thèse : *Jean-Pierre Jessel, Jean-Marc Cieutat*
Rapporteurs : *Gaëlle Calvary, Guillaume Moreau*

Résumé

Le système M.A.R. T.S a été conçu avec l'objectif de proposer un nouveau mode de médiation au musée. Lorsque la visite au musée se déroule sans guide conférencier, les outils de médiation classiques comme l'audio-guide ou les étiquettes montrent rapidement leurs limites en termes d'interaction humaine. Nous proposons donc de permettre au visiteur d'interagir avec un guide humain virtuel appuyé de nos paradigmes d'interaction « Sélection Documentation Reconstruction », puisant leurs principes de la réalité augmentée. L'objectif, par le biais de nos propositions, est de créer les conditions propices à un contact riche avec la collection, et de donner au visiteur les outils nécessaires pour qu'il puisse vivre une expérience d'apprentissage riche au musée. Les expérimentations que nous avons menées vont dans ce sens, montrant l'efficacité de ce nouveau mode de médiation. Qui plus est, les principes d'interactions sont suffisamment généraux pour dépasser le cadre original de la visite guidée au musée et peuvent être exploitée pour une visite guidée à l'extérieur.

Mots clés : Réalité augmentée mobile, vision par ordinateur, agent conversationnel 3D, informatique mobile.

**Design Engineering of Mobile
Augmented Reality Systems :
Applications for Cultural Tourism**

Abstract

M.A.R.T.S system was built in order to propose a new mode of museum mediation. When the museum visit take place without a human guide, classical mediation tools such as labels or audio-guide rapidly show their limits in terms of human interaction. We therefore propose to allow visitors to interact with a virtual human guide supported by our interaction paradigms "Selection Documentation Reconstruction", which are based on augmented reality. The aim behind our proposals is to ensure favorable conditions for a rich contact with collections, and provide the visitor with necessary tools enabling him to live a worthwhile learning experience in the museum. The experiments we have conducted go in this direction, showing the effectiveness of this new mode of museum mediation. Moreover, interactions principles are sufficiently general to be applied to other cultural heritage activities such as outdoor guided tours.

Keywords:

Mobile augmented reality, computer vision, 3D conversational agent, mobile computing.

Remerciements

Je souhaite remercier en premier lieu mon co-directeur de thèse, M. Jean-Marc Cieutat, maître de conférences (HDR) du laboratoire ESTIA recherche, pour m'avoir accueilli au sein de son équipe. Je lui suis également reconnaissante pour le temps conséquent qu'il m'a accordé, ses qualités pédagogiques et scientifiques, sa grande patience, sa franchise et sa sympathie. J'ai beaucoup appris à ses côtés et je lui adresse ma gratitude pour tout cela.

J'adresse de chaleureux remerciements à mon co-directeur de thèse, M. Jean-Pierre Jessel, professeur de l'université Paul sabatier de Toulouse, pour son attention de tout instant sur mes travaux, pour ses conseils avisés et son écoute qui ont été prépondérants pour la bonne réussite de cette thèse.

J'adresse de sincères remerciements à mes rapporteurs , Mme. Gaëlle Calvary, professeur de l'institut national polytechnique de Grenoble de et à M. Guillaume Moreau, professeur de l'École Centrale de Nantes pour avoir accepté de rapporter ce mémoire de thèse et pour l'intérêt qu'ils ont porté à mon travail.

Je suis très honorée que M. Samir Otmane ait accepté de présider mon jury de thèse.

Je souhaite remercier très sincèrement mes examinateurs, M. David Oyarzun et M. Sébastien Bottecchia pour l'intérêt qu'ils ont porté à mon travail.

Je tiens à remercier tous les membres des équipes ESTIA recherche et IRIT, pour leur aide, leur soutien et leur amitié.

Enfin, je remercie mes parents, ma famille et mes amis pour leur soutien au cours de ces trois années et sans lesquels je n'en serais pas là aujourd'hui.

Liste des Abréviations

HMD	Head Mounted Display
IHM	Interface Homme Machine
MAR	Mobile Augmented Reality
POSIT	Pose from Orthography and Scaling with Iterations
RA	Réalité Augmentée
SDK	Software Development Kit
SURF	Speeded Up Robust Feature (algorithme)

Table des matières

Résumé	i
Abstract	v
Remerciements	vii
Table des matières	xi
Table des figures	xv
Liste des tableaux	xvii
INTRODUCTION GENERALE	1
1 Fondements de la Réalité Augmentée	7
1.1 La réalité augmentée	8
1.1.1 Terminologie	8
1.1.2 Définition du concept de la réalité augmentée	8
1.2 Notre définition de la réalité augmentée	11
1.3 Réalité augmentée Mobile	12
1.4 Fondements de la Réalité Augmentée	12
1.4.1 Architecture matérielle d'un système de RA	12
1.4.2 Composants technologiques d'un système de RA	13
1.4.3 Reconnaissance	15
1.4.3.1 Approche globale	15
1.4.3.2 Approche locale	16
1.4.4 Pose	17
1.4.5 Suivi	18
1.4.5.1 La technologie Radio	19
1.4.5.2 La technologie Ultrason	20

1.4.5.3	La technologie inertielle	20
1.4.5.4	La technologie optique	21
1.4.6	Recalage	26
1.4.7	Rendu	26
1.4.8	Modèles d'affichage	27
Partie 1		31
2	Le visiteur, la visite guidée au musée et la réalité augmentée	33
2.1	Le visiteur	33
2.1.1	Les motivations de visite	33
2.1.2	L'expérience visiteur	34
2.2	Apprendre au musée	37
2.2.1	L'apprentissage	38
2.2.2	Les théories cognitivistes d'apprentissage	40
2.2.3	La connaissance	41
2.3	L'expérience d'apprentissage au musée	42
2.3.1	L'expérience	42
2.3.2	L'expérience d'apprentissage	44
2.3.3	Mesurer l'expérience d'apprentissage	45
2.4	Le musée : le rôle de médiation	47
2.4.1	Définition de la médiation	47
2.4.2	Les problèmes liés à la médiation dans les musées	48
2.5	Les moyens classiques de la médiation au musée	50
2.5.1	Guide conférencier-animateur	50
2.5.2	Audio Guide	50
2.5.3	Les textes	50
2.5.4	Conclusion sur les moyens classiques de la médiation	51
2.6	L'intérêt de la Réalité Augmentée	52
2.7	Les systèmes actuels de R.A pour la visite guidée	53
2.7.1	Conclusion sur les systèmes de RA	60
2.8	Naissance d'une problématique	60
3	La visite guidée au musée : Propositions	63
3.1	Les modes de représentation du savoir	63
3.2	Propositions pour soutenir l'apprentissage au musée	65
3.2.1	Guide humain virtuel	65
3.2.1.1	Importance de simulation de la présence humaine dans un musée	65
3.2.1.2	Qu'est-ce qu'un guide humain virtuel?	66

3.2.1.3	Communication verbale et non verbale	68
3.2.2	Paradigmes d'interaction	70
3.2.2.1	Sélection	70
3.2.2.2	Reconstruction	73
3.2.2.3	Documentation	73
3.2.2.4	Adaptations de l'interaction à l'environnement de l'utilisateur	74
3.3	Conclusion	76
4	Système M.A.R.T.S : Mobile Augmented Reality Touring System (Concep- tion et Expérimentation)	77
4.1	Le système M.A.R.T.S	78
4.1.1	Rappel de fonctionnement	78
4.1.2	Diagramme de déploiement	79
4.1.3	Choix d'implémentation	81
4.1.4	Vue d'ensemble de l'architecture interne	82
4.1.4.1	Diagramme de relation	82
4.1.4.2	Description du rôle des composants	82
4.2	M.A.R.T.S : implémentation du prototype expérimental	85
4.2.1	Reconnaissance	85
4.2.1.1	Algorithme	86
4.2.1.2	Évaluation et Résultats	89
4.2.2	Le suivi	90
4.2.3	Guide humain virtuel	91
4.2.4	Paradigmes d'interaction	93
4.2.4.1	Sélection	93
4.2.4.2	Reconstruction	95
4.2.4.3	Documentation	95
4.2.4.4	Interface utilisateur	95
4.2.4.5	Adaptations de l'interaction à l'environnement de l'utilisateur	97
4.3	Évaluation expérimentale	98
4.3.1	Ce que l'on cherche à évaluer	98
4.3.2	Description de l'évaluation	99
4.3.2.1	Descriptions des objets exposés et des moyens de communication	99
4.3.2.2	Sujets et plan de recherche	100
4.4	Analyse et interprétations	100
4.4.1	Méthode d'analyse des résultats	101
4.4.1.1	Nombre de réponses correctes	101
4.4.1.2	Les données du questionnaire	101

4.4.2	Résultats et interprétations	102
4.4.2.1	Nombre de réponses correctes	102
4.4.2.2	Les données du questionnaire	104
4.4.2.3	Discussion des résultats obtenus	107
Partie 2		109
5 M.A.R.T.S : Vers un Système de RAM pour la visite guidée à l'extérieur		111
5.1	Visite guidée à l'extérieur	112
5.2	La navigation	112
5.2.1	Principe	113
5.2.2	Implémentation	116
5.2.3	Etude Expérimentale	118
5.3	M.A.R.T.S en environnement urbain	124
5.4	Reconnaissance du patrimoine architectural	125
5.4.1	Algorithme	126
5.4.2	Géométrie des points et des lignes de fuite	126
5.4.3	Détection des points de fuite	128
5.4.4	Alignement	131
5.4.5	Représentation	134
5.4.6	Mesure de Similarité	135
5.4.7	Expérience et Résultats	136
Conclusion et perspectives		139
Annexes		145
A Fichier KML		147
B Projet Aquitaine Euskadi		149
C SDK de Réalité Augmentée Mobile		153
Bibliographie		159

Table des figures

1.1.1	Continuum entre réalité et virtualité	9
1.4.1	Briques technologiques "building blocks" d'un système de RA . . .	13
1.4.2	Constitution technologique d'un système de RA modifiée	15
1.4.3	Projection orthographique mise en œuvre par POSIT	18
1.4.4	marqueurs ARToolKit de type « template »	22
1.4.5	Structure d'un QR Code	23
1.4.6	Boucle de système de Suivi de RAPID [Klein 03]	24
1.4.7	Occultation contradictoire dans l'image gauche [Fukiage 12]	28
1.4.8	video see-through HMD [Azuma 97]	29
1.4.9	optical see-through HMD [Azuma 97]	29
2.2.1	Le modèle constructiviste d'apprentissage au musée de Hein	41
2.3.1	Generic Learning Outcomes (GLO)	46
2.7.1	NaviCam fournit des informations textuelles à l'utilisateur	54
2.7.2	Virtuoso	55
2.7.3	Navigation dans les différents thèmes d'une œuvre d'art et accès à l'interprétation	55
2.7.4	A gauche le système « salle de présentation », à droite le système de guidage	56
2.7.5	UMPC équipé d'une webcam	57
2.7.6	Une fenêtre en RA qui fait apparaître le repentir « Pentimento » à l'utilisateur	58
2.7.7	Déchiffrer les codes hiéroglyphes à l'aide de Passport	59
3.2.1	Principe de « Sélection » montrant l'arc de triomphe provisoire à l'entrée de la place Réduit de Bayonne	72
3.2.2	Chaland monoxyle de moyen âge	73
3.2.3	Le principe de la "Documentation" virtuelle	74
4.1.1	Principe de Fonctionnement	79
4.1.2	Design général de M.A.R.T.S	80

TABLE DES FIGURES

4.1.3	Vue d'ensemble de l'architecture basée composants du système M.A.R.T.S.	83
4.2.1	Comparaison avec FLANN	90
4.2.2	Flux de données	91
4.2.3	Structure du moteur d'avatar	92
4.2.4	Mise en œuvre de la « Sélection »	94
4.2.5	Interface utilisateur de M.A.R.T.S	96
4.4.1	Table de moyennes	103
4.4.2	Effet inter-sujet sur le nombre de réponses correctes	103
4.4.3	Notes moyennes pour chaque mode en fonction des questions.	104
4.4.4	Degré d'accord entre les différents sujets sur le classement des sys- tèmes	105
5.2.1	Prototypes possibles	114
5.2.2	Schéma tactile de Pocket Navigator citePielot :2010	114
5.2.3	Changement d'orientation	115
5.2.4	Prototype matériel d'HaptiNav	117
5.2.5	Ceinture vibro-tactile	117
5.2.6	Structure de la partie applicative de système	118
5.2.7	HaptiNav	118
5.2.8	Trajet emprunté lors des expérimentations	119
5.2.9	Erreurs et désorientations	121
5.2.10	Diagramme de moyenne d'erreurs	121
5.2.11	Diagramme de nombre des désorientations	122
5.4.1	Géométrie des points de fuite [Kim 06]	128
5.4.2	Géométrie des lignes de fuite [Lai 09]	129
5.4.3	Géométrie realtive à l'erreur	130
5.4.4	Alignement de point de vue. De la droite vers la gauche : l'image test, l'image modèle, l'image rectifiée	134
5.4.5	Construction de modèle de lignes d'un bâtiment	136
A.1	Structure d'un fichier KML	148
C.1	. Architecture logicielle globale du SDK MAR	155
C.2	Liens entre les composants du SDK	156

Liste des tableaux

INTRODUCTION GENERALE

Contexte

Tourisme culturel

La culture et le tourisme, étroitement liés, sont de nature à renforcer l'attractivité et la compétitivité de régions et de pays. De plus en plus, la culture est une composante importante du produit touristique parce qu'elle permet de se différencier sur un marché mondial très encombré¹. En ce sens, La culture est un instrument de valorisation des territoires dans la concurrence interurbaine. Selon l'ICOMOS², le tourisme culturel signifie : « une forme de tourisme dont l'objectif est de permettre, entre autres, la découverte des monuments et des sites. ».

Par ailleurs, les études réalisées³ par le cluster Goazen⁴ auprès des visiteurs du pays basque, mettent en exergue l'importance du tourisme culturel. En effet, les visites liées au patrimoine culturel constituent 47% de l'ensemble des activités pratiquées par les touristes et se classent ainsi au deuxième rang. Les musées et les monuments historiques constituent les principaux sites culturels fréquentés par les touristes.

Plusieurs travaux portant sur le tourisme, à l'instar de Bourgère [Bourgère 12], défendent une dimension éducative du tourisme culturel. Dans cette optique, les pratiques touristiques sont considérées comme un large espace d'apprentissages pour le touriste. De plus, d'autres chercheurs comme Van Winkle et al. [Van Winkle 12],

1. OECD (2009), The Impact of Culture on Tourism, Paris.
<http://www.oecd.org/fr/cfe/tourisme/>

2. ICOMOS (International Scientific Committee on Cultural Tourism.)
<http://www.icomos.org/tourism/>

3. Etude qualitative des clientèles touristiques en Pays Basque (2010), consultable depuis le site : <http://www.bayonne.cci.fr/>

4. Cluster tourisme du Pays Basque. <http://www.bayonne.cci.fr/GOazen/Reseau-GOazen.html>

mettent l'accent sur une expérience d'apprentissage naissant de l'activité touristique culturelle.

Au musée, la recherche de connaissances et d'émotions s'attache à une expérience d'apprentissage tel que défini par [Falk 00][Falk 05]. Cet axe sera traité dans la première partie de ce mémoire. D'un autre côté, la visite des villes et des pays d'arts et d'histoire, riches en monuments historiques,⁵s'inscrit dans une quête du patrimoine culturel, historique et architectural. C'est donc à l'activité touristique à l'extérieur que nous allons nous intéresser à la deuxième partie de ce mémoire.

Tourisme culturel et numérique

Le tourisme constitue un secteur essentiel, pour l'économie française, de par son importance et de son impact sur d'autres secteurs d'activités qui lui sont connexes [CNT 10]. Face à ce constat, les offices de tourisme ont engagé depuis quelques années une profonde réflexion pour répondre aux nouvelles attentes de leurs clientèles à travers l'exploitation de plusieurs stratégies. En animant les visites et en engageant la communication avec les touristes, qui deviendront plus tard des ambassadeurs spontanées de la destination, l'office de tourisme participe dynamiquement à la promotion du secteur touristique⁶.

En ce sens, mettre le numérique au service du tourisme et des touristes, représente une stratégie par laquelle les professionnels du tourisme visent à renforcer l'attractivité de leur destination. Avec le développement des technologies de l'information et de la communication (TIC), le tourisme numérique ou le « eTourisme » voit le jour. Les sites web, les applications pour téléphones mobiles, les bornes d'information interactives, les vitrines interactives, les audio-guides, entre autres, forment aujourd'hui la panoplie des outils numériques permettant de valoriser l'offre touristique [Mante 13]. Dans cette optique, les musées accordent à la médiation au public, une attention fondamentale, bien devant les missions de conservation, d'étude ou d'acquisition. De plus, les différentes formes de mises en scène numériques du patrimoine culturel, sont représentatives de l'effort concédé par les offices de tourisme pour la valorisation de celui-ci.

5. Selon le ministère de la culture et de la communication, Bayonne fait partie des villes et des pays d'art et d'histoire. Le réseau français regroupe 179 villes et pays d'art et d'histoire attachées à la valorisation et à l'animation de l'architecture et du patrimoine

6. Résumé des propos du Jean-Luc Boulin, directeur de la Mission OTSI et Pays Touristiques d'Aquitaine (MOPA), dans un interview accordé au magazine de l'aquitaine numérique (édition du mai-juin 2011).

Problématique : améliorer l'expérience d'apprentissage vécue par un touriste

Un touriste voyageant sans guide-accompagnateur, libre de choisir ses excursions et acteur de son voyage, a besoin d'aide. Il est évident qu'être touriste, c'est avant tout se déplacer hors de son environnement habituel. Or, pour évoluer dans un nouvel environnement, il faut être capable de le comprendre de mieux en mieux jusqu'à avoir l'impression qu'il lui devienne familier. A l'extérieur, se trouver en face d'un bâtiment historique méconnu, laisse naître le besoin de découvrir et d'appréhender tous ses différents aspects. Apprendre de son voyage se révèle être donc un besoin et une activité qui caractérisent le mieux ce l'on appelle le tourisme culturel.

Par ailleurs, assister le touriste dans sa visite au musée, lui transmettre les connaissances propres aux œuvres d'art, rendre les visites encore plus attrayantes constituent les rôles essentiels que pourraient remplir un guide touristique. En l'absence de ce dernier, un système de guide numérique doit donc retranscrire pleinement ces rôles.

Avec la réalité augmentée, nous pouvons désormais envisager un système d'aide à la visite. Notre recherche va donc être axée sur la possibilité d'avoir un système d'aide à la visite permettant de transmettre efficacement de nouvelles connaissances au visiteur. Nous nous sommes alors appuyés sur la réalité augmentée pour proposer de nouvelles modalités de communication de savoir expérimentées ici dans le cadre d'activités de visite au musée.

Structure du mémoire

La structure globale de ce mémoire se présente sous deux axes différents. Le premier concerne l'amélioration de l'expérience vécue par le touriste en tant que visiteur du musée ; Il est développé dans les chapitres deux, trois et quatre. Le second axe se rapporte aux activités touristiques pratiquées à l'extérieur, notamment en milieux urbains riches en bâtiments patrimoniaux. Cette partie sera développée dans le chapitre 5. Voici donc les détails des différents chapitres de ce mémoire :

Chapitre 1 Ce chapitre permettra de retracer l'évolution de la définition de la réalité augmentée. À l'issue de cet historique, nous proposerons au lecteur notre propre définition afin d'établir une compréhension commune de ce terme au fil du mémoire. Nous nous consacrerons ensuite à présenter les fondements technologiques qui constituent la réalité augmentée ; l'intention

étant de nous éclairer sur ses principes nous ayant permis de proposer notre solution.

Chapitre 2 Quels sont les problèmes liés à la médiation au musée ? Pourquoi certaines solutions de médiation sont-elles inefficaces ? Quel pourrait être l'intérêt d'introduire les nouvelles technologies comme la réalité augmentée dans le processus d'assistance à la visite au musée ? Est-ce que les propositions actuelles permettent de soutenir efficacement l'expérience d'apprentissage du visiteur ? C'est à ces questions que le chapitre 2 va répondre, et les manquements constatés nous amèneront à formuler notre problématique.

Chapitre 3 Dans ce chapitre nous détaillerons les différents modes de représentation du savoir. Etant donné les préférences de chaque individu pour un mode et/ou pour l'autre, Hooper-Greenhill [Hooper Greenhill 94] insiste sur l'importance de rendre compte des différents styles d'apprentissages des visiteurs. C'est à la lumière de ces observations que nous énoncerons nos propositions, en termes de perception et d'interaction, visant à « simuler » les différents modes de représentation de connaissances, et cela, grâce à la réalité augmentée.

Chapitre 4 Dans ce chapitre, nous allons présenter le système M.A.R.T.S, système se révélant être l'implémentation de nos propositions de perception et d'interaction. Nous avons été également en mesure d'expérimenter une partie de nos principes et nous en présenterons les premiers résultats.

Chapitre 5 Enfin, dans le chapitre 5 nous montrerons que la visite guidée à l'extérieur est largement différente de la visite guidée au musée. Partant de ce constat, nous présenterons au chapitre 5 comment étendre le système M.A.R.T.S à un fonctionnement à l'extérieur.

Lecture du mémoire

Afin de faciliter la lecture du mémoire, des symboles sont ajoutés dont voici la légende :

- Ce symbole précède une réflexion ou une discussion. 
- Vous trouverez ce symbole lorsqu'il sera conseillé au lecteur d'approfondir le sujet par une lecture bibliographique. 
- Ce symbole précède une information importante à retenir pour la bonne compréhension du mémoire. 

— Vous trouverez ce symbole lors de la présence de définitions.



Chapitre 1

Fondements de la Réalité Augmentée

Introduction

La réalité augmentée est un concept qui a fait ses premières apparitions il y a près de 50 ans chez Sutherland [Sutherland 65] [Sutherland 68]. En revanche, ce concept n'a connu sa vraie évolution qu'à la dernière décennie, grâce à la miniaturisation des composants électroniques, à l'augmentation de la puissance de calcul et à l'autonomie des ordinateurs.

La réalité augmentée est par essence transdisciplinaire. Dans le domaine des sciences et des techniques, la réalité augmentée appartient principalement au domaine des STIC (Sciences et Techniques de l'Information et de la Communication). Cependant, elle n'est pas entièrement incluse dans ce cadre. De nombreuses disciplines rentrent donc en jeu dans les avancées du domaine : la vision par ordinateur, l'informatique, l'électronique, l'ergonomie, la psychologie, etc.

L'objectif de ce chapitre n'est pas de faire un état de l'art exhaustif sur la réalité augmentée, mais plutôt de nous éclairer sur ses principes et ses technologies nous ayant permis de proposer notre solution. Ainsi, ce chapitre pourrait se découper en deux parties :

- La première retrace la réalité augmentée à travers l'évolution de ses définitions depuis sa création. À l'issue de cet historique, nous proposerons alors au lecteur notre propre définition du terme « réalité augmentée ».
- La deuxième partie quant à elle, vise à présenter les principes technologiques indispensables à ce domaine.

1.1 La réalité augmentée

1.1.1 Terminologie

L'expression réalité augmentée est une association de deux mots « réalité » et « augmentée ». Les définitions étymologiques des deux termes, tels que présentées dans le dictionnaire de Larousse, sont les suivantes :

- Réalité : n.f (latin médiéval *realitas*, du latin classique *realis*, de *res*, chose)
1. Caractère de ce qui est réel, de ce qui existe effectivement. 2. Ce qui est réel, ce qui existe en fait, par opposition à ce qui est imaginé, rêvé, fictif.
- Augmenter : v.tr (bas latin *augmentare*, du latin classique *augere*, accroître)
1. Rendre quelque chose plus grand, plus considérable. 2. Rendre quelque chose plus important, plus intensif.

En français, le terme est un non-sens puisque, puisqu'il est sémantiquement incorrecte de vouloir augmenter la réalité. Toutefois, le grand public s'étant habitué à cette appellation, il est donc vain aujourd'hui de tenter de la changer. A l'origine, l'expression "réalité augmentée" existe depuis les années 1990. Elle dérive de la traduction anglaise de "augmented reality", proposée par Caudell et Mizell [Caudell 92] [Milgram 94] lorsqu'ils qualifient une application permettant d'aider des techniciens pour certaines opérations de maintenance dans le domaine aéronautique. Comme le fait remarquer Olivier Hugues [Hugues 11], la finalité de la réalité augmentée, laisse voir que, de point de vue logique, c'est plutôt la perception de la réalité que l'on cherche à augmenter, soit en restant fidèle aux règles de la réalité, soit en modifiant le réel par l'imaginaire.

1.1.2 Définition du concept de la réalité augmentée

Le terme réalité augmentée a été employé pour la première fois par Caudell et al. [Caudell 92] pour désigner l'acte de superposition d'informations numériques (générées par ordinateur) sur le monde réel. Ces auteurs remarquent que l'augmentation visuelle vise à aider l'utilisateur dans sa tâche en cours. « This technology is used to "augment" the visual field of the user with information necessary in the performance of the current task, and therefore we refer to the technology as "augmented reality" (AR) » ([Caudell 92], page 660).

Il est à noter que c'est qu'en 1994 que l'on a vu s'éveiller une discussion sur la définition de la RA ; celle-ci a été lancée par Milgram et al. [Milgram 94]. Ces auteurs ont signalé que jusqu'à cette époque-là les chercheurs continuaient à utiliser l'expression « Réalité Augmentée » sans véritablement la définir. A ce propos, ils

évoquent deux définitions employées par le journal « *Telemanipulator and Telepresence Technologies* » en 1994 :

1. « Augmenter la rétroaction naturelle de l'utilisateur avec des indices virtuels ».
2. « une forme de la réalité virtuelle qui présente aux participants une vision claire du monde réelle grâce à un afficheur téléporté¹ transparent ».

Comme le fait remarquer Bottecchia [Bottecchia 10], dans l'une ou l'autre des définitions, rien ne laisse entrevoir la relation qui doit exister entre le réel et le virtuel. De plus, la deuxième définition est clairement orientée vers la modalité visuelle et exclut donc toute autre modalité. D'un autre côté, il est à noter que dans Milgram et al. [Milgram 94], ceux-ci n'ont pas proposé de définition pour la RA. Cependant, ils cherchaient plutôt à établir une relation entre la RV et la RA. Dans cette optique, ils ont décrit un continuum entre le monde réel et le monde virtuel, appelé réalité mixte, où la réalité augmentée évolue, proche du monde réel tandis que la virtualité augmentée évolue, proche du monde virtuel.



FIGURE 1.1.1 – Continuum entre réalité et virtualité

Azuma [Azuma 97] a proposé en 1997 une première définition de la réalité augmentée qu'il a ensuite complétée en 2001 [Azuma 01]. Selon lui, un système de réalité augmentée ajoute des objets virtuels au monde réel de telle sorte qu'ils semblent coexister dans le même espace que le monde réel. En plus, tout système de Réalité Augmentée doit satisfaire les trois propriétés suivantes :

- Combinaison du réel et du virtuel : les entités virtuelles en 3D doivent être intégrées au monde réel qui, lui, est naturellement en 3D.

1. Un afficheur téléporté est appelé en anglais Head Mounted Display (HMD). Cette technologie d'affichage va être présentée plus loin dans ce chapitre.

- Interactivité en temps réel : cela exclut notamment les films, même si la condition précédente est respectée.
- Recalage visuel en trois dimensions : cela permet de faire coïncider visuellement les entités virtuelles avec la réalité pour créer l'illusion de coexistence entre les deux.

Il est évident que la définition de RA d'Azuma laisse entrevoir une relation de coexistence entre le réel et le virtuel. Sans l'exprimer explicitement dans leur définition, les auteurs remarquent que leur définition de la RA peut être appliquée aux sens autres que le sens visuel. Néanmoins, cette définition s'avère restrictive vu qu'elle exclut les augmentations en deux dimensions (comme les annotations textuelles) qui ne nécessitent pas de recalage en 3D. Partant de ce constat, Fuchs et al. [Fuchs 01] ont proposé cette définition de la RA :

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

Bien que cette définition paraisse beaucoup moins restrictive que celle proposée par Azuma [Azuma 01], elle ne laisse profiler aucun lien entre les mondes réels et virtuels. Afin de mettre en avant la relation existante entre le réel et le virtuel, Bottecchia [Bottecchia 10] a proposé cette définition de la RA :

«Combinaison de l'espace physique avec l'espace numérique dans un contexte sémantiquement lié» ([Bottecchia 10], page 23).

En revanche, cette définition ne laisse rien entrevoir sur le but de la liaison entre le réel et le virtuel et donc la sur finalité de la RA en elle-même. Dans ce contexte, nous précisons au lecteur que la finalité de la RA a été définie par Hugues et al. [Hugues 11] comme suit :

« La finalité de la réalité augmentée est de permettre à une personne de réaliser des activités sensorimotrices et cognitives dans un espace mixte associant l'environnement réel et un environnement artificiel. La réalité augmentée peut proposer soit une modélisation du monde réel sur la base d'un monde qui imite ou symbolise certains aspects du monde réel, soit la création d'un monde imaginaire ne correspondant à rien d'actuel» ([Hugues 11], page 49).

Les auteurs précisent que cette finalité se décompose en deux catégories. La première se rapporte au type d'interaction que l'utilisateur entretient avec l'environnement mixte. Les termes "activité sensorimotrice" sont employés pour signifier l'idée que la personne perçoit et agit physiquement avec les entités et éléments du monde mixte, qu'ils soient artificiels ou naturels. La seconde catégorie concerne la diversité des environnements représentés dans ce monde mixte. L'environnement mixte appartenant aux deux premiers cas, imitant ou symbolisant le monde réel, est considéré comme un environnement pseudonaturel, à la différence du second cas, considéré comme un environnement imaginaire.

1.2 Notre définition de la réalité augmentée

Afin de tenir compte des manquements soulignés dans la section précédente, nous proposons notre propre définition de la RA dont voici l'énoncé :



La réalité augmentée est la **combinaison** de **données sensorielles** avec le monde réel de telle manière que, poursuivant un **but bien déterminé**, elles semblent **coexister** avec le monde réel et lui être **liées**.

Notre définition de la RA met donc en avant ces différents aspects :

1. **La multimodalité** : les données rajoutées au monde réel sont multi sensorielles.
2. **La non restrictivité** : ces données peuvent être de plusieurs natures : numériques (images de synthèse, textes virtuels ...) ou analogiques (bras à retour d'effort), en 2D ou en 3D, etc.
3. **Le lien entre le réel et le virtuel** : la combinaison du monde réel et du monde virtuel doit laisser voir pour l'utilisateur une illusion de coexistence physique et un lien sémantique.
4. **La finalité de la combinaison du réel et du virtuel** : comme mis en avant par Hugues et al. [Hugues 11], la finalité de RA est de deux ordres : le type d'interaction (sensorimoteur ou cognitif) et la représentation souhaitée (symbolique, imitative ou imaginaire).

En ce sens, notre définition n'exclut pas les définitions précédentes et les remarques de leurs auteurs, mais vient les compléter tout en tirant profit. En effet, notre définition de la réalité augmentée reprend les aspects (1), (2) et (3) mis en avant respectivement par [Azuma 01],

[Fuchs 01] et [Bottecchia 10]. **En plus, elle tient compte de l'aspect (4) défini par Hugues et al. [Hugues 11]. Celui-ci traite la finalité de la RA et met donc l'accent sur le rôle de l'utilisateur dans la construction de l'environnement augmenté. A notre connaissance, ces quatre aspects réunis ne figurent dans aucune définition précédente de la RA.**

1.3 Réalité augmentée Mobile

L'essor de l'informatique mobile a favorisé le développement des applications de réalité augmentée. Effectivement, la prolifération des ordinateurs portables et portés, devenus plus petits ces dernières années et plus accessibles dans un secteur fortement concurrentiel, et la facilité de l'accès à l'information ont permis de trouver de nouveaux cadres d'applications à la réalité augmentée. L'usage de la réalité augmentée est devenu possible en intérieur comme en extérieur. La réalité augmentée mobile vise d'ailleurs un public plus large que par le passé, comme les utilisateurs possèdent leurs propres dispositifs mobiles et savent déjà comment les manipuler.

Hollerer et al. (2004) [Hollerer 04] fixent six composants nécessaires au fonctionnement d'un système de réalité augmentée mobile :

- Plateforme mobile de calcul.
- Dispositif mobile d'affichage.
- Le suivi et l'alignement.
- Modélisation de l'environnement.
- Dispositif(s) porté(s) d'interaction et technologies d'interaction.
- Communication sans fil et technologies de stockage de données.

1.4 Fondements de la Réalité Augmentée

1.4.1 Architecture matérielle d'un système de RA

En général, l'architecture matérielle, commune à tous les systèmes de RA, fixes ou mobiles, comporte principalement quatre éléments :

- **Caméra** : permet de filmer la scène réelle
- **Ordinateur** : chargé d'effectuer les calculs se rapportant au rendu, suivi, etc.

- **Dispositif d’affichage** : permet à l’utilisateur de visualiser la scène augmentée. Il peut être un HMD [Azuma 97], l’écran d’une tablette, l’écran d’un Smartphone.
- **Capteurs** : différents capteurs peuvent être utilisés, comme les capteurs de localisation et de suivi : GPS, Wifi, centrale inertielle, etc.

Dans certains systèmes de réalité augmentée, une seule entité matérielle peut regrouper un ou plusieurs éléments parmi ceux présentés ci-dessus. Effectivement, certains smartphone ou tablettes, largement utilisés aujourd’hui par le grand public, peuvent intégrer en même temps tous ces composants matériels.

1.4.2 Composants technologiques d’un système de RA

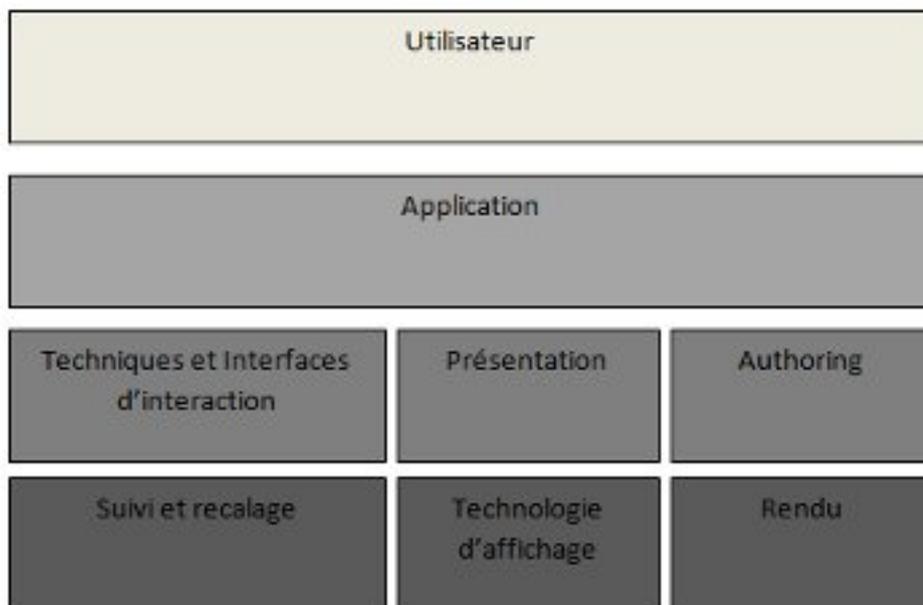


FIGURE 1.4.1 – Briques technologiques "building blocks" d’un système de RA

Bimber et al. [Bimber 05] posent dans leur livre intitulé “Spatial Augmented Reality Merging Real and Virtual Worlds”, la constitution générale d’un système de RA. La couche de base représentée dans la figure 1.4.1 décrit les composants fondamentaux d’un système de RA ; Bimber et al. les appellent également les « building blocks ». Ces composants sont : le suivi, le recalage, la technologie d’affichage et le rendu. Dans la couche située au dessus de la couche inférieure, l’on trouve des modules plus avancés à savoir : techniques et interfaces d’interaction, présentation et authoring. Au troisième niveau, se situe la couche application qui

représente l'interface utilisateur.

Couche de base : c'est le plus bas niveau du système de RA incluant les composants fondamentaux, tant logiciel que matériel, à savoir : le suivi, le recalage, les technologies d'affichage et le rendu des images générées. La plupart des efforts de recherche en réalité augmentée à l'heure actuelle se concentrent à ce niveau. Nous allons détailler de manière plus précise ces trois briques dans les sections suivantes de ce chapitre.

Deuxième Couche : ce niveau intermédiaire, principalement implémenté sous forme logicielle, s'occupe des méthodes d'interactions avec l'utilisateur, de la façon de présenter les informations virtuelles et des moyens pour les mettre en scène.

Couche application : ce niveau représente l'application de réalité augmentée en elle-même, application complètement tournée vers l'usage dévolu à la spécificité de la tâche.

Couche utilisateur : dans la mesure où une application de RA est destinée à un utilisateur final, l'acceptance de l'application est nécessaire par ce dernier. Ce niveau fait intervenir donc plusieurs disciplines en sciences cognitives et en ergonomie.

Bimber et al. [Bimber 05] soulignent que la réalité augmentée représente plus qu'une interface homme machine, vu que, la couche application est largement dépendante de la couche de base. Il est évident que la plupart des efforts de recherche en réalité augmentée à l'heure actuelle, se concentrent au niveau de cette dernière, qui demeure la plus challengeuse en comparaison avec les autres. En revanche, nous attirons l'attention du lecteur sur le fait que Bimber et al. [Bimber 05], ont omis les briques technologiques représentant la reconnaissance et le calcul de pose de la caméra. En effet, comme le soulignent Mooser et al. [Mooser 07], la majorité des travaux en RA se sont consacrés à résoudre le problème de reconnaissance d'objets et celui de pose de la caméra. La reconnaissance a pour but d'identifier l'objet en question afin de lui attribuer l'augmentation appropriée. Bien évidemment, le système de RA a besoin avant même de procéder au suivi, d'identifier la cible. D'un autre côté, la position et l'orientation de l'objet réel constituent les informations géométriques nécessaires permettant d'aligner précisément le contenu virtuel avec ce dernier.

Compte tenu de ce qui précède nous proposons de compléter le schéma de Bimber et al. de la manière suivante :

Nous allons par la suite, détailler les briques technologiques qui composent le nouveau niveau de base.

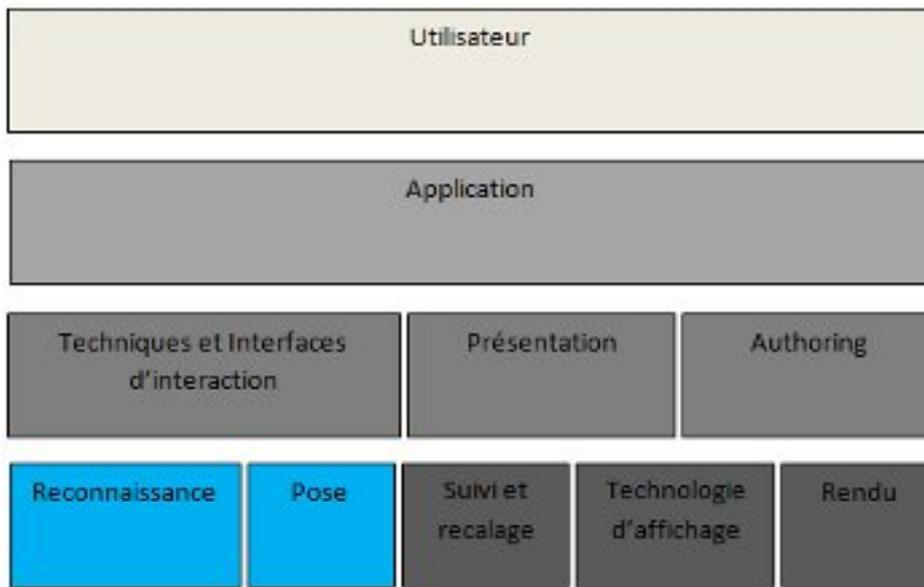


FIGURE 1.4.2 – Constitution technologique d'un système de RA modifiée

1.4.3 Reconnaissance

La reconnaissance automatique d'objets dans les images 2D est un problème difficile de la vision par ordinateur qui possède de nombreuses applications : robotique, vidéosurveillance, réalité augmentée, etc. En ce qui concerne la réalité augmentée, l'application doit identifier l'objet figurant dans la scène afin de lui attribuer la bonne augmentation. Pour la reconnaissance d'objets spécifiques, Grauman [Grauman 11] distingue deux types d'approches : la première est dite globale, la seconde est appelée locale. Ces deux approches sont présentées en détail ci-dessous.

1.4.3.1 Approche globale

Cette approche consiste à considérer l'objet dans sa totalité, en tenant compte de son apparence globale. Dans ce contexte, l'on trouve plusieurs techniques, à savoir : les moments de l'image [Hu 62], l'histogramme de couleur [Swain 91], l'Eigenface [Turk 91], le Fisherface [Peter N. 97], etc.

Hu [Hu 62] a introduit la famille d'invariant, portant son nom, en utilisant les moments géométriques. La méthode de Hu a été largement utilisée afin de reconnaître plusieurs classes d'objet indépendamment de leur orientation, taille et position.

L'histogramme de couleur mesure la fréquence avec laquelle chaque couleur apparaît dans l'image. Celui-ci permet de former une description holistique de l'image en utilisant tout simplement la distribution des couleurs.

Dans le cas de l'Eigneface, Chaque image est considérée comme un vecteur dans un espace ayant autant de dimensions que de pixels dans l'image. Les caractéristiques de l'image sont extraites par une méthode mathématique de réduction de dimensionnalité basée sur l'analyse en composante principale (ACP)². Belhumeur et al. [Peter N. 97] proposent d'améliorer la séparation de classes, assurée par l'ACP, en utilisant la méthode d'analyse discriminante linéaire de Fisher (appelé FLD ou LDA selon les communautés).

Ces techniques (appelées également descripteurs globaux) étant peu robustes face à l'occultation, plusieurs auteurs se sont tournés donc vers les descripteurs locaux.

1.4.3.2 Approche locale

Les descripteurs locaux sont calculés non pas sur toute l'image, mais sur des zones spécifiques, par exemple dans le voisinage des points d'intérêts (tels ceux de Harris [Harris 90], SIFT [Lowe 99] et SURF [Bay 08]) ou encore sur des régions d'intérêt stables telles que les régions MSER [Matas 02].

Dans ces techniques locales, en particulier celles basées sur les points d'intérêt, le processus de reconnaissance consiste généralement en trois étapes [Grauman 11] :

1. Extraire les caractéristiques locales de l'image référence et de l'image test. Dans cette étape, on peut utiliser les détecteurs de caractéristiques images³.
2. Calculer les correspondances entre les points d'intérêt. Cette deuxième étape peut faire appel aux méthodes de calcul de similarités qui se basent essentiellement sur le calcul du plus proche voisin. Dans ce contexte, l' on peut citer des algorithmes comme kd-trees [Friedman 77], LSH⁴ [Indyk 98], k-means tree [Kanungo 02]...
3. Vérifier si les points d'intérêts appariés respectent une configuration géométrique cohérente. Cela veut dire que ces derniers sont liés par une transformation géométrique commune. Cette étape peut être effectuée en utilisant

2. Appelé PCA en anglais.

3. appelées Image features en anglais, renvoient dans ce cas au points d'intérêts et aux descripteurs locaux.

4. Locality-sensitive hashing

des méthodes comme l'homographie [Agarwal 05] pour le calcul de la transformation géométrique. Ensuite, afin d'éliminer les points aberrants, on peut utiliser RANSAC [Fischler 81] par exemple.

1.4.4 Pose

La pose de la caméra est définie par une matrice $3 \times 4 [R | t]$. R s'agit d'une matrice 3×3 qui représente la rotation de la caméra, t est un vecteur décrivant la translation de cette dernière. Le problème de calcul de pose de la caméra à partir de correspondances de n points est connu sous le nom de perspective- n -point (PnP). Pour trois points, la pose de la caméra n'est pas unique. Plusieurs travaux en vision par ordinateur se sont proposés à résoudre le problème (P3P) [Haralick 94]

[Quan 99]. Pour $n \geq 4$, la solution devient unique à condition que les points ne soient pas coplanaires ou alignés. Dans ce cas, l'on peut citer la méthode POSIT⁵ couramment utilisée en réalité augmentée :

POSIT

L'algorithme POSIT est inventé par Dementhon et al. [DeMenthon 95], il permet de calculer la position et l'orientation d'un objet dans une scène réelle. POSIT se base sur la projection orthographique réduite, appelée en anglais SOP, acronyme de « scaled orthographic projection ». POSIT estime la pose (position et orientation) en utilisant les couples représentant les correspondances de points, p_i et p'_i , comme le montre la figure 1.4.3.

A la convergence, SOP ressemble à la projection perspective. L'approximation SOP conduit à un système d'équations linéaires, ce qui donne directement la rotation et la translation, sans avoir besoin d'une pose initiale. Le SOP de p_i est le point schématisé dans la figure 1.4.3 par \hat{p}'_i , avec une certaine valeur d'échelle. Cette valeur d'échelle peut être trouvée en faisant coïncider \hat{p}'_i avec la projeté de p_i , suivant la projection perspective, noté par p'_i . POSIT affine itérativement cette valeur d'échelle.

La boucle de POSIT se déroule en quatre étapes :

1. la valeur d'échelle w_i est initialement fixé à 1 pour chaque correspondance.
2. Estimer les paramètres de la pose à partir de système d'équation linéaire, formé par les correspondances de points.
3. Calculer la nouvelle valeur d'échelle avec la formule suivante $w_i = \frac{p_i^T K}{t_x} + 1$

5. POSIT est implémenté par la bibliothèque de vision par ordinateur OpenCV.

4. Refaire la boucle à partir de l'étape 2 jusqu'à ce que le changement en devienne inférieur à un certain seuil ou jusqu'à ce que le nombre d'itération maximum w_i soit atteint.

Notons que, pour calculer la pose, POSIT n'a besoin de connaître que quatre points de l'espace 3D (non coplanaires), ainsi que leur projection en 2D.

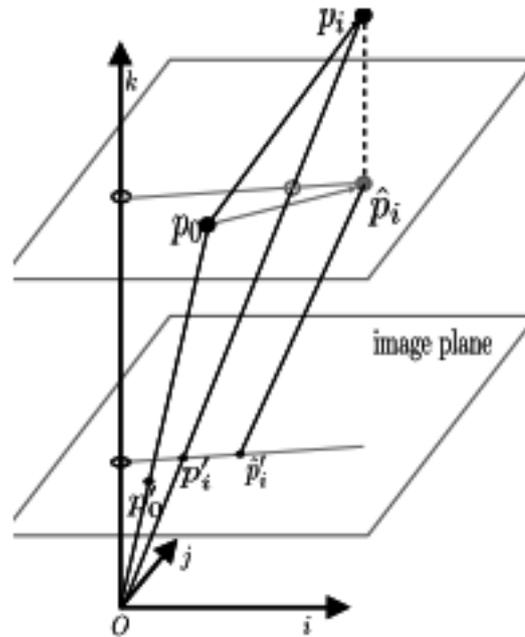


FIGURE 1.4.3 – Projection orthographique mise en œuvre par POSIT

1.4.5 Suivi

En réalité augmentée, le suivi « Tracking » est un processus fondamental. Le suivi permet de déterminer en continu la position et l'orientation de la scène réelle ou d'un objet dans l'espace physique, dans le but de faire coïncider visuellement le virtuel sur le physique. Les types de suivi sont classifiés selon la technologie employée ou la nature du capteur. Les principales technologies utilisées aujourd'hui pour le suivi sont :

1.4.5.1 La technologie Radio

Le suivi Radio est basé sur les ondes magnétiques. Une onde électromagnétique consiste en l'énergie générée par une oscillation, chargée de particules électriques, dans l'espace. La génération d'onde électromagnétique est connue sous le nom d'émission de fréquence Radio RF [Torres-Solis 10]. Les solutions radios permettent d'estimer l'emplacement d'une cible mobile dans l'environnement par estimation d'une ou plusieurs propriétés d'une onde électromagnétique envoyé par un émetteur et reçu par une station mobile. Ces propriétés dépendent généralement de la distance parcourue par le signal radio et les caractéristiques de l'environnement.

La Radio Frequency IDentification (RFID) est la plus connue des technologies utilisant la localisation par proximité [Ahriz Roula 10]. Celle-ci est basée sur l'échange des signaux radio entre un lecteur RFID et une radio-étiquette RFID. Le premier est constitué d'une antenne, d'un processeur et d'une alimentation. La radio-étiquette RFID est constituée d'une antenne et d'une mémoire; elle peut être active, c'est-à-dire posséder sa propre alimentation, ou passive, c'est-à-dire être alimentée par le signal radio reçu. Le lecteur RFID émet régulièrement des requêtes d'informations qui sont captées par des radio-étiquettes. Ce signal fournit aux radio-étiquettes passives l'énergie dont elles ont besoin pour répondre à cette requête. La portée de ce signal radio étant limitée, les étiquettes qui répondent sont localisées dans la zone à proximité du lecteur RFID. Selon la taille de la zone à couvrir et le nombre de personnes à localiser, le déploiement de cette technologie peut être effectué de deux manières :

- dans le premier cas, les lecteurs RFID constituent la partie fixe et les personnes à localiser sont dotées de radio-étiquettes ;
- dans le second cas, la structure à couvrir est équipée de radio-étiquettes et les personnes à localiser de lecteurs RFID.

Le coût des lecteurs RFID étant plus élevé que celui des étiquettes, le choix de l'architecture de déploiement dépend du coût envisagé pour l'infrastructure. Le réseau constitué de lecteurs/étiquettes RFID peut être remplacé par un réseau Bluetooth utilisé principalement pour la connexion sans fil entre des équipements par liaison radio de courte portée.

Les réseaux WiFi installés dans certains lieux publics peuvent être considérés comme une infrastructure pour la localisation. Un réseau WiFi est constitué d'un ensemble de points d'accès, chacun permettant de couvrir une zone allant jusqu'à quelques centaines de mètres. La taille de cette zone dépend de l'environnement

de propagation. Selon les spécifications de du WiFi (IEEE 802.11), chaque point d'accès diffuse régulièrement un signal sur un canal de contrôle qui permet au mobile de se connecter au réseau par le point le plus favorable. Ces canaux sont scrutés périodiquement par les mobiles afin de pouvoir utiliser un autre point d'accès si nécessaire. Parmi les informations diffusées sur les canaux de contrôle figure une identification du point d'accès. Les mobiles recevant cette information, peuvent alors se localiser à proximité (dans la zone de couverture) de ce point d'accès. Le temps de propagation du signal entre l'émetteur et le récepteur est une grandeur physique qui permet le calcul de la distance. Ceci peut être obtenu par rapport à des points d'accès à un réseau Wifi [Ahriz Roula 10], en estimant le temps de propagation à partir du timestamp contenu dans les signaux de contrôle.

1.4.5.2 La technologie Ultrason

Il s'agit de capteurs télémétriques utilisant la variation du « temps de vol » pour déterminer la distance séparant la source ultrasonore d'un élément de l'environnement. En utilisant ce principe de mesure du temps aller-retour par trois capteurs et respectivement capté par trois microphones, il est possible de déterminer les 6 degrés de liberté composant la position-orientation de l'objet à suivre. Toutefois, même si ces technologies sont peu onéreuses et que leur précision est de l'ordre du centimètre, elles peuvent se heurter à des erreurs de mesure principalement dues au phénomène d'absorption ou de réflexions multiples (échos) des matériaux environnants. On notera également une fréquence d'échantillonnage faible ne permettant pas une utilisation dans des applications nécessitant une haute réactivité.

1.4.5.3 La technologie inertielle

La centrale inertielle est un dispositif de mesures composé généralement d'un accéléromètre trois axes, d'un gyroscope trois axes (mesurant les taux de rotation) et d'un magnétomètre trois axes (mesurant le champ magnétique terrestre). L'accéléromètre retourne le vecteur de gravité (le vecteur pointant vers le centre de la terre) quand il est en état de repos. En cas de mouvement, l'accéléromètre donne l'accélération actuelle par rapport aux trois axes, en m/s^2 . Le magnétomètre 3 axes permet d'effectuer les mesures du champ magnétique. Les informations provenant de l'accéléromètre et du magnétomètre suffisent pour calculer l'orientation. Toutefois, les sorties de ces deux capteurs sont imprécises. Le gyroscope, est plus précis que les deux premiers capteurs et a un temps de réponse très court. L'inconvénient de son utilisation pour le suivi se manifeste dans l'accumulation d'erreurs d'intégration. Effectivement, le gyroscope fournit les vitesses de rotation angulaire pour chacun des trois axes. Pour déduire l'orientation relative au trois vecteurs

de vitesse, les vecteurs de vitesse doivent être intégrés au cours du temps. L'intégration incrémentale donne lieu à des erreurs qui se cumulent en générant l'erreur connue sous le nom de la dérive « drift » du gyroscope. Pour pallier aux imprécisions relatives aux capteurs de la centrale inertielle, plusieurs solutions [Haid 04] [Bachmann 99][Gustafsson 02] consistent à fusionner les sorties de ces capteurs inertiels, en leur appliquant des filtres comme le filtre complémentaire, le filtre de kalman et le filtre particule.

1.4.5.4 La technologie optique

Ce type de suivi est également appelé suivi visuel ou suivi basé vision. Il peut utiliser une caméra (suivi monoculaire), ou deux caméras (suivi stéréoscopique). Au niveau de la technologie optique, on peut distinguer deux grandes catégories : le suivi avec marqueur et celui sans marqueur. Le suivi avec marqueur consiste à incruster des marqueurs fiduciaires dans la scène réelle. La deuxième catégorie se base sur les caractéristiques images comme par exemple les points d'intérêt et les segments de lignes. Les deux catégories sont présentées en détail ci-dessous.

Les marqueurs fiduciaires

Plusieurs types de marqueurs sont aujourd'hui utilisés en réalité augmentée. Les plus connus sont les marqueurs templates et les marqueurs à codes-barres 2D [Siltanen 12].

Les marqueurs templates sont en noir et blanc et consistent en un simple motif entouré d'une bordure noire. Ce genre de marqueur est utilisé par ARToolKit⁶. La figure suivante montre des exemples de marqueurs templates.

Les marqueurs de type codes-barres 2D que l'on emploie aujourd'hui souvent en réalité augmentée, sont les QR codes. Le terme QR code est acronyme de « Quadratic Residue Code » [Interlando 11]. Les QR codes ont été inventé en 1994 par Denso, l'une des majeures compagnies du groupe Toyota. Ils ont été approuvé comme un standard ISO/ IEC18004 en Juin 2000. Un QR code est en réalité un symbole matriciel, qui contient généralement les éléments suivants : le motif de recherche, le motif d'alignement, le motif de synchronisation et la zone calme. La figure suivante montre la structure de ce marqueur :

6. <http://www.hitl.washington.edu/artoolkit/>

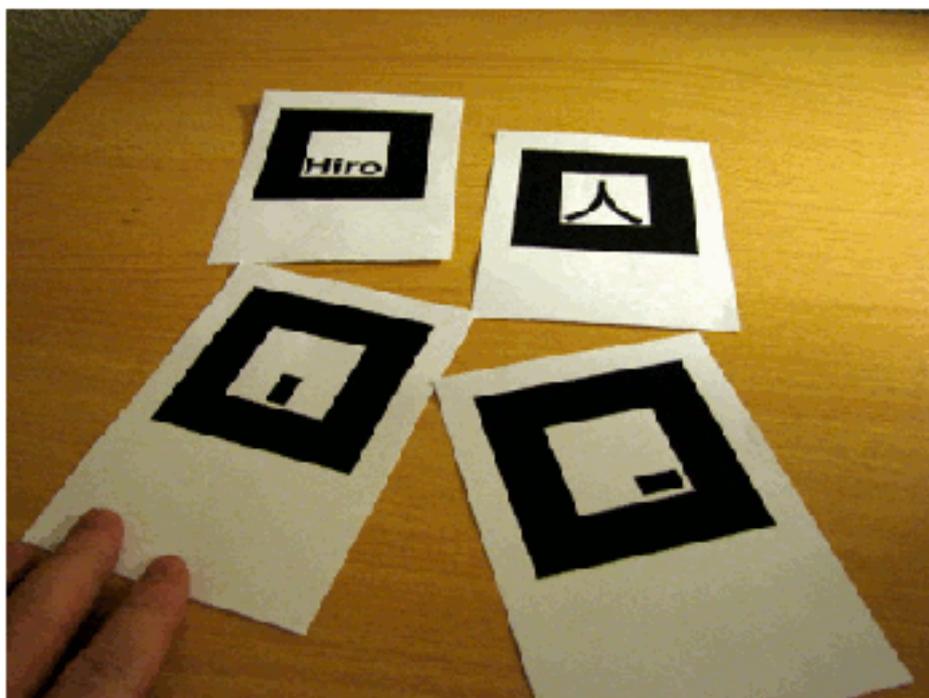


FIGURE 1.4.4 – marqueurs ARToolKit de type « template »

Les données sont enregistrées en binaire dans la zone donnée du QR Code, elles sont représentées en gris dans la figure 1.4.5. Au moment de la création d'un QR code, les nombres binaires de '0' et '1' sont converties respectivement en des cellules noires et blanches.

Le suivi sans marqueur

Dans ce qui suit, nous présentons quelques exemples d'algorithmes de suivi, basés vision qui sont souvent utilisés en RA :

RAPiD

RAPiD (Real-time Attitude and Position Determination) est un algorithme de suivi visuel monoculaire. Il a été inventé par Harris et al. [Harris 90] en 1990. RAPiD se base sur le modèle 3D pré-calculé de l'objet ciblé dans la scène réelle. L'idée clé sur laquelle est fondée cet algorithme, consiste à sélectionner des points de contrôle sur les arêtes de l'objet. Ensuite, la position et l'orientation de l'objet est déduite par comparaison de ces points dans le modèle et l'image de la scène réelle. Etant donné que RAPiD opère en temps réel, il a été adopté par plusieurs auteurs de travaux de RA, en l'occurrence par Klein et al. [Klein 03]. Le processus

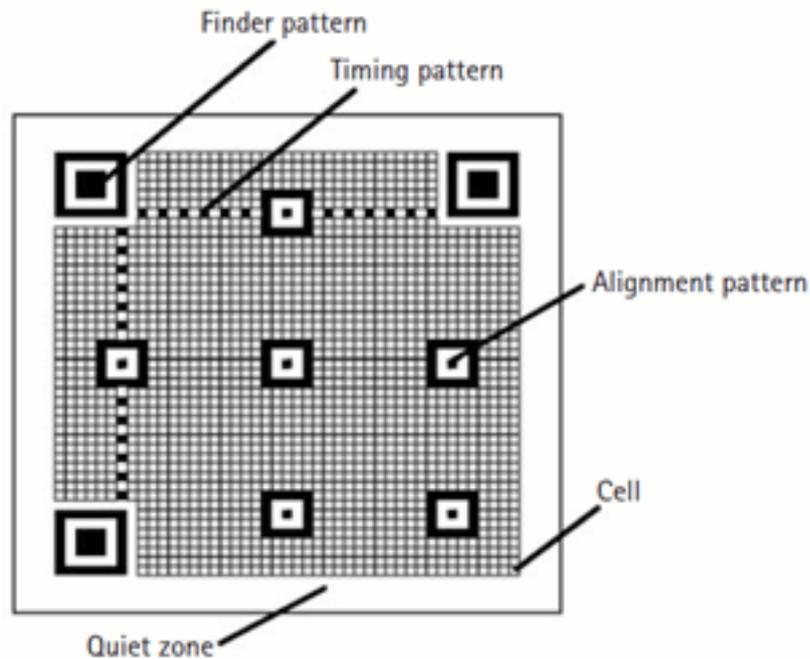


FIGURE 1.4.5 – Structure d'un QR Code

de suivi mis en œuvre par RAPID est illustré par la figure 1.4.6. Il se déroule en quatre étapes :

- Etape 1 : l'image vidéo est acquise depuis la caméra.
- Etape 2 : projection du modèle basée sur l'estimation présente de la pose de la caméra.
- Etape 3 : calcul des vecteurs d'erreurs ponctuelles (distances) entre le modèle projeté et son image dans la vidéo de la scène réelle.
- Etape 4 : la pose est mise à jour en se basant sur le vecteur de mouvement afin de minimiser les erreurs mesurées à l'étape précédente.

PTAM

PTAM est l'acronyme de « Parallel Tracking and Mapping », il s'agit d'un algorithme de suivi monoculaire, proposé par Georg Klein et al. [Bimber 05] en 2007, permettant de faire le suivi dans des environnements non préparés « unknown environments ». Cet algorithme divise le fameux algorithme SLAM « Simultaneous Localization and Mapping » en deux threads séparés ; un thread traitant le suivi, l'autre traitant la cartographie « mapping » représentant l'environnement.

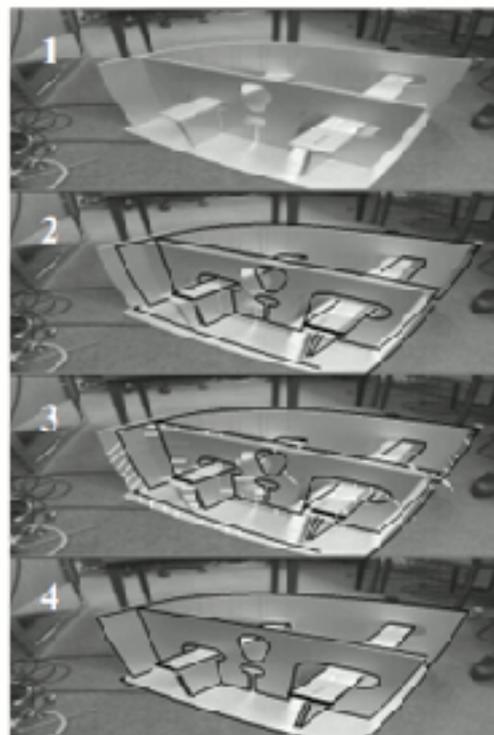


FIGURE 1.4.6 – Boucle de système de Suivi de RAPID [Klein 03]



Nous proposons au lecteur souhaitant avoir un complément d'information sur SLAM de consulter le travail de Leonard et al. [Leonard 91], les premiers à avoir proposé cet algorithme.

Le deuxième thread permet de construire la carte 3D de points caractéristiques « features points » observés depuis les frames vidéo. Le point j de la carte (p_{jw}) a les coordonnées $p_{jw} = (x_{jw}, y_{jw}, z_{jw})^T$ dans le repère de la frame monde w . Le processus de cartographie permet également de stocker des images clés. Il s'agit d'images prises par la caméra à différents instants. Notons que la construction de la carte de points caractéristiques se déroule en deux étapes distinctes :

Premièrement, la carte de l'environnement est initialisée en utilisant la stéréovision. A ce stade, la première image-clé est stockée par le processus de cartographie, et 1000 2D patch « patch-tracks » sont initialisées avec les coins maximaux FAST. Les patches 2D sont suivis à travers le mouvement de la caméra. Ensuite, la deuxième image-clé acquise par la caméra, est utilisée afin de calculer la correspondance de points caractéristiques avec l'image clé précédente. Ceci permet à l'algorithme stéréo cinq-points stéréo [Stewenius 06] et à l'algorithme RANSAC [Fischler 81] d'estimer la matrice essentielle et de trianguler les points caractéristiques. La carte 3D obtenue suite à la triangulation, est par la suite affinée grâce au « bundle adjustment » [Triggs 00].

Deuxièmement, la carte est sans cesse affinée et élargie à fur et à mesure que de nouvelles images clés sont ajoutés par le thread de suivi.

Le processus de suivi est responsable de tracer le mouvement de la caméra. A chaque nouvelle image acquise par la caméra, ce processus exécute les étapes suivantes :

- Une pose au préalable « prior pose » E'_{CW} est estimée en se basant sur le modèle de mouvement « motion model » M , de la manière suivante :

$$E'_{CW} = ME_{CW} = \exp(\mu)ME_{CW} \quad (1.4.1)$$

Le vecteur de mouvement μ est calculé par l'équation suivante :

$$\mu' = \underset{\mu}{\operatorname{argmin}} \sum_{j \in s} \operatorname{Obj} \left(\frac{|e_j|}{\sigma}, \sigma_T \right) \quad (1.4.2)$$

e_j est le vecteur d'erreur calculé de la façon suivante :

$$e_j = \begin{pmatrix} \hat{u}_j \\ \hat{v}_j \end{pmatrix} \text{camProj} (\exp(\mu)ME_{CW}) \quad (1.4.3)$$

$Obj(\cdot, \sigma_T)$ est la fonction objective de bi-poids « biweights » de Tukey [Huber 81].

$(\hat{u} \hat{v})^T$: représente la position du patch qui en se référant à l'unité de pixel.

- Les points de la carte, construite par le thread de cartographie « mapping », sont projetés sur l'image suivant la pose au préalable, calculée à l'étape précédente.
- Un nombre réduit (50) de points caractéristiques de la carte, est cherché dans l'image.
- La pose de la caméra est mise à jour partir des correspondances de point calculées à l'étape précédente, en utilisant l'équation 1.4.2.
- Un nombre plus important de points caractéristiques (1000) est projeté sur l'image, afin de chercher leurs correspondances.
- Finalement, la pose finale est calculé à partir des tous les correspondances.

1.4.6 Recalage

Le recalage est un problème courant en réalité augmentée. Le recalage ou « registration » décrit le procédé de superposer et d'unifier les objets virtuels sur le réel, dans le but de les fusionner. Selon Anastassova et al. [Anastassova 07], le recalage vise précisément à faire coïncider visuellement, haptiquement, auditivement les objets du monde réel aux enrichissements qui leur sont associés. L'objectif est de réduire au maximum le décalage perçu par l'utilisateur. Ce procédé de recalage est cité par Azuma [Azuma 01] comme une propriété qu'un système de RA doit satisfaire. La technologie n'aboutit pas encore à un recalage parfaitement fiable. En effet, le recalage dépend directement de la qualité du suivi.

1.4.7 Rendu

Suivant le type de modalité, le rendu peut être graphique, haptique, audio, etc. Le rendu graphique 3D s'inscrit dans le cadre de synthèse d'image tridimensionnelle. Dans ce qui suit, nous allons nous intéresser au rendu graphique. En réalité, le terme de rendu se réfère à des calculs effectués par le moteur de rendu pour traduire la scène géométrique 3D en une image 2D. Au cours de ce processus, la texture, l'illumination et les ombres sont combinés ensemble, afin de déterminer la valeur de couleur de chaque pixel dans l'image 2D. A l'heure actuelle, les moteurs de rendu les plus connus sont Mental Ray et V Ray, utilisés respectivement par

Autodesk Maya et 3DS Max. Comme le soulignent Bimber et al. [Bimber 05], il existe plusieurs techniques utilisées pour le rendu à savoir :

Rastérisation (*rasterization*) : la rastérisation intervient dans le processus de génération d'une image sur les cartes graphiques, permettant de projeter un maillage exprimé en trois dimensions sur un écran à deux dimensions. La rastérisation correspond, d'une façon plus précise, à la discrétisation de données vectorielles (maillages) en rasters (matrices de pixels).

Lancer de Rayon (*Ray Tracing*) : le lancer de rayon consiste en une technique d'illumination de la scène numérique. En fait, pour chaque pixel de la scène, un (ou plusieurs) rayon (s) de lumière sont tracés à partir de la caméra jusqu'à l'objet 3D le plus proche. La couleur de chaque pixel est calculée en se basant sur l'interaction des rayons de la lumière avec les objets, en termes de réfraction et de réflexion. Cette technique, relativement lente, est utilisée en rendu photo-réaliste.

Radiosité (*Radiosity*) : la radiosité ou la radiance représente une technique d'éclairage (ou illumination) d'une scène 3D. Contrairement à la précédente, cette technique est indépendante de la caméra. La radiosité consiste à simuler la couleur de la surface en tenant compte de l'éclairage indirect. Pour créer un rendu photo-réaliste, l'on utilise le lancer de rayon, la radiosité. Cependant, comme le font remarquer Chen et al. [Chen 12], c'est le rendu non photo-réaliste qui est souvent utilisé afin de maintenir la cohérence temporelle.

En plus de la contrainte temps-réel, le rendu en RA doit respecter la règle de cohérence spatiale. En effet, la superposition des objets virtuels sur la scène réelle provoque souvent une occultation contradictoire, qui peut être, un objet réel caché par objet virtuel qui doit être normalement en arrière-plan. Dans de tels cas, l'utilisateur constate que l'objet virtuel n'appartient pas à la scène réelle, ce qui diminue le sentiment de présence et perturbe l'expérience de réalité augmentée. Pour parvenir à maintenir la cohérence spatiale, plusieurs techniques [Kanbara 02] [Kim 03] ont utilisés qui consistent en général à reconstituer la profondeur de la scène à fin de palier au problème d'occultation contradictoire.

1.4.8 Modèles d'affichage

Pour combiner le réel et le virtuel, deux modèles d'affichage sont utilisés, l'affichage direct (optical see-through) et l'affichage indirect (video see-through) [Azuma 97]. Le modèle "optical see-through" permet à l'utilisateur de visualiser directement la scène à travers le dispositif HMD « Head-Mounted Displays ». Cependant, le



FIGURE 1.4.7 – Occultation contradictoire dans l’image gauche [Fukiage 12]

modèle “video see-through” consiste à acquérir le point de vue de l’utilisateur, puis la video du monde réel est combiné avec le rendu (augmentations graphiques) orienté suivant le point de vue de l’utilisateur.

Conclusion

Au fil de la première partie de ce chapitre, nous avons pu nous rendre compte que les concepts de la réalité augmentée ne sont pas nouveaux. C’est finalement l’évolution technologique qui a permis un regain d’intérêt dans ce domaine. Nous avons pu discuter de la richesse conceptuelle de ces idées, et il ressort de ce fait une grande diversité d’interprétations sur la définition même de la réalité augmentée. En ce sens, pour avoir une vision commune au fil du document avec le lecteur, nous avons proposé notre propre définition de la réalité augmentée :



La réalité augmentée est la combinaison de données sensorielles avec le monde réel de telle manière que, poursuivant un but bien déterminé, elles semblent co-exister avec le monde réel et lui être liées.

Comme nous l’avons déjà précisé, cette définition, bien que globale, n’exclut pas les précédentes et profite ainsi de leurs points forts.

Au niveau de la deuxième partie de ce chapitre, nous avons exposé les moyens technologiques permettant de mettre en œuvre la RA. Nous avons donc voulu présenter les fondements de ce qui caractérise l’implémentation technologique de

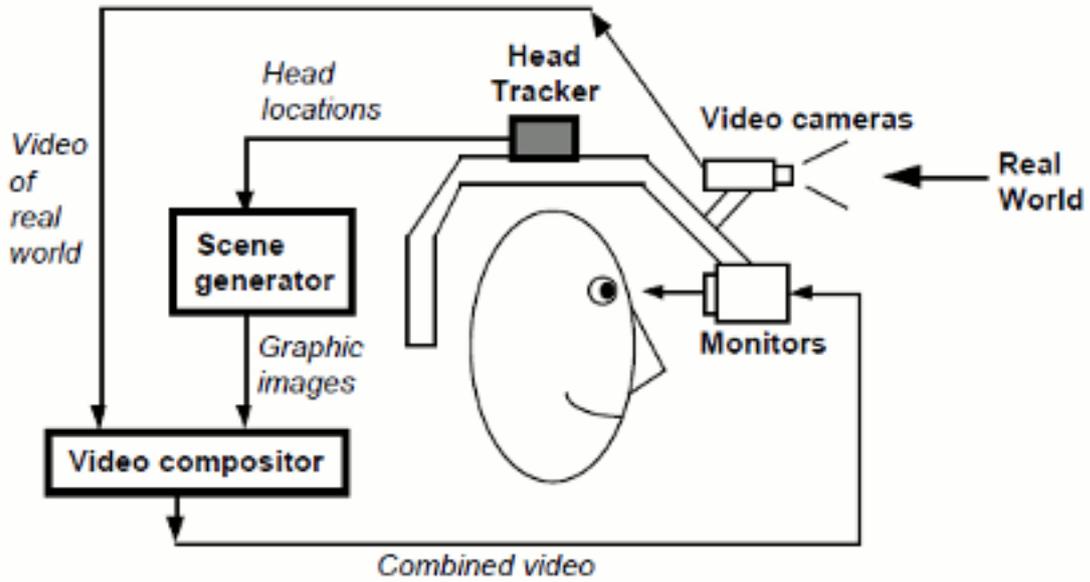


FIGURE 1.4.8 – video see-through HMD [Azuma 97]

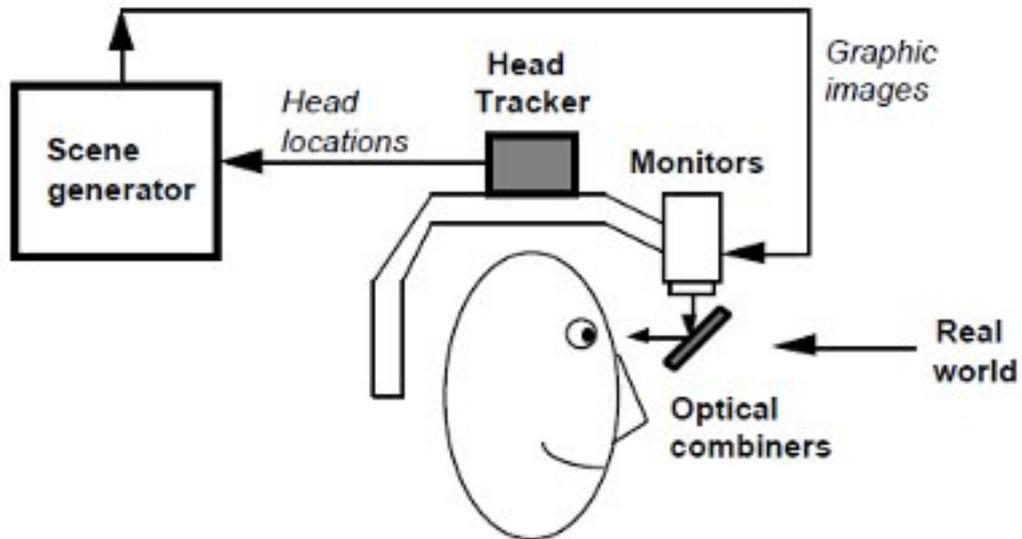


FIGURE 1.4.9 – optical see-through HMD [Azuma 97]

cette dernière. Nous nous sommes principalement intéressés à la modalité visuelle de ce vaste domaine, la vue étant généralement privilégiée par l'humain aux autres sens.

Partie 1

« Much of what sophisticates
loftily refer to as the
"complexity" of the real world
is in fact the inconsistency in
their own minds... »"

(Thomas Sowell)

Chapitre 2

Le visiteur, la visite guidée au musée et la réalité augmentée

Introduction

Si le visiteur est bien au centre des préoccupations du musée, quelles en sont ses attentes ? Quelles sont les expériences recherchées par le visiteur ? Comment le musée répond-il aux besoins de ses publics ?

Au niveau de la deuxième partie de ce chapitre, nous allons donc étudier l'intérêt d'introduire de nouvelles technologies comme la réalité augmentée dans l'assistance à la visite. A travers l'examen de l'état de l'art, nous allons alors tenter de répondre à cette question : Est ce que les propositions actuelles permettent de soutenir efficacement l'activité de visite du musée ?

2.1 Le visiteur

2.1.1 Les motivations de visite

La question que l'on se pose à ce niveau est pourquoi visiter un musée ? Les enjeux de cette question semblent être essentiels, elle définit une notion d'attente du visiteur envers le musée.

Dans cette optique, plusieurs travaux de recherche ont tenté de recenser les raisons pour lesquels les visiteurs viennent au musée. Pour Packer et al. [Packer 02], les visiteurs vont au musée pour cinq raisons différentes : l'apprentissage et la découverte « Learning and Discovery », l'amusement passif « Passive Enjoyment »,

la relaxation « Restoration », l'interaction sociale « Social Interaction » et l'accomplissement personnel « Self-fulfilment ». Selon Mottaz-Baran [Mottaz Baran 01], les différentes raisons d'aller au musée sont : l'intérêt pour le thème traité, le désir de se cultiver, le plaisir et l'émotion esthétique et la recherche de référentiel. D'un autre côté, les travaux de Falk et al. [Falk 98] ont permis d'identifier trois raisons majeures motivant la visite de musées : l'apprentissage, l'amusement et l'établissement de liens sociaux.

En se basant sur les travaux cités ci-dessus, l'on peut constater qu'il existe plusieurs raisons pour la visite des musées. Néanmoins, il semble évident que les deux raisons principales, également opposées, soient : apprendre et s'amuser. De plus, Packer et al. [Packer 02] montrent à l'aide d'une étude réalisée auprès des visiteurs des musées en Australie, que le fait d'apprendre et découvrir constitue le premier motif de visite, suivi par le désir de s'amuser et de s'épanouir. Bien évidemment, ce résultat met en relief l'importance que présentent l'apprentissage et la découverte pour le visiteur de musée.

Par ailleurs, l'apprentissage fut considéré par les recherches en muséologie comme une expérience recherchée par le visiteur [Falk 00][Packer 06][Kotler 99] : « learning experience ». L'expérience d'apprentissage se situe dans un contexte plus large qu'est l'expérience visiteur. En ce sens, avant de nous consacrer dans le cadre de cette thèse à l'expérience d'apprentissage, nous allons nous intéresser dans la section suivante à l'expérience visiteur.

2.1.2 L'expérience visiteur

L'expérience visiteur « visitor experience » ou encore l'expérience de musée « museum experience » comme appelée par Falk et al. [Falk 92], constitue depuis plusieurs années un sujet de recherche à la fois pertinent et complexe. Comme le souligne Daengbuppha [Daengbuppha 09], la complexité de cette dernière revient à l'aspect expérientiel et à la multiplicité de valeurs de consommation recherchées par les visiteurs. "... The complexity of visitor experiences in terms of how 'individual visitors' construct the meaning of experience explained by experiential aspect as well as the multiplicity of visitor experiences in terms of experience consumption practice..." ([Daengbuppha 09], page 224).

Les sociologues à l'instar de Cohen [Cohen 79], furent les premiers à introduire le concept de l'expérience en tourisme. En effet, en se basant sur une étude empirique, Cohen [Cohen 79] dresse le modèle typologique de touristes dans lequel il distingue l'expérientiel comme un type de tourisme. En effet, dans sa typologie,

Cohen [Cohen 79] souligne que la quête d'expérience constitue en elle-même un style de consommation définissant ainsi un type de tourisme à part entière. Un peu plus tard, l'expérience devient reconnue comme un aboutissement ou un produit de la visite au musée [McLean 97].

Les dimensions de l'expérience visiteur

Comme le souligne Mason [Mason 12], l'expérience visiteur décrit l'expérience globale de l'individu au musée. Elle peut encore désigner le sentiment de satisfaction que le visiteur éprouve vis-à-vis de sa visite au musée. Mason [Mason 12] a défini principalement trois éléments influençant l'expérience visiteur, qu'il appelle dimensions de contexte "dimensions of the context". Selon lui, les dimensions de contexte de l'expérience visiteur sont le visiteur, le contexte du musée et le contenu du musée.

Visiteur : les attentes, les motivations et les comportements des visiteurs sont des éléments essentiels qui interviennent dans l'étude de l'expérience visiteur ; ce sont des éléments à prendre en compte dans la conception des outils de médiation pour la visite guidée. D'un autre côté, Mason [Mason 12] souligne l'importance de la participation active des visiteurs dans la création d'une visite de qualité, par l'intermédiaire de l'interaction avec les objets exposés et aussi par l'interaction sociale avec les autres visiteurs.

Contexte du musée : le contexte du musée intègre principalement l'environnement physique du musée ainsi que ses objectifs et missions. L'environnement physique de musée inclut entre autres les conditions d'illumination, l'organisation et la présentation des expositions. La structure muséale doit veiller à ce que le visiteur soit perpétuellement orienté à l'intérieur du musée en utilisant des plans, des indications visuelles, etc. Il est aussi évident que bien accueillir le visiteur, bien interpréter et labéliser les œuvres d'art est au cœur des missions du musée.

Contenu du musée : le contenu de musée renvoie principalement aux objets exposés et à la médiation offerte au visiteur. Aujourd'hui, avec l'essor de la technologie mobile, les applications mobiles s'imposent dans la création des visites guidées. Il est à noter que l'expérience visiteur est à ne pas confondre avec l'expérience utilisateur introduite par Donald Norman¹. En effet, comme le fait remarquer Mason [Mason 12], l'expérience visiteur a un spectre plus large que l'expérience utilisateur, qui renvoie à l'étude subjective de l'utilisabilité d'un produit

1. Donald Norman utilise le terme expérience utilisateur pour la première fois dans son livre *The Design of Everyday Things*.

ou d'un système informatique. En ce sens, l'expérience utilisateur d'une application informatique dédiée à la médiation fait partie de l'expérience visiteur.

Les types d'expériences visiteur

L'expérience visiteur est fortement liée à la valeur recherchée par ce dernier. En effet, en se basant sur la théorie de valeurs de consommation « theory of consumption values » [Sheth 91] et donc sur une perspective de marketing, Bourgeon-Renault et al. [Bourgeon-Renault 07] ont établi une classification des expériences visiteurs. Les valeurs attachées par les publics (considérés dans ce cas comme des clients) aux musées renvoient aux valeurs de consommations suivantes :

une valeur fonctionnelle ou capacité des musées à proposer une offre de qualité caractérisée par les thèmes abordés, les œuvres exposées, les lieux et les services mis à la disposition des publics.

une valeur sociale liée à la visite pratiquée avec un ou plusieurs groupes sociaux (famille, amis...).

une valeur émotionnelle ou capacité des musées à susciter des émotions, la sensibilité au « beau » ou à l'esthétisme, un sentiment de découverte, d'évasion, de surprise, d'étonnement, de dépassement de soi et d'émerveillement, le développement de l'imaginaire, la nostalgie, une impression de liberté, le plaisir d'être dans les lieux, le désir de nouveauté et de variété, la perception mystique du lieu, la spiritualité...

une valeur épistémique mettant l'accent sur la capacité des musées à susciter la curiosité et à stimuler le désir de la connaissance.

une valeur conditionnelle s'analysant comme le résultat d'une situation spécifique ou d'un ensemble de circonstances (présence de monde ou non, entrée payante ou gratuite, présence d'un guide, occasion particulière de visite...).

Dans le but d'identifier les types d'expériences satisfaisantes pour les visiteurs "satisfying experiences", Pekarik et al. [Pekarik 99] ont mené une étude empirique qui a permis de dresser cette liste d'expériences :

L'expérience de l'objet : se focaliser sur quelque chose à l'extérieur de soi, voir des objets réels et concrets. Voir des œuvres d'arts de valeur qui incarnent la beauté.

L'expérience cognitive : se focaliser sur les aspects intellectuels de la visite, acquérir des nouveaux savoirs, élargir le spectre de ses connaissances, mieux appréhender des choses déjà connues.

L'expérience introspective : se focaliser sur sa personne, sur les sentiments suscités par la visite. Se laisser prendre par son imagination, ses souvenirs...

L'expérience sociale : se concentrer sur l'interaction sociale avec les gens que

l'on accompagne, les autres visiteurs ou avec l'équipe de musée.

D'un autre côté, en partant d'une perspective de marketing, Kotler [Kotler 99] distingue quatre catégories d'expériences allant de l'émotionnel au cognitif, qu'il appelle 'range of recreational experiences' : l'excitation, le divertissement, la contemplation et l'apprentissage. Les types d'expériences recherchés au cœur de la visite au musée sont donc pluriels. Nous pouvons constater qu'essentiellement les catégories : sociale, émotionnelle et éducative sont omniprésentes mais avec des appellations différentes. La catégorie émotionnelle est décrite dans certains travaux comme Introspective ou excitative. L'expérience d'apprentissage « learning experience »² apparaît dans certains travaux avec les noms éducative [Packer 06], épistémique ou cognitive.

Comme nous avons pu le voir dans la section précédente, venir au musée pour découvrir et acquérir des nouvelles connaissances et savoirs constitue une principale motivation de visite. Cependant, la question que l'on se pose par la suite est : est-ce que l'on vient au musée pour apprendre ou pour expérimenter l'apprentissage « to experience learning » ? En confrontant ce constat aux travaux de Packer [Packer 06], l'on peut simplement déduire que c'est essentiellement pour le deuxième objectif. "...learning for fun to refer to the phenomenon in which visitors engage in a learning experience because they value and enjoy the process of learning itself" ([Packer 06], page 299). "Although most visitors don't come with a deliberate intention to learn, they do seek, or are unconsciously drawn into, an experience that incorporates learning" ([Packer 06], page 340). Ainsi, c'est de cette perspective expérientielle que nous allons étudier l'apprentissage au musée.

2.2 Apprendre au musée

La question que l'on peut légitimement se poser est : « Y-a-t-il réellement un apprentissage au musée ? ».

En effet, comme nous avons pu le voir dans la section précédente, apprendre et découvrir constituent une principale motivation de visite. En plus, plusieurs travaux dans le domaine muséal, considèrent maintenant le rôle d'un musée comme un rôle multi faces incluant plusieurs tâches comme la conservation, la documentation, la présentation et l'explication des artefacts, etc. En particulier, beaucoup voient aujourd'hui les musées comme une importante source d'informations et un lieu pour permettre aux gens d'apprendre à travers les expositions [Falk 05] [Falk 00]

2. Dans la suite de cette thèse, c'est l'appellation « expérience d'apprentissage » que nous allons adopter

[Roberts 97]. De plus, le musée fut considéré comme un endroit éminemment propice à l'apprentissage par la découverte.

Par ailleurs, l'ICOM³ donne une définition du musée qui met en avant le rôle éducatif et de transmission de savoirs qu'assume le musée : « Un musée est une institution permanente, sans but lucratif, au service de la société et de son développement, ouverte au public et qui fait des recherches concernant les témoins matériels de l'homme et de son environnement, acquiert ceux-là, les conserve, les communique et notamment les expose à des fins d'études, d'éducation et de délectation »⁴. On voit donc ici que la fonction du musée est de communiquer un savoir. Si apprendre n'est pas toujours l'objectif principal du visiteur, comme on a pu le voir précédemment, la vocation première du musée, est-elle, bien de communiquer et de permettre au visiteur d'apprendre.

2.2.1 L'apprentissage

Du point de vue de la phénoménologie⁵, apprendre est à la fois comprendre un phénomène particulier qu'on peut qualifier d'objet d'apprentissage (quoi) [Beaty 97]. Apprendre est aussi réaliser un acte d'apprentissage (comment). En ce sens, l'apprentissage peut être étudié en tant que phénomène expérimenté ayant pour objectif de constituer le « quoi ». Le sens qu'incarne l'apprentissage pour l'individu est appelé conception de l'apprentissage.

A ce niveau du mémoire, nous allons donc tenter de répondre à la question suivante : « Que signifie apprendre ? ».

Les théories épistémologiques permettent de mettre en place plusieurs conceptions⁶ de l'apprentissage que Marton et al. [Marton 93] classent en deux catégories :

- Une conception reproductive de l'apprentissage : augmenter ses connaissances, mémoriser et reproduire, appliquer.
- Une conception de recherche de sens : comprendre, voir quelque chose de manière différente, changer sa vision du monde.

Marton et al. [Marton 93] mettent donc en évidence une corrélation entre approche et conception : l'approche de surface est associée à la conception reproductive (l'ap-

3. Organisation internationale des musées et des professionnels de musée

4. <http://icom.museum/la-vision/definition-du-musee/L/2/>

5. la phénoménologie, initiée par Husserl (1992)

6. Conception de l'apprentissage : il s'agit de la vision de l'apprentissage, du modèle ou de la manière privilégiée de considérer l'apprentissage.

prenant applique, exploite le savoir qu'il possède) et l'approche en profondeur à la conception de recherche de sens.

Loin du contexte formel de l'apprentissage dans lequel se trouve l'apprenant, le visiteur de musée quand à lui adopte une approche de profondeur. En effet, comme le souligne Roberts [Roberts 97], donner un sens à ce qu'il voit constitue la principale préoccupation de visiteur de musée en son expérience des collections : "En définissant la connaissance par rapport à un contexte social et historique plus large, on donne aux visiteurs la possibilité de trouver du sens à ce qu'ils savent et disent. Ils partagent maintenant avec le musée la responsabilité ainsi que le contrôle de ce que signifie leur expérience des collections. On voit donc s'éroder plusieurs principes anciens comme l'autorité du conservateur, le caractère sacré des objets, et même le prestige de l'institution elle-même en tant que source et dispensatrice de connaissances" ([Roberts 97], page 132).

Bien évidemment, cette vision profonde de l'apprentissage revient au caractère libre qui caractérise l'apprentissage dans un cadre informel comme celui de musée. En ce sens, l'apprentissage dans les musées comme vue par Falk et al. [Falk 00], est personnel et est orienté par les connaissances et les ressentis. De plus, ils qualifient cet apprentissage d'apprentissage libre « free-choice learning ».

D'un autre côté, les études réalisées par Falk et al. [Falk 98][Falk 05] démontrent également que les détails rappelés par les individus après la visite, sont généralement en lien étroit avec leurs connaissances et intérêts préalables. Ils rappellent à cet effet qu'un sujet apprend mieux les informations qui l'intéressent.

Si nous essayons de résumer les caractéristiques de l'apprentissage dans les musées sur lesquels s'accordent la plupart des chercheurs ici cités, nous pouvons retenir que l'apprentissage :

- a lieu au contact d'objets réels,
- s'appuie sur une approche de profondeur,
- y est volontaire,
- est stimulé par l'intérêt personnel.

L'enjeu par la suite est de savoir comment l'on apprend et comment faire apprendre au visiteur ? Quels types de connaissances transmettre aux visiteurs ? On s'intéressera donc aux constats sur l'apprentissage dans les musées, puis aux méthodes proposées pour décrire et mesurer cet apprentissage.

2.2.2 Les théories cognitivistes d'apprentissage

De point de vue des sciences cognitives, l'apprentissage est un processus non observable de réorganisation des structures cognitives. Ce processus est finalisé, visant à acquérir de nouvelles connaissances, d'habiletés, d'attitudes ou à modifier des acquis antérieurs, ce qui se traduit par un changement durable des comportements dont l'objectif est de s'adapter à soi-même et à son environnement.

Les théories d'apprentissage tentent de définir la manière dont toutes les connaissances doivent être acquises. Les sciences cognitives comportent de nombreux courants de pensée lorsqu'il s'agit d'aborder la question des mécanismes d'acquisition et de transformation des connaissances. Si l'on considère les principaux, alors nous retiendrons les cinq suivants : le béhaviorisme, le constructivisme, le cognitivisme, le connexionisme et l'énaction. Mais ces courants ne sont pas traités avec les mêmes égards suivant les domaines où ils s'appliquent. En effet, dans le cadre muséologique, c'est le modèle constructiviste qui est plutôt retenu [Hein 98].

La théorie du constructivisme [Piaget 69], prend résolument des principes en opposition avec le béhaviorisme en faisant intervenir le côté personnel de l'apprenant dans le processus d'apprentissage. L'individu représente les informations reçues en regard de ses propres expériences, de ce qu'il s'agit déjà. L'apprenant assimile les nouvelles connaissances en les modifiant et en l'adaptant à une situation nouvelle. L'apprenant ne transfère donc pas simplement le savoir dans sa mémoire, mais construit ses propres interprétations à partir des interactions qu'il a avec le monde. La pédagogie constructiviste aurait donc une tendance à transmettre le savoir en mettant l'apprenant au défi, en stimulant et en remettant en cause d'anciennes connaissances pour en produire de nouvelles. C'est ce double mécanisme d'assimilation-accomodation, que Piaget [Piaget 69] nomme équilibration.

Les travaux de Hein [Hein 98] ont cherché à trouver les liens entre les théories d'apprentissage cités plus haut et l'apprentissage dans le cadre de musée. En effet, après une analyse extensive des différents composants des théories d'apprentissage, Hein [Hein 98] suggère un modèle pour l'apprentissage au musée fondé sur le constructivisme. En effet, le modèle de Hein [Hein 98], met l'accent sur les visiteurs : les individus s'engageant dans un processus d'apprentissage libre et participent activement à construire un nouveau sens en se basant sur leurs nouvelles⁷ et antérieurs connaissances.

7. Les nouvelles connaissances réfèrent ici aux connaissances acquises lors de la visite

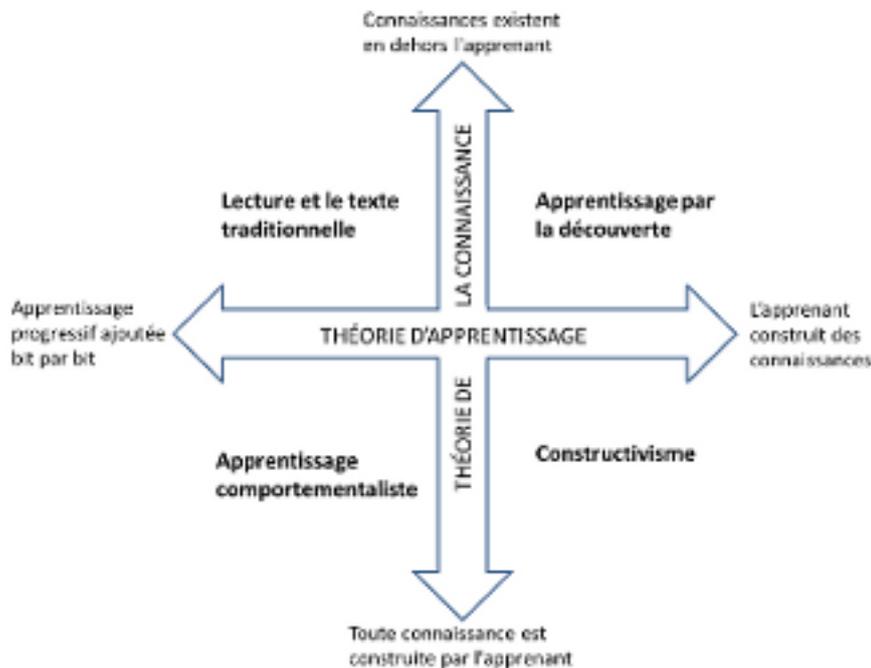


FIGURE 2.2.1 – Le modèle constructiviste d'apprentissage au musée de Hein

2.2.3 La connaissance

Qu'il s'agisse du domaine purement cognitif, du domaine affectif, social, sensoriel ou moteur, les sciences cognitives considèrent qu'il y a fondamentalement trois catégories de connaissances : les connaissances déclaratives, procédurales et conditionnelles. Elles ont été identifiées suite aux travaux de [Anderson 83] et de [Tardif 92]. Chacune d'elle fait appel à des stratégies d'apprentissage différentes :

Les connaissances déclaratives : elles sont fondamentalement des connaissances plutôt statiques que dynamiques. Pour permettre l'action, elles doivent être traduites en procédures et en conditions, c'est-à-dire en connaissances procédurales ou conditionnelles. Exemples des connaissances déclaratives : définitions de mots, description de l'objet, connaissance des noms des personnages, des villes...

Les connaissances procédurales : elles correspondent au comment de l'action, aux étapes et aux procédures permettant de la réaliser. En comparaison avec les connaissances déclaratives, les connaissances procédurales sont dynamiques. Exemples de connaissances procédurales : réaliser une expérience, faire un plan de travail, résoudre un problème...

Les connaissances conditionnelles : elles se réfèrent aux conditions de l'ac-

tion. Elles concernent le quand et le pourquoi. A quel moment et dans quel contexte est-il approprié d'utiliser telle ou telle stratégie, telle ou telle démarche, d'engager telle ou telle action ? Ces connaissances sont aussi celles responsables du transfert de l'apprentissage, autrement dit de la capacité d'utiliser des connaissances dans un contexte différent de celui pour, ou dans lequel elles ont été acquises. Exemples de connaissances conditionnelles : distinguer des formes déférentes, reconnaître un objet, un style artistique. . .

Si l'on se réfère aux travaux réalisés par Loffler-Laurion [Loffler-Laurion 93] et par Chamberland [Chamberland 91], les musées transmettent au visiteur des connaissances présentées sous plusieurs formes : des connaissances chiffrées, des records, des explications et descriptions, des comparaisons entre thèmes et objets, etc. En se basant sur les types de connaissances en sciences cognitives présentés plus haut, l'on peut classer les connaissances recensées par Loffler-Laurion [Loffler-Laurion 93] et Chamberland [Chamberland 91] en deux types : les connaissances conditionnelles et les connaissances déclaratives. Les connaissances conditionnelles incluent les comparaisons [Loffler-Laurion 93], les valeurs d'utilisation [Chamberland 91] et les comparaisons des contextes [Chamberland 91]. Les connaissances déclaratives incluent les connaissances chiffrées [Loffler-Laurion 93], la description, l'identification, etc. Il est à noter que les connaissances conditionnelles peuvent permettre à l'individu de distinguer entre plusieurs styles d'art. En effet, comme le met en évidence le modèle GLO [Hooper-Greenhill 03] détaillé plus loin dans ce chapitre, c'est surtout ce dernier type de connaissances qui permettrait au visiteur d'acquérir de nouvelles compétences.

2.3 L'expérience d'apprentissage au musée

Avant de présenter dans les paragraphes suivants l'expérience d'apprentissage, il convient d'examiner les termes « apprentissage » et « expérience ». Nous avons déjà défini dans la section précédente le premier terme dans un contexte générale puis dans le contexte de musée. Nous allons donc définir ici le terme « expérience ».

2.3.1 L'expérience

Si l'on se réfère à la psychologie, l'expérience comme vue par Jodelet [Jodelet 06] renvoie à un phénomène et à un concept dont la signification change d'une discipline à l'autre. Elle distingue deux définitions. La première se rapporte à la psychologie sociale, où l'expérience s'apparente au sens commun se formant par le biais des situations concrètes auxquelles se trouve confronté le sujet tout le long

du flux de son vécu. En ce sens, l'expérience renvoie à une attitude naturelle qui se déploie dans l'espace-temps de vie de l'individu.

La deuxième définition de l'expérience se rapporte à la psychologie expérimentale ou cognitive où l'expérience se situe comme un cadre empirique de la relation du sujet à son monde d'objets, elle détermine la connaissance en tant qu'elle pourvoit les informations qui font l'objet d'un processus de traitement aboutissant à des structures ou des réseaux de conservation de ces informations en mémoire. Ces structures et réseaux (que certains qualifient de représentations mentales), seront réactivés pour traiter les nouvelles informations.

Dans son étude des représentations sociales⁸, Jodelet [Jodelet 06] souligne que l'expérience ne peut être étudiée indépendamment de la notion du vécu, qui renvoie à un état que le sujet éprouve et ressent de manière émotionnelle. À côté de cette dimension vécue, l'expérience comporte une dimension cognitive dans la mesure où elle favorise une expérimentation du/et sur le monde et concourt à la construction de la réalité et des connaissances. Jodelet [Jodelet 06] met ainsi l'accent sur l'utilité de recourir à l'analyse de l'expérience, comme connaissance et comme vécu, pour comprendre comment du sens se produit, dans des situations et contextes sociaux et spatio-temporels particuliers.

Dans un contexte spatio-temporel particulier qu'est la visite du musée, c'est l'expérience en sa signification donnée par Jodelet [Jodelet 06] que nous cherchons à étudier. Cette manière d'aborder l'expérience est non seulement pertinente mais susceptible de contribuer à un enrichissement de l'approche de l'expérience visiteur. Effectivement, la définition de l'expérience que nous avons adoptée s'accorde avec les études effectuées sur les visiteurs des musées. En effet, Chamberland [Chamberland 91] grâce à une étude empirique réalisée au musée de sciences naturelles Georges-Préfontaine de Montréal, a démontré que des aspects cognitifs et affectifs interviennent dans le contact du visiteur avec les objets exposés au musée. Ces aspects se réunissent pour donner un sens à ce que ce dernier voit. L'expérience qu'a le visiteur des objets met donc en jeu deux dimensions : une dimension de connaissance et une dimension qui est de l'ordre de l'éprouvé, de l'implication psychologique de l'individu et qui a un effet de résonance directe agissant essentiellement au niveau du ressenti.

8. constituent un patrimoine commun à la culture d'appartenance, et se présentent comme un savoir que chacun des membres partage [Cavallo 93].

2.3.2 L'expérience d'apprentissage

Dans cette section, nous allons faire le point sur le concept de l'expérience d'apprentissage tel que défini par Falk et al. [Falk 00]. Comme nous avons pu le voir dans les sections précédentes, l'apprentissage dans un musée présente un caractère volontaire, libre et dirigé par les propres choix du visiteur. Partant de ce constat, Falk et al. [Falk 00] ont fixé les facteurs qui influencent l'apprentissage au musée. En effet, après avoir examiné plusieurs travaux de recherche, Falk et al. [Falk 00] ont proposé un modèle incluant douze facteurs qu'ils appellent le modèle contextuel d'apprentissage. Ce modèle est fondé essentiellement sur la théorie constructiviste de l'apprentissage à savoir :

Le contexte personnel

1. Motivations et attentes du visiteur
2. Connaissances préalables
3. Expériences préalables
4. Intérêts préalables
5. Choix et contrôle

Le contexte socioculturel

6. Médiation sociale au sein du groupe
7. Médiation assurée par les gens en dehors du groupe social

Le contexte physique

8. Organismes à l'avance
9. Orientation dans l'espace physique
10. Architecture et environnement à grande échelle
11. Conception et exposition aux objets exposés
12. Événements ultérieurs et expériences de renforcement à l'extérieur du musée

Le contexte physique : il signifie que l'apprentissage se produit dans l'environnement physique du musée. Il se traduit par un dialogue entre cet environnement physique et le visiteur. Ainsi, il faut s'attendre à une réaction des visiteurs au contexte physique du musée. Ce dernier comprend à la fois des propriétés appelées « à grande échelle » (l'éclairage, le bruit, affluence) et les propriétés à petite échelle (expositions et autres objets).

Le contexte socioculturel : exprime le fait que les êtres humains soient des créatures extrêmement sociales. Ils sont les produits de leurs cultures et relations sociales. Par conséquent, il faut s'attendre à ce que l'apprentissage au musée soit situé dans le contexte socioculturel du visiteur. Le facteur d'ordre socioculturel qui affecte principalement l'apprentissage au musée est le contexte culturel que représente le musée dans la société. D'un autre côté, les visiteurs de musées sont

fortement influencés par les interactions et les collaborations avec les individus de leur propre groupe social ainsi qu'avec d'autres individus à l'extérieur de leur propre groupe social comme les guides, les artistes ou même les autres groupes de visiteurs.

Le contexte personnel : implique que l'apprentissage au musée soit fortement lié aux propres choix du visiteur, ses motivations et ses attentes. De plus, l'expérience d'apprentissage du visiteur peut être influencée par ses connaissances et ses expériences antérieures.

Nous pouvons conclure que la conception d'apprentissage incarné par le modèle de Falk et al. [Falk 00] présente une quête à la recherche de sens, une quête à éveiller un dialogue entre le visiteur et son environnement physique et socioculturel : "The view of learning embodied in this framework is that learning can be conceptualized as a contextually driven effort to make meaning in order to survive and prosper within the world ; an effort that is best viewed as a continuous, never-ending dialogue between the individual and his other physical and sociocultural environment" ([Falk 05], page 745).

En outre, comme nous avons pu le voir plus haut dans ce chapitre, apprendre au musée consiste à comprendre et à rechercher le sens qu'incarnent les objets exposés. Cette approche profonde de l'apprentissage [Marton 93] est celle mise en avant par le modèle contextuel de Falk et al. [Falk 00]. Compte tenu de tout cela, c'est sur ce modèle que va se fonder la conception de notre système de réalité augmentée pour la visite guidée au musée.

2.3.3 Mesurer l'expérience d'apprentissage

Essayons maintenant de répondre à la deuxième question : « Comment atteindre l'expérience d'apprentissage au musée ? »

Plusieurs travaux de recherche [Hooper-Greenhill 03][Gammon 03] ont proposés des modèles qui permettent de mesurer l'expérience d'apprentissage du visiteur. Dans le cadre de cette thèse, nous allons nous intéresser au modèle GLO «Generic Learning Outcomes» donné par Hooper-Greenhill et al. [Hooper-Greenhill 03]. En effet, ce modèle conformément aux travaux réalisés par Chamberland [Chamberland 91] met l'accent sur les deux dimensions épistémologique et affective ayant un impact significatif sur l'expérience du visiteur dans son contact avec les objets exposés. GLO a été mis en place par le ministère (MLA) des musées, des librairies et archives au royaume Uni pour démontrer l'impact de ces dernières institutions dans

l'apprentissage continu. D'un autre côté, le modèle GLO est aujourd'hui largement utilisé par plusieurs musées au Royaume Uni [Graham 13]. Ce modèle est appliqué avec succès par les services de médiation de ces musées dans leurs études des expériences d'apprentissage des visiteurs. GLO reconnaît l'apprentissage comme une expérience dont les aboutissements comprennent : une augmentation de la connaissance et de la compréhension, une augmentation de compétences, un changement dans les valeurs et les attitudes, des sentiments d'amusement, d'inspiration et de créativité et une incitation à l'activité et à la progression.

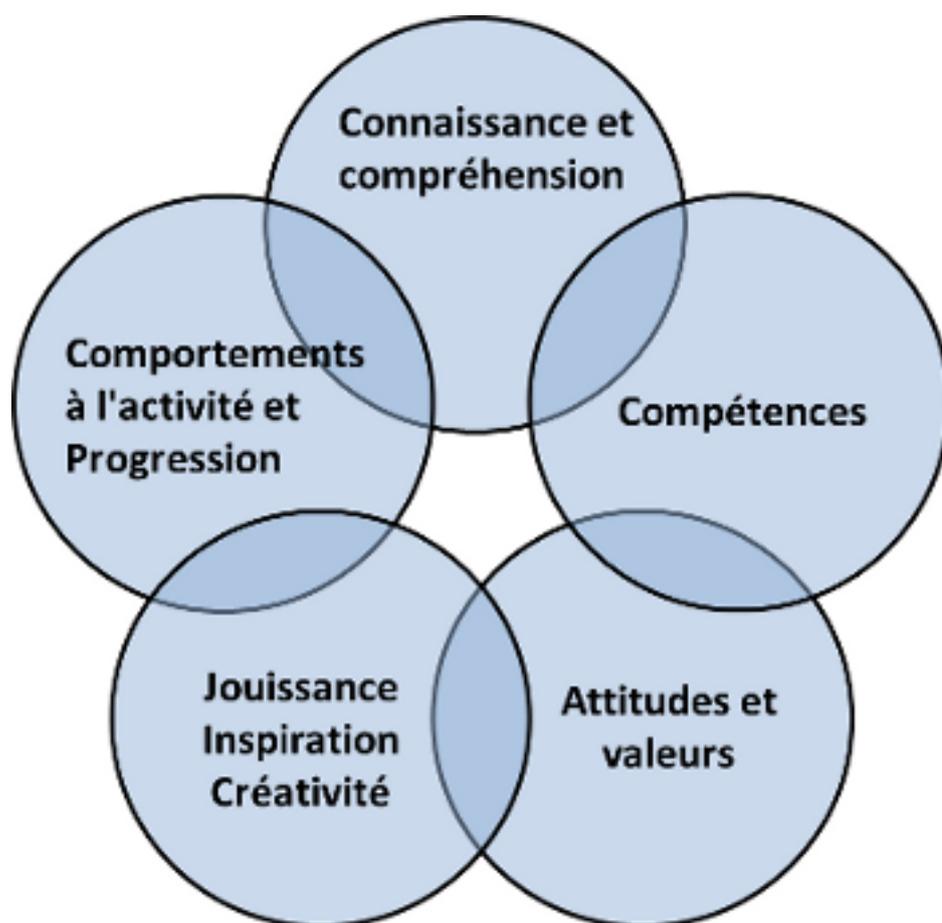


FIGURE 2.3.1 – Generic Learning Outcomes (GLO)

Connaissance et Compréhension : la visite au musée permet d'acquérir des nouvelles connaissances, de comprendre des nouveaux aspects sur des thèmes diversifiés. Elle permet également de donner du sens aux objets exposés et à faire des liens avec la vie quotidienne ou les connaissances pré-acquises.

Compétences : les compétences que l'on peut tirer d'une visite au musée sont de l'ordre intellectuel, sociale, communicatif, etc. A la sortie d'un musée d'art par exemple, l'on peut commencer à distinguer les styles d'art de différentes époques ou de différentes cultures.

Attitudes et Valeurs : à l'issue d'une visite au musée l'on peut changer son attitude envers une culture, une époque historique, un domaine artistique ou scientifique. En ce sens, le musée peut inculquer plusieurs valeurs comme la tolérance, le respect, l'empathie, etc.

Amusement, Inspiration et Créativité : lors de l'expérience d'apprentissage au musée, le visiteur éprouve un sentiment se situant entre l'ennui et l'amusement. En outre, la visite au musée peut inspirer l'individu et l'inciter à devenir plus créatif dans le cadre de sa vie quotidienne.

Activité, Comportements et Progression : ce cinquième volet du modèle inclut les intentions d'activités et d'actions que suscite la visite au musée. Il intègre également les changements de comportements de visiteur ainsi que la progression qui peut avoir lieu dans sa vie.

2.4 Le musée : le rôle de médiation

2.4.1 Définition de la médiation

Au sens large du mot, la médiation signifie d'après le dictionnaire Larousse : Médiation n.f (bas latin *mediatio*, -onis, de *mediare*, s'interposer) : le fait de servir d'intermédiaire, en particulier dans la communication : Sensibiliser le public par la médiation de la radio.

Dans le contexte muséal, la médiation culturelle est aujourd'hui un service à part entière [Lugez 04]. Elle constitue le mécanisme permettant d'établir le dialogue entre les œuvres et leurs récepteurs, les spécialistes et les publics. Les premiers efforts de réflexion puis d'effectuation en termes de médiation culturelle ont été menés dans des musées nord-américains. Cette dynamique ne s'est développée en France qu'à partir des années 1960 avec une volonté politique marquée de démocratisation culturelle et de développement des expositions temporaires. Depuis la fin des années 1980 et le début de la décennie suivante, les outils de connaissance des publics, initiatives et expérimentations se sont multipliées dans diverses institutions muséales.

Par ailleurs, comme le fait remarquer Peyrin [Peyrin 12], la médiation dans les musées recouvre aujourd'hui un ensemble de services d'accompagnement des visiteurs. Elle remplit des missions diverses de diffusion et de transmission des savoirs, de mise en relation des œuvres et publics [Lugez 04]. En plus de cette dimension éducative, la médiation porte aussi sur le confort du visiteur, le plaisir et la détection du public. Elle fournit au visiteur les renseignements pratiques dont il a besoin et suscite son envie. Elle fait donc intervenir des champs d'action variés : activités pédagogiques et actions éducatives, animations, diffusion culturelle, aide à la visite, etc. D'un autre côté, comme le signale Lugez [Lugez 04], le médiateur doit connaître les publics auxquels il s'adresse, leurs attentes et leurs besoins.

Comme nous avons pu le voir au début de ce chapitre, apprendre et découvrir constitue la première raison pour laquelle l'on visite un musée [Packer 02]. En ce sens, nous allons nous consacrer dans ce qui suit à la médiation dans son aspect éducative : transmission de connaissances et diffusion de savoirs.

Dans ce contexte, nous allons présenter dans ce qui suit les outils d'aide à la visite. Nous allons commencer par présenter les outils de médiation traditionnels, ensuite, nous allons examiner l'intérêt du numérique et en particulier l'apport de la réalité augmentée. Mais avant cela nous allons nous consacrer aux problèmes liés à la médiation dans les musées et qui ont un effet négatif sur l'expérience d'apprentissage au musée.

2.4.2 Les problèmes liés à la médiation dans les musées

Si le musée est bien un lieu éducatif, il n'en ressort pas moins certaines interrogations concernant l'impact d'une mauvaise médiation sur l'apprentissage ayant lieu au musée. Selon une étude réalisée au Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie en 2006 [Alibert 06], 52% des visiteurs pensent que dans les musées ils sont livrés à eux même sans avoir suffisamment d'explication et sans être bien guidé. Cette même étude rapporte que 56% des gens interrogés disent qu'ils iraient dans les musée uniquement s'ils disposaient plus d'information sur les expositions. Pour Burigana et al. [Burigana 13], le manque de stimulation du visiteur constitue un frein à la découverte et à l'acquisition des nouveaux savoirs. Dans cette optique, les raisons qui empêchent l'apprentissage au musée sont essentiellement une mauvaise conception de l'exposition. En outre, Dufresne-Tassé et al. [Dufresne-Tassé 96] estiment que l'apprentissage de l'adulte au musée est le résultat d'un questionnement. Les questions que se pose le visiteur l'amèneront

à apprendre. En ce sens, les auteurs précisent que le manque de questionnement revient soit à des raisons personnelles (manque d'intérêt pour le musée ou manque d'implication), soit à l'absence ou l'insuffisance d'informations fournies au visiteur.

Nous pouvons conclure de ce qui précède qu'une mauvaise médiation au musée est essentiellement dû à :

- **L'insuffisance ou absence d'informations fournies au visiteur.**
- **Le manque de stimulation des visiteurs Ces lacunes au niveau de la médiation constituent un obstacle pour un questionnement personnel et donc empêchent le visiteur d'acquérir de nouvelles connaissances.**

Dans cette optique, plusieurs travaux ont tenté d'étudier les critères à améliorer afin de rendre le musée plus efficace en matière de médiation. Ritchhart et al. [Ritchhart 07] par exemple, propose de créer un concept de « culture de pensée » lors de la visite guidée du musée, c'est-à-dire une expérience collective dans laquelle la pensée est l'élément central. Pour l'auteur, il y a un enjeu très important pour les médiateurs de musée de créer une culture de pensée pour que la visite bénéficie davantage aux visiteurs. Créer une culture de pensée consiste à stimuler le visiteur et donner un sens à la visite. Les solutions proposées sont les suivantes :

- Provoquer des attentes : stimuler le visiteur en expliquant ce qu'il va voir.
- Exploiter les ressources du musée : comparer des objets exposés côte à côte, par exemple.
- Utiliser au mieux le temps (en réalité trop court pour s'imprégner des objets), en effectuant des rappels par exemple (A quel autre objet cet objet vous fait-il penser ?)
- Inciter les visiteurs à utiliser le musée comme une ressource personnelle : source d'inspiration pour dessiner ou pour entreprendre une activité...
- Créer des habitudes au cours de la visite : toujours commencer par la description de l'objet, respecter un ordre défini.
- Favoriser le langage : encourager des discussions sur les objets, fournir des explications.
- Utiliser l'espace de façon à favoriser les interactions : s'asseoir en cercle, créer des activités.
- Encourager les interactions.

Dans ce qui suit, nous allons présenter les outils permettant de propulser la médiation au musée. Nous allons voir comment ces outils peuvent répondre aux problèmes cités ci-dessus ainsi que leurs éventuelles avantages et inconvénients.

2.5 Les moyens classiques de la médiation au musée

2.5.1 Guide conférencier-animateur

La mission du guide conférencier (appelé également guide animateur) [Lugez 04] consiste à permettre aux visiteurs de découvrir et d'apprécier le patrimoine culturel. Pour exercer son activité, ce médiateur doit avoir des connaissances approfondies dans différentes disciplines : histoire, histoire de l'art, communication, ethnographie, économie, géographie, etc. Il doit encore disposer de qualités relationnelles, avoir le sens de la pédagogie et également manifester une bonne capacité d'adaptation aux attentes des visiteurs.

Il est à souligner que la médiation humaine reste très efficace et très appréciée par le public, la visite accompagnée laisse plus volontiers une empreinte sur le visiteur [Merleau-Ponty 10]. Les visites guidées classiques sont des moyens de transmission de savoirs qui conviennent à un public habitué à ce type de médiation. Les rencontres entre médiateurs et visiteurs sont à l'origine d'un apport intellectuel et émotionnel permettant alors au musée de remplir pleinement sa mission de médiation.

2.5.2 Audio Guide

Apparus dans l'industrie au début des années 90⁹, les audio-guides constituent un dispositif d'aide à la visite [Deshayes 02]. Ils présentent un outil individuel sonore qui permet aux visiteurs de recevoir des compléments d'information sur les œuvres exposées du musée. Ils ont pour objectif de faciliter la compréhension des œuvres par les publics. Ils assurent également la traduction en langues étrangères pour les publics internationaux. Comme le signale Deshayes [Deshayes 02], les audio-guides fournissent des clés d'interprétation et ont une forte vocation pédagogique. L'un des succès de l'audio-guide réside dans le fait qu'il permet au visiteur de regarder l'œuvre et d'écouter les commentaires en même temps.

2.5.3 Les textes

De nature diverse et conçus de façon complémentaire, les textes sont intégrés à l'espace physique de musée et permettent d'identifier l'objet exposé. Ces derniers doivent tenir compte des différents niveaux d'appropriation possible par les publics. De ce fait, leur conception réclame des compétences variées et une connaissance des objets, de la communication et de la mise en forme. Lugez [Lugez 04]

9. <http://ocim.revues.org/396>

distingue trois types de textes présents au musée :

Le texte affiché : ce type de texte se trouve intégré à la muséographie . Ils correspondent aux panneaux placés dans les espaces d'exposition, aux supports du discours qu'elle véhicule et vecteurs d'informations scientifiques et techniques. Ils constituent une unité autonome mais ne sont consultés intégralement que par une minorité de visiteurs.

Le texte signalétique : la signalétique a plusieurs fonctions : orientation, information ou interprétation. Elle indique et porte l'attention sur des éléments remarquables et permet au visiteur de construire son parcours. Il s'agit surtout des dépliants aussi perçus comme une manière de toucher l'œuvre indirectement.

L'étiquette : appelée aussi le cartel, et est un petit texte écrit sur un support de petites dimensions, placé à proximité d'un objet exposé. Elle désigne, commente ou décrit l'objet auquel elle est associée et a un double rôle de désignation et de médiation. En ce sens, elle identifie l'objet exposé et facilite son appropriation par le visiteur. L'étiquette contient un titre et un texte descriptif et ne doit pas compter plus de 50 mots en moyenne¹⁰.

2.5.4 Conclusion sur les moyens classiques de la médiation

Comme le souligne Merleau-Ponty [Merleau-Ponty 10], la médiation humaine, pour tout le potentiel qu'elle présente, s'avère très efficace et très appréciée par le visiteur. Le guide conférencier s'engage dans une communication bidirectionnelle dans le but d'assurer une fonction d'interprète d'œuvres d'art. En revanche, comme dans le cas de musée basque et de l'histoire de Bayonne , les visites guidées par des conférenciers sont d'habitude plus chères que les visites libres. De plus, elles ne sont disponibles que deux jours dans semaine. Ceci pourrait être susceptible de créer une certaine réticence chez les publics envers ce genre de visite.

Par ailleurs, les audio-guides ont montré un grand succès depuis leur introduction dans les musées. Généralement gratuits et faciles à utiliser, les audio-guides sont appréciés par les visiteurs choisissant d'effectuer une visite libre (sans guide-conférencier). Néanmoins, Deshayes [Deshayes 02] remarque que contrairement au médiateur humain, ce dispositif n'offre aucune possibilité de dialogue avec le visiteur. Les textes, en particulier les étiquettes, représentent une source d'information

10. Etiquettes, cartels et notices. Direction des musées de France (DMF), MusÉofiches, 1998, 2 p. Disponible sur : <http://www.culturecommunication.gouv.fr/content/download/57184/442743/file/MUSOIMG.pdf>

importante pour les visiteurs. Mise en place par des professionnels de la conservation et de la communication, elles permettent d'interpréter les objets exposés et peuvent être appréhendés par différents types de visiteurs. Cependant, il est à noter que l'étude de McManus [McManus 92] réalisées auprès de 583 visiteurs, a montré que 48% des visiteurs ne lisent pas du tout les textes descriptifs, uniquement 12.5% des visiteurs semblent lire attentivement le texte. D'un autre côté, les gens qui lisent les étiquettes s'approprient souvent le sens et les utilisent dans leur reformulation et échanges avec leurs accompagnateurs. Cependant, dans quelques cas ces mêmes textes peuvent être des sources de confusion chez certains visiteurs [Kawashima-Bertrand 99].

Face à cela, le progrès des techniques audio-visuelles, la révolution numérique ont démultiplié les possibilités de transmission des informations et de la mise en valeur des collections. C'est ce que nous allons voir dans la section suivante.

2.6 L'intérêt de la Réalité Augmentée

Bien que les années 60 aient vu naître la réalité augmentée, la concrétisation de son implémentation, dans le domaine muséal date du début des années 90. Sa vocation est alors d'accompagner un visiteur dans sa découverte de musée en fournissant de l'information contextualisée, et ceci, en temps réel.

Pour les visiteurs

Au sein des musées et lieux d'exposition, la réalité augmentée a déjà été testée à plusieurs reprises [Damala 08][Miyashita 08][Mannion 12][Bationo Tillon 10][Bationo Tillon 11]. Elle permet de compléter le parcours classique du visiteur avec des informations complémentaires des sources d'information disponibles sur place (comme les étiquettes) [Miyashita 08]. Comme dans [Bationo Tillon 10][Bationo Tillon 11], la RA a été utilisé pour révéler des éléments invisibles dans l'œuvre d'art. D'un autre côté, plusieurs travaux rapportent que l'utilisation des systèmes de RA au musée, constitue pour les visiteurs une expérience enrichissante, intéressante et fortement apprécié par les visiteurs [Miyashita 08].

Pour les musées

Par ailleurs, l'une des forces de la réalité augmentée est d'offrir de multiples interprétations et des contenus divers aux visiteurs, sans encombrer l'espace réel des salles d'exposition. Il s'agit également d'un moyen trouvé pour exposer des œuvres qui ne peuvent pas sortir des réserves [Mannion 12]. D'un autre côté, la réalité augmentée s'avère un moyen publicitaire permettant au musée d'attirer plus

de visiteurs.

2.7 Les systèmes actuels de R.A pour la visite guidée

Plusieurs musées¹¹ dans le monde sont aujourd'hui conscients des avantages qu'ils pourraient retirer de la réalité augmentée. Mais l'introduction d'une technologie émergente dans le milieu muséale se fait généralement en plusieurs étapes dépendantes des améliorations progressives de la recherche dans le domaine. Ainsi, étalés sur près de deux décennies, divers projets, prototypes, démonstrateurs sont apparus. L'exhaustivité étant vaine, nous avons choisi de présenter des exemples de systèmes faisant référence dévolus à l'aide à la visite et ayant essentiellement des buts éducatifs.

NaviCam 1995

NaviCam [Rekimoto 95] est considéré comme l'un des systèmes pionniers de la réalité augmentée mobile. Le prototype proposé en 1995 peut contenir soit un PDA¹² soit des lunettes « see-through » de RA. L'utilisation de ce prototype au musée, permet de présenter à l'utilisateur des informations contextuelles concernant les œuvres d'arts exposées. NaviCam se base sur des codes à barre en couleurs afin d'identifier l'œuvre d'art en question. Les auteurs de NaviCam [Rekimoto 95] trouvent que les étiquettes accompagnant les œuvres d'art sont très basiques pour les experts, très difficiles pour les enfants et les étrangers. Afin de pallier à ce problème, NaviCam présente des informations contextuelles appropriées.

Virtuoso 2005

Virtuoso [Wagner 06] est un jeu éducatif conçu pour être joué à plusieurs visiteurs au musée. L'objectif du jeu est de trier une collection d'œuvres d'art en fonction de leurs dates de création. Le joueur doit placer à l'aide de son PDA chaque œuvre d'art sur un marqueur. Le jeu nécessite deux joueurs ou plus afin d'ordonner les œuvres d'art. A l'aide de l'application Virtuoso, le joueur peut accéder à des interprétations concernant l'œuvre d'art. Quand le joueur se trouve confus, il peut s'adresser à un personnage virtuel portant le même nom que l'application. Ce dernier présente des explications, des conseils aux joueurs afin de les aider à terminer le classement chronologique des œuvres d'art. Les sujets ayant

11. <http://www.museum-id.com/ideas.asp>

12. Personal Digital Assistant



FIGURE 2.7.1 – NaviCam fournit des informations textuelles à l'utilisateur

expérimenté Virtuoso reprochent le fait que le personnage virtuel ne leur regarde pas pendant l'interaction.

AR Guide 2008

Le CEDRIC¹³, en collaboration avec le musée des Beaux Arts de Rennes¹⁴, a développé pour celui-ci un prototype de guide en réalité augmenté pour l'aide à la visite [Damala 08]. Le prototype proposé s'appelle AR Guide. Ce dernier consiste en un PC portable UMPC¹⁵ de type Samsung, et équipé d'une caméra ordinaire, permettant ainsi un affichage déporté des augmentations.

Les auteurs d'AR Guide utilisent une technique sans marqueur pour la détection des objets exposés. Cependant ils ne précisent pas la nature de la technique utilisée. Une fois que l'œuvre d'art est détectée, AR Guide superpose à ce dernier un menu de cinq items : description, technique, iconographie, contexte et artiste. Pour chaque item, l'utilisateur peut accéder à divers contenus multimédia permettant d'interpréter l'œuvre d'art en question. Les divers médias employés sont : du texte,

13. Conservatoire National des Arts et Métiers, Paris, France

14. <http://www.mbar.org/>

15. Ultra Mobile PC

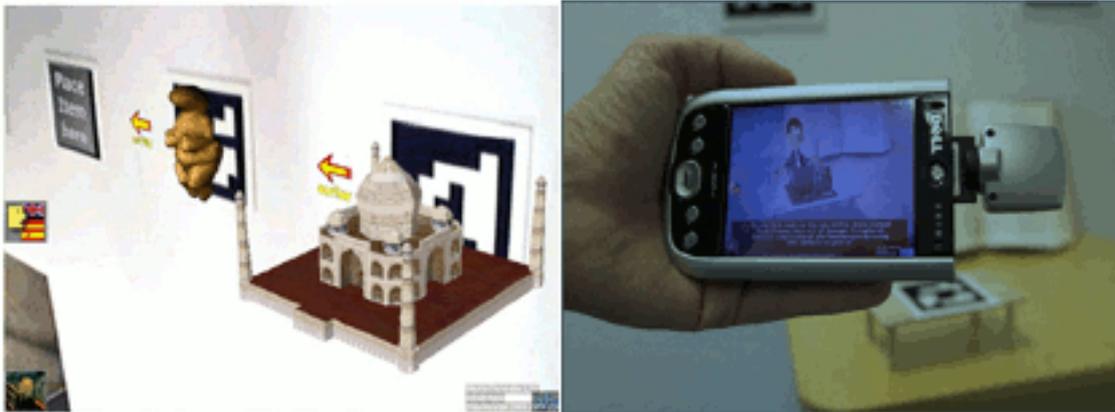


FIGURE 2.7.2 – Virtuoso

des interprétations audio, des vidéos, des images 2D et 3D. Les augmentations présentées à l'utilisateur sont donc riches. Néanmoins, les utilisateurs d'AR guide ont critiqué la redondance des informations présentés via différents médias (Par exemple les mêmes informations audio ont été reproduite d'une manière textuelle).



FIGURE 2.7.3 – Navigation dans les différents thèmes d'une œuvre d'art et accès à l'interprétation

Le projet LDML 2006-2009

Le projet LDML [Miyashita 08] a vu le jour dans le cadre d'une collaboration entre le musée du Louvre de Paris et la compagnie japonaise DNP (Dai Nippon

Printing Co. Ltd)¹⁶. Ce projet a permis de fournir deux systèmes : un système de guidage « guidance » et un système d'appréciation des œuvres d'art. Le premier système permet de guider l'utilisateur dans l'espace d'exposition (entre les salles de présentation) dans un ordre bien déterminé. Le deuxième système est utilisé à l'intérieur de chaque salle de présentation, il permet de fournir à l'utilisateur des informations l'aidant à connaître les caractéristiques de l'œuvre d'art.



FIGURE 2.7.4 – À gauche le système « salle de présentation », à droite le système de guidage

Le système de salles de présentation permet d'assurer principalement les fonctionnalités suivantes :

- Afficher du texte virtuel pour l'interprétation de l'œuvre d'art
- Afficher le modèle 3D de l'œuvre d'art dans le but de permettre à l'utilisateur de le manipuler. Dans ce cas, les auteurs énoncent la possibilité de voir le verso de l'œuvre.

Le système de guidage emploie un personnage virtuel pour guider le visiteur à travers les salles de présentation du musée. En plus, des ballons flottants virtuels ont été utilisés. Ce choix a été mis en œuvre afin de dissimuler les imprécisions de suivi, étant donné que les ballons flottants effectuent naturellement des mouvements de translation. Ceci est dans le but de ne pas nuire à l'expérience de l'utilisateur avec la technologie de la RA. Il est à noter que dans le cadre du projet LDML, le personnage virtuel a été uniquement utilisé pour communiquer des informations de direction et non pas pour assurer l'interprétation des œuvres d'art.

16. <http://www.dnp.co.jp/eng/>

Le matériel utilisé par le système des salles de présentation comprend un UMPC auquel est connectée une caméra via un câble (côté utilisateur) et un ordinateur de bureau. Le système de guidage quant à lui utilise un PC de type Fujitsu BIBLO LOOX-P. Les auteurs utilisent un suivi hybride se basant sur un suivi optique sans marqueur couplé avec la technologie inertielle. En ce qui concerne le développement des deux systèmes, ils utilisent le SDK commercial de RA Unifeye, fourni par la compagnie Metaio¹⁷.

Les auteurs précisent que l'une des limitations techniques du système de guidage consiste dans le fait qu'il ne fonctionne que pour certaines stations¹⁸ dans le musée. En plus, les augmentations prennent beaucoup de temps pour apparaître ce qui fait que les utilisateurs quittent l'endroit de l'augmentation en pensant que la RA n'y est pas disponible.

Le projet GAMME 2007-2010

GAMME [Bationo Tillon 10][Bationo Tillon 11] est acronyme de Guide Augmenté Mobile pour les Musées et Expositions. Il s'agit d'un projet de recherche industrielle financé par l'ANR et alliant entre autre l'IRISA¹⁹ et Orange Labs. Le prototype matériel proposé se base sur un UMPC de type Samsung Q1 et une webcam (caméra arrière).



FIGURE 2.7.5 – UMPC équipé d'une webcam

Le prototype matériel proposé dans le cadre de ce projet s'appuie sur un suivi hybride sans marqueur basé sur l'algorithme SLAM et sur le système inertielle

17. <http://www.metaio.com/>

18. Désignent des endroits dans le musée où existe des panneaux de guidage.

19. Plus précisément l'équipe Lagadic qui est une équipe de recherche INRIA installée à Rennes.

[Servant 10]. Afin d'assurer des fonctions éducatives au sein de musée, le prototype de guide présenté se base sur les principes énoncés par Goodman [Goodman 87]. Le premier prototype de GAMME [Bationo Tillon 10] présente grâce à la RA, les fonctionnalités analytiques suivantes :

- Trouver les couleurs complémentaires dans un œuvre d'art.
- Mettre en évidence le repentir « Pentimento »²⁰ sur un tableau.
- Explorer l'arrière d'un tableau d'un tableau ou l'on peut trouver la signature de l'artiste.
- Superposer des œuvres d'art connexes dans le but de comparaison avec celle en cours.

Les fonctionnalités de repentir et de comparaison furent avantageuses pour les utilisateurs. Cependant, les deux autres n'ont suscité aucun intérêt chez eux. En plus, les utilisateurs estiment ne pas avoir reçu suffisamment d'interprétations et d'informations concernant l'œuvre d'art.



FIGURE 2.7.6 – Une fenêtre en RA qui fait apparaître le repentir « Pentimento » à l'utilisateur

Courant 2011, les protagonistes de GAMME, ont proposé un deuxième prototype [Bationo Tillon 11]. Ce prototype, complète le premier en créant un effet émotionnel grâce à la musique, aux compositions et aux séquences audio. Néanmoins, il est à noter qu'ils ont gardé la même partie analytique proposé au niveau du premier prototype.

20. Dans le lexique des arts visuels, et en particulier dans le technolecte de la peinture, le terme « repentir » désigne une altération de l'œuvre [Leone 10]

Passeport 2012

Le système Passeport [Mannion 12] a été développé dans le cadre d'un partenariat entre le musée Britannique²¹ et la compagnie Samsung. Il s'agit d'un jeu collaboratif et éducatif de RA, se basant sur un ancien livre égyptien. Passeport emploie la plateforme gratuite de RA, Junaio²². Il se base sur le Smartphone Galaxy Apollo de Samsung et utilise la technologie de marqueurs fiduciaux. Ces marqueurs permettent aux joueurs de visualiser des animations de modèles 3D représentant des répliques d'œuvres d'art. Les animations présentées permettent aux utilisateurs de déchiffrer des codes hiéroglyphes²³, dans le but de compléter une séquence de mots existante dans l'ancien livre égyptien.



FIGURE 2.7.7 – Déchiffrer les codes hiéroglyphes à l'aide de Passport

21. <http://www.britishmuseum.org/visiting.aspx?lang=fr>

22. <http://dev.metaio.com/junaio/>

23. Un hiéroglyphe est un caractère de la plus ancienne écriture égyptienne.

2.7.1 Conclusion sur les systèmes de RA

Les systèmes de réalité augmentée dévolus à l'aide à la visite dans le domaine muséal et ayant des objectifs éducatifs, font ces dernières années l'objet de beaucoup d'attention par les musées. Parmi les systèmes de réalité augmentée proposés dans ce domaine, l'on peut distinguer deux catégories : des jeux sérieux collaboratifs [Wagner 06][Mannion 12] et des systèmes analytiques²⁴ ayant pour fonction principale l'interprétation des objets exposés [Damala 08][Miyashita 08][Bationo Tillon 10][Bationo Tillon 11]. La première catégorie est essentiellement destinée aux enfants, aux groupes et aux visites en familles. La deuxième catégorie vise autant les visites individuelles que les visites en groupe. Comme nous pouvons le constater, c'est cette dernière catégorie qui fait l'objet de plus d'intérêt pour les musées.

En ce qui concerne la deuxième catégorie, nous venons de voir que les solutions proposées sont parfois assez différentes en termes d'interaction et d'implémentation, et pourraient donc être complémentaires aux interprétations apportées au visiteur. Nous pensons donc que les concepts de la réalité augmentée permettraient de « pousser » plus loin les détails présentés au visiteur et ceci en assurant une forte connexion entre ces derniers et le monde réel.

Comme le fait remarquer Miyashita et al. [Miyashita 08], d'un point de vue technique, il reste tant de défis techniques à lever afin de permettre aux systèmes de RA, de passer du stade de prototype au stade des systèmes complètement déployés dans les musées.

2.8 Naissance d'une problématique

Depuis que l'éducation et l'ouverture au grand public sont au centre de leurs missions, les musées d'art et d'histoire mettent en place des structures pour transmettre les connaissances et savoirs qu'ils produisent à un maximum de visiteurs. L'objectif, par le biais de la médiation, est de créer les conditions propices à un contact riche avec la collection, et de donner au visiteur les outils nécessaires pour qu'il puisse, à partir de ses expériences muséales, créer du sens. Les informations que le musée transmet sont les connaissances autour des objets de la collection, fruits des recherches menées par les conservateurs. L'information présente (et présentée) dans le musée est d'un large spectre. Outre les savoirs produits par les conservateurs, l'information réside aussi dans l'objet lui-même, dans son histoire,

24. Cette appellation est adoptée de [Bationo Tillon 11]

que sa matérialité même raconte.

Nous avons présenté dans ce chapitre les problèmes liés à la médiation dans les musées. On peut par exemple citer le manque d'accompagnement des visiteurs. En ce sens, les visiteurs se sentent livrés à eux-mêmes et estiment ne pas avoir suffisamment d'informations.

Au fil de ce chapitre, nous avons présenté les outils traditionnels de la médiation. Nous avons montré que ces derniers à cause des inconvénients qu'ils présentent ne permettent pas de bien remplir l'une des plus importantes missions de la médiation, qui est faire du musée un réel environnement pour l'acquisition des nouveaux savoirs. Mais avec l'avancement de nouvelles technologies de l'information et de la communication, de nouvelles modalités d'assistance à la visite sont apparues. En ce sens, nous avons passé en revue l'ensemble des prototypes d'aide du futur. Nous avons proposé une classification en deux catégories de ces systèmes. La première dite des « systèmes analytiques » regroupe l'ensemble des systèmes permettant d'interpréter les œuvres d'art. La deuxième renvoie aux jeux sérieux collaboratifs. La première catégorie fait l'objet de plus d'intérêt vu qu'elle est destinée à une plus large audience et se trouve appropriée aux visites en groupe comme aux visites individuelles. En plus, les systèmes s'inscrivant sous cette catégorie, malgré quelques limitations techniques, font preuve d'une maturité technologique leur permettant de prétendre à une intégration totale et relativement proche dans le domaine muséale. L'objectif principal de ces systèmes est de présenter à l'utilisateur des informations, en connexion avec le monde réel, afin de les aider à appréhender l'œuvre d'art. Nous avons vu que cela n'est pas forcément le cas !

Notre problématique pourrait alors se formuler de la manière suivante : « **Comment présenter les informations concernant l'œuvre d'art de telle sorte à impliquer pleinement le visiteur du musée dans une expérience d'apprentissage ?** »

Avec la réalité augmentée, nous pouvons désormais envisager un système d'aide à la visite permettant des capacités de présentation de l'information et d'interaction intéressantes. C'est ce que nous allons présenter dans le chapitre suivant. Nous précisons au lecteur qu'il ne s'agit pas de proposer un outil supplémentaire qui ne ferait que répéter les schémas actuels d'aide à la visite ; mais plutôt d'interroger les différentes études de recherche, motivations et pratiques propres au musée, afin de voir si et où la technologie pourrait intervenir pour enrichir le rapport aux œuvres.

Conclusion

Nous nous sommes consacré au début de ce chapitre à étudier les attentes des visiteurs envers le musée. En se basant sur plusieurs travaux de recherches dans le contexte muséal, nous avons pu identifier le besoin d'apprendre et de découvrir comme une quête centrale du visiteur.

D'un point de vue cognitif, l'apprentissage au musée suit une stratégie constructiviste [Hein 98]. D'un point de vue phénoménologique, l'apprentissage est une expérience pour l'individu [Marton 93]. Falk et al. [Falk 00] concilient les deux visions en proposant un modèle contextuel de l'expérience d'apprentissage au musée basé essentiellement sur une approche constructiviste. Nous allons donc au niveau des deux chapitres suivants, baser nos propositions pour soutenir l'expérience d'apprentissage au musée sur le modèle de Falk et al. [Falk 00].

Chapitre 3

La visite guidée au musée : Propositions

Introduction

Il est évident qu'au cœur de toute activité de médiation se trouve l'homme. En ce sens, il est à noter que chaque individu a une préférence pour un mode de représentation de l'information par rapport à un autre [Bruner 60]. De ce fait, comment concevoir un système d'aide à la visite qui prend en compte les spécificités de chaque visiteur ?

Nous avons vu au chapitre précédent la préférence des visiteurs pour la médiation assurée par les guides conférenciers ou animateurs. Comment donc faire en sorte que ce mode de médiation soit représenté ?

C'est à la lumière de toutes ces constatations que nous allons proposer, grâce aux NTIC, et plus particulièrement à la réalité augmentée, nos paradigmes d'interaction visant à soutenir l'expérience d'apprentissage au musée.

3.1 Les modes de représentation du savoir

Selon Bruner [Bruner 60], la représentation des connaissances passe par plusieurs modes appelés encore canaux de communication « communication channels » par Hooper-Greenhill [Hooper Greenhill 94]. Ces modes de représentation du savoir permettent la transmission des informations à autrui. Bruner [Bruner 60] distingue trois modes :

Mode symbolique : le mode symbolique est une représentation abstraite de l'information par des mots « outils linguistiques » ou des codes divers comme les symboles mathématiques. Hooper-Greenhill [Hooper Greenhill 94] précise que dans le cadre du musée, le mode symbolique est très présent à travers les textes descriptifs des objets exposés.

Mode iconique : le mode iconique se base essentiellement sur les images afin de représenter un objet ou une information. Il permet l'apprentissage selon des représentations 2D ou 3D de l'information. Dans le cadre muséal, cette forme de transmission des connaissances passe par l'usage des symboles iconographiques : des images [Hooper Greenhill 94].

Mode éactif : le mode éactif propose que l'on apprenne par l'action et la manipulation. L'information passe par l'action. Connaître, c'est d'abord agir. On connaît quelque chose parce qu'on « sait le faire ». Pour apprendre, on a besoin de manipuler les données, de les percevoir par les sens. Dans le contexte du musée, l'apprentissage via le mode éactif passe par les personnes responsables de la médiation (comme les guides), les événements, les activités. « . . . The enactive mode, consists of learning through using real things, such as objects, or learning through people and events or through activities » ([Hooper Greenhill 94], page 144).

Au cours du développement intellectuel de l'être humain, les modes de représentation évoluent du mode éactif au mode symbolique en passant par le mode iconique. Cependant, Bruner [Bruner 60] souligne qu'il ne s'agit pas de trois stades de développement liés à l'âge et à la maturation. Une fois que les trois systèmes de représentation sont acquis, ils commencent à fonctionner en parallèles pour traiter l'information. D'après cet auteur, « Ce qui est toujours intéressant, en ce qui concerne la nature du développement intellectuel, est que ce dernier semble parcourir ces trois systèmes jusqu'à ce que l'homme soit capable de les commander tous les trois. ». Plus encore, la personne tend à privilégier pour l'apprentissage l'un et/ou l'autre des trois modes selon ses choix, ses préférences, ses habitudes et encore selon la situation dans laquelle elle se trouve.

En ce sens, Hooper-Greenhill dans le chapitre « Museums : ideal learning environments » de son livre [Hooper Greenhill 94], insiste sur le fait que si l'on présente un seul mode de communication d'information au musée, cela va restreindre le nombre de visiteurs avec lesquels les objets exposés vont pouvoir communiquer. D'après lui : « The more opportunities for different modes of contact with ideas that are offered in any exhibition or presentation, the more possibilities will open up for the communication process » ([Hooper Greenhill 94], page 146).

C'est à la lumière de ces indications que nous rendons compte dans nos propositions, puisant leurs principes de la réalité augmentée, des trois modes de représentation du savoir. Comme nous allons le voir dans la section suivante, le mode éactif est représenté par un guide humain virtuel simulant la présence d'un guide humain et appuyé par le paradigme « Sélection ». Le mode symbolique est traduit par le paradigme d'interaction appelé « Documentation ». Le paradigme « Reconstruction » renvoie quant à lui au mode iconique.

3.2 Propositions pour soutenir l'apprentissage au musée

3.2.1 Guide humain virtuel

Les humains virtuels se veulent un moyen de communication intuitif et naturel, qui rappelle la communication entre un être humain et un autre. En effet, en s'appuyant sur un mode de communication multimodal, un guide virtuel peut joindre une expression à la parole, mais également un geste à la parole [Bolt 80]. L'idée exposée dans cette thèse repose sur l'utilisation des humains virtuels comme interface avec le monde réel, afin d'augmenter les visites guidées.

3.2.1.1 Importance de simulation de la présence humaine dans un musée

Les études réalisées par McManus [McManus 92], consultante des musées en Grande Bretagne, auprès de 1572 visiteurs, mettent en évidence le besoin d'interaction verbale des visiteurs au sein d'un musée. McManus [McManus 92] rapporte que lorsque les visiteurs lisent les textes explicatifs, ils ont l'impression que quelqu'un leur parle ou leur raconte une histoire. De plus, les visiteurs parlent de l'objet exposé, posent des questions, émettent des constations ou répondent aux commentaires contenus dans le texte explicatif associé à l'objet. En ce sens, comme le confirme cet auteur, le visiteur cherche à établir un rapport de conversation avec une personne susceptible de lui présenter l'objet exposé :

D'un autre côté, Hooper-Greenhill [Hooper Greenhill 94] insiste sur l'importance de l'apprentissage « face à face » dans le cadre de musée. Cette forme d'apprentissage fait intervenir les ressources humaines de musée responsables de la médiation qui sont essentiellement les guides (animateur ou conférencier). Le rôle du guide au sein du musée est ici très important. Il entame une relation avec le public de façon à ce qu'il y ait un retour et que le visiteur puisse interagir d'une façon active

avec les connaissances qui lui sont présentées. Comme nous avons pu le voir dans la section précédente, ce modèle de représentation et de transmission de connaissances fait appel au mode enactive.



Nous précisons au lecteur que la “coprésence” est à l’origine un terme propre à la sociologie. Il s’agit d’un concept qui décrit les conditions dans lesquelles un individu humain interagit physiquement avec un autre, face à face [Zhao 03].

Ce concept a été repris en réalité virtuelle pour décrire le fait que l’on sent l’existence des autres participants ainsi que leur présence effective dans l’environnement virtuel. Cela traduit aussi le sentiment que l’on est en train d’interagir avec des personnes réelles [Casanueva 01].

Nous pouvons conclure que les recherches en muséologie [McManus 92], en conjonction avec la sociologie [Hooper Greenhill 94], insistent sur l’importance de la présence d’un guide humain pour l’accompagnement du visiteur. Le rôle de ce dernier s’avère important dans le processus de transmission des connaissances. En outre, nous avons pu voir au chapitre précédent l’appréciation des publics pour cette forme de médiation.

Ainsi, à la lumière de ces constatations, nous nous proposons de simuler la présence d’un guide humain. Ceci peut se faire par le biais d’un guide humain virtuel. En effet, plusieurs travaux [Gerhard 05][Luo 14] en réalité virtuelle et en réalité augmentée ont prouvé la capacité des humains virtuels de créer un sentiment de coprésence chez l’utilisateur que ce soit dans un environnement virtuel ou dans un environnement mixte.

Nous allons voir un peu plus loin dans cette section que les guides humains virtuels peuvent faire intervenir une communication verbale et non verbale. Ceci rejoint notre but de produire un environnement d’apprentissage face à face au sens de Hooper-Greenhill [Hooper Greenhill 94].

3.2.1.2 Qu’est-ce qu’un guide humain virtuel ?

Définition terminologique

D’un point de vue étymologique, l’expression « guide humain virtuel » est une association entre l’expression « humain virtuel » et le terme « guide ». Comme

apparaissant dans le dictionnaire Larousse, le terme guide signifie : personne qui donne une direction morale, intellectuelle, qui dirige quelqu'un, un pays, leurs actions, etc. Dans le cadre de cette thèse, nous désignons par guide celui qui va être chargé d'accompagner, de diriger et d'informer l'utilisateur lors de sa visite au musée.

En ce qui concerne l'expression « humain virtuel », d'après Fuchs et al. [Fuchs 09], il s'agit d'une modélisation de l'être humain, donnant l'illusion à l'utilisateur de percevoir et d'interagir avec un être humain. Les recherches sur les humains virtuels s'inscrivent à la croisée de plusieurs disciplines comme l'informatique, les sciences de l'ingénierie, les sciences humaines et les sciences cognitives. Étant donné que les sciences humaines s'intéressent principalement à appréhender les comportements humains, nous avons trouvé la définition d'Alessi et al. [Alessi 00] fort intéressante :

Un humain virtuel est une représentation numérique 3D du corpus humain. Afin de doter la représentation numérique 3D d'une allure humaine, les niveaux suivants doivent être respectés et modélisés :

- Apparence humaine : avoir une structure comme la structure du squelette humain.
- Animation : l'ensemble des mouvements du corps et du visage du corpus virtuel traduisant les mouvements et les émotions humaines. L'animation de l'humain virtuel doit respecter les lois de la biomécanique humaine.
- Autonomie : la simulation du raisonnement humain permettant de doter l'humain virtuel d'une certaine autonomie, ce qui va lui permettre d'agir et de se comporter comme les êtres humains.
- Interaction : ce niveau dépend du type de l'application et des fonctions assumées par l'humain virtuel à l'intérieur de l'environnement virtuel.

Modéliser l'être humain dans sa globalité, reste jusqu'à nos jours un sujet de recherche très ambitieux voir utopique. Étant donné la complexité de la modélisation de l'être humain, plusieurs travaux de développement et de recherche se sont intéressés à la reproduction en partie de l'être humain. En fait, selon l'objectif de l'application, on cherche à modéliser un et/ou l'autre de ces niveaux caractérisant l'être humain : le niveau de motricité, le niveau de comportement, le niveau de l'interaction avec l'utilisateur, etc.

Définition Fonctionnelle

Selon Fuchs et al. citeFuchs :2009, un humain virtuel peut avoir différents usages :

Un avatar, c'est-à-dire une représentation de l'utilisateur immergé dans l'environnement virtuel;

Un personnage virtuel figuratif dans l'environnement virtuel;

Un personnage virtuel qui va interagir avec l'utilisateur.

Dans le cadre de cette thèse, c'est le troisième cas d'usage qui nous intéresse. Toujours selon Fuchs et al. citeFuchs :2009, l'interaction entre l'utilisateur réel et l'humain virtuel peut être sensorimotrice et(ou) cognitive. Au niveau sensorimoteur, l'utilisateur et l'humain virtuel peuvent participer ensemble à une tâche physique : par exemple, ils manipulent ensemble un objet virtuel, l'utilisateur pouvant ressentir les efforts physiques de l'humain virtuel via une interface à retour d'effort. Au niveau cognitif, l'utilisateur et l'humain virtuel peuvent dialoguer : par exemple, lors d'une séance de formation en environnement virtuel, un tuteur virtuel explique à l'utilisateur/apprenant comment faire une tâche.

Lors d'une visite au musée, c'est plutôt une interaction de type cognitif que l'on cherche à mettre en œuvre. Comme nous allons le voir dans la section suivante, le guide humain virtuel est chargé d'interpréter et d'expliquer les œuvres d'arts. En plus, répondre aux questions et aux commentaires de l'utilisateur nous paraît très intéressant dans la création d'une visite interactive. Cependant, cet aspect n'est pas encore traité au niveau du système MARTS¹.

3.2.1.3 Communication verbale et non verbale

La communication verbale passe par la voix. Elle désigne l'ensemble des éléments d'information transmis par l'humain virtuel en utilisant la parole. Si une partie de la communication humaine est verbale, les expressions faciales, le regard, la désignation, les gestes et la distance physique entre les individus sont autant de canaux de communication additionnels [Short 93]. Grâce à cela, les personnes tendent à minimiser leurs efforts de communication en mixant les modalités, plutôt que de faire usage d'informations verbales redondantes [Resnick 91]. **Ceci met en lumière toute l'importance que revêt l'utilisation en même temps de plusieurs mécanismes de communication lorsque l'interlocuteur virtuel vise à transmettre des informations à l'utilisateur.**

1. Le système MARTS sera détaillé au chapitre suivant.

Les gestes de communication

Bien qu'il existe une multitude de classifications de gestes exécutés durant différents travaux [McNeill 92] [Kwenzi Mikala 93], tous sont d'accord pour faire une distinction entre les gestes déictiques et les gestes représentant le mouvement des mains. Étant donné que la tâche principale de l'interlocuteur virtuel est d'interpréter un objet exposé et de raconter son histoire au visiteur, nous avons trouvé le travail de Kwenzi-Mikala [Kwenzi Mikala 93] très intéressant. Voici selon lui les définitions des gestes accompagnant la narration orale et leurs fonctions possibles lors de la présentation d'un objet exposé :

- Les gestes référentiels : ce sont des gestes liés à l'activité discursive et qui renvoient à une personne ou à un objet. Ces gestes sont essentiellement des gestes déictiques et les gestes illustratifs. Les gestes déictiques sont des gestes de présentation, de désignation ou de pointage. Les gestes illustratifs quant à eux renvoient aux gestes spatiographiques, pictographiques et kinémimiques décrivant respectivement les rapports spatiaux entre le locuteur et l'objet, une forme et les qualités structurales d'un objet ou les qualités morphologiques d'une personne et la mimique d'une action.
- Les gestes expressifs : ce sont des gestes liés à l'activité discursive et qui sont des marques d'expression du narrateur, qui qualifient ses sentiments et traduisent ses émotions. Les gestes expressifs sont surtout des mimiques faciales, par exemple le narrateur sourit pour manifester sa joie.
- Les emblèmes : appelés aussi quasi linguistiques, les emblèmes sont des gestes conventionnels propres à chaque culture. Ils peuvent être des substituts de la parole.
- Les gestes phatiques : ce sont des gestes qui servent à établir et à maintenir le contact entre le narrateur et l'auditoire.
- Les gestes rythmiques : ils englobent le battement des mains qui accompagne le discours du narrateur et aussi les gestes qui marquent la fin des énoncés.
- Les gestes extra-communicatifs : On range dans cette catégorie les gestes autocentrés non liés au sujet du locuteur : les mouvements de confort, les grattages et l'essuyage des mains. De l'étude qualitative réalisée par Kwenzi-Mikala [Kwenzi Mikala 93], il en sort que 86% des gestes sont référentiels dans le cadre d'une communication orale. Ceci nous a amené à expliquer ci-dessous les couplages entre les gestes référentiels et la parole dans le discours du guide humain virtuel.

Les références ostensives

Comme nous venons de le voir, les gestes référentiels sont des gestes très récurrents dans la communication orale. Les recherches en plusieurs disciplines - de la linguistique à la psychologie de l'apprentissage - soutiennent l'idée que ces gestes ont un lien très étroit avec les mots qu'ils accompagnent [Kelly 08][Louwerse 05][Roth 01]. En particulier, les gestes déictiques sont utilisés pour clarifier ou améliorer un message [Kelly 08]. Le geste de désignation est alors utilisé pour se référer à un objet et le désigner par le doigt afin de mieux attirer l'attention de l'individu sur ce dernier. Rappelons au lecteur que la principale mission du guide humain virtuel est d'apporter des nouvelles connaissances au visiteur, d'interpréter et d'expliquer les objets exposés. En ce sens, comme mis en avant par Spencer et al. [Kelly 08] et Roth [Roth 01], leur utilisation s'avère très intéressante pour renforcer la compréhension de l'individu du sujet traité.

Comme nous l'avons déjà précisé, dans une situation de communication, nous retrouvons le geste et le discours qui lui est associé. Généralement, le discours contient alors des termes déictiques en rapport à la désignation. Les deixis sont les termes ayant recours à la situation de communication. Dans l'exemple « Regardez cet œuvre d'art ! », le groupe nominal « cet œuvre d'art » constitue une forme de deixis. À cet énoncé est associé le geste de désignation, que l'on appelle alors ici geste déictique (ou geste ostensif). L'association d'un geste déictique et de deixis est alors nommée « référence ostensive ».

Pour résumer, quand il s'agit de montrer une œuvre d'art particulière au visiteur, le locuteur virtuel utilise une référence ostensive de cette manière : il dit « regardez cette œuvre d'art » tout en pointant du doigt l'œuvre exposée. En revanche, dans certains cas, quand il s'agit de désigner une zone particulière sur un objet, il est possible que le geste de désignation puisse introduire une ambiguïté, notamment à cause de l'imprécision de la direction de pointage ou encore à cause de la complexité de sa mise en œuvre. Nous verrons dans nos propositions comment contrecarrer cet effet, grâce à une deuxième utilisation des références ostensives combinant notre paradigme d'interaction appelé « Sélection » et la parole. Cette manière d'utiliser les références ostensives est expliquée dans la section suivante.

3.2.2 Paradigmes d'interaction

3.2.2.1 Sélection

Quand on veut expliquer au visiteur un objet exposé, il est très récurrent que l'on a besoin de se référer à des zones ou endroits précis sur ce dernier. Au musée basque par exemple, pour présenter l'œuvre d'art intitulée « Entrée du Duc d'Or-

léans en 1839 » (figure), on a besoin au cours des interprétations de désigner entre autres l'arc de triomphe provisoire construit de bois et le véhicule du duc. C'est aussi le cas de l'œuvre intitulée « Le confluent de la Nive et de l'Adour ». Pour ce tableau particulier, le visiteur doit pouvoir localiser « la monumentale porte de France » et « le buste de Louis XV » situés au premier plan du tableau. Au deuxième plan du tableau, on doit pouvoir distinguer entre autres « le clocher des Augustins » et le « dôme bulbeux ». Il est évident qu'en l'absence d'une référence visuelle, le visiteur se trouvera perdu face à tous ces éléments². Il se peut aussi qu'il se sente confus voir même ennuyé et perdra l'envie de poursuivre la présentation de l'œuvre d'art.

Pour éviter ce genre de problèmes, nous pensons qu'il est indispensable de mettre en place un paradigme d'interaction permettant à l'utilisateur de distinguer facilement l'élément référencé sur l'objet exposé. Dans cette optique, au lieu de mettre en œuvre des explications redondantes pour situer un tel élément, nous faisons appel au paradigme d'interaction que nous appelons « Sélection ». Cette simple interaction est représentative du geste déictique. Le guide humain virtuel peut clairement pointer une zone particulière (zone d'intérêt) sur un objet exposé. Le visiteur verra donc apparaître devant ses yeux un modèle iconographique (rectangle) à l'endroit exact où il doit porter son attention, levant ainsi toute ambiguïté (Figure 3.2.1).

Nous précisons au lecteur que lorsqu'il s'agit de désigner plusieurs petites zones sur un objet, le pointage avec le doigt, comme présenté dans la section précédente, s'avère non approprié. En effet, le guide humain virtuel va perpétuellement cacher l'œuvre aux yeux du visiteur ce qui peut gêner ce dernier. D'un autre côté, le pointage du doigt de certains éléments peut entraîner de l'ambiguïté pour l'utilisateur ayant pour cause la mauvaise perception de l'interaction. Ce dernier fait revient à la complexité de la mise en œuvre technique de ce geste avec une forte précision quand il s'agit de référencer des petits éléments. En effet, comme le montre l'étude expérimentale réalisée par Clair et al. [Clair 11] il est difficile, voir impossible de distinguer de petites cibles avec le simple geste de la main. Dans ce contexte, les auteurs expliquent que les participants se sont trouvés dans la majorité des cas incapables de distinguer la bonne cible désignée par un robot en utilisant le geste de la main. Pour pallier à ce problème, les auteurs proposent de combiner plusieurs modalités pour mettre en place le geste de la désignation. À la lumière de ces constatations, nous proposons d'utiliser une représentation symbolique du geste déictique « Sélection » pour le référencement de petites cibles sur un objet exposé. La mise en œuvre de ce paradigme d'interaction est expliquée au chapitre suivant.

2. Dans le reste de ce mémoire, nous appelons ces éléments zones d'intérêt.

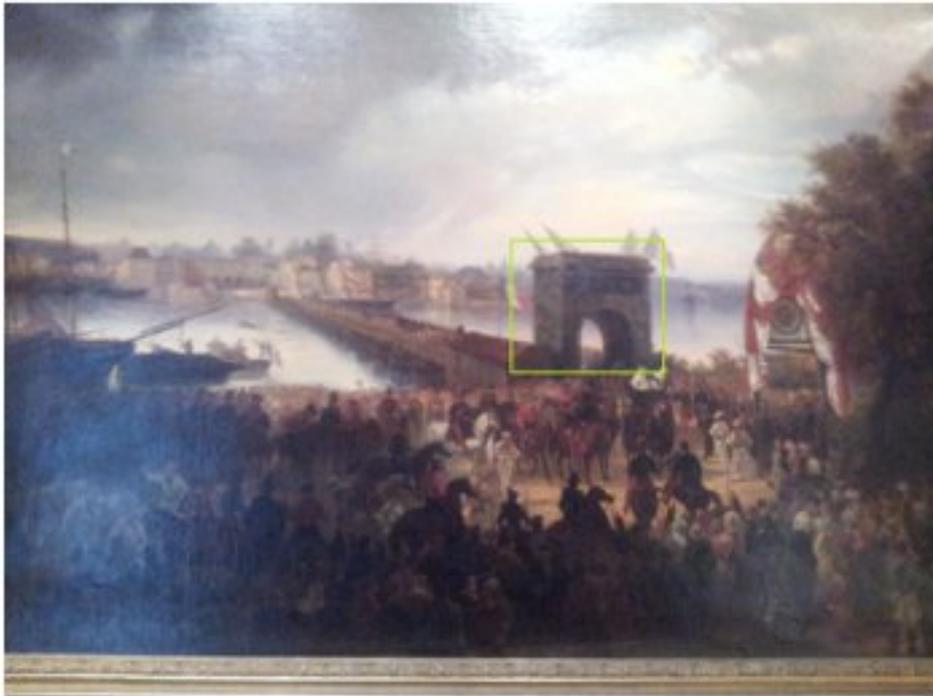


FIGURE 3.2.1 – Principe de « Sélection » montrant l'arc de triomphe provisoire à l'entrée de la place Réduit de Bayonne

3.2.2.2 Reconstruction

Grâce à la réalité augmentée, nous avons la possibilité de reconstruire en 3D un objet dégradé (figure 3.2.2). En effet, nous pouvons ajouter des éléments virtuels au regard du visiteur de manière à faire apparaître les morceaux manquants ou le modèle 3D entier de l'objet exposé si ce dernier est dans un état très dégradé. Le but donc est de montrer au visiteur à quoi ressemblerait l'objet exposé en son état original. Pour cette seconde interaction, le visiteur verra alors apparaître devant lui la représentation 3D de l'objet considéré, et ce de manière complètement recalée par rapport à la réalité. Clairement fondée sur le principe de visibilité augmentée de la RA, la « reconstruction » oblige en amont à posséder les modèles 3D adéquats et d'être en mesure de recalibrer l'information virtuelle.



FIGURE 3.2.2 – Chaland monoxyle de moyen âge

3.2.2.3 Documentation

Comme nous l'avons pu le voir au début de ce chapitre, Hooper-Greenhill dans le chapitre intitulé « Museums : ideal learning environments » de son livre [Hooper Greenhill 94], insiste sur l'importance de combiner les trois modes de représentation d'informations au musée. Ceci est dans le but de se mettre à la portée des différents visiteurs

ayant des préférences pour l'un et/ou l'autre de ces modes lors de l'apprentissage. Le mode symbolique est très présent aux musées à travers le texte. Au musée basque par exemple, on trouve des étiquettes ayant pour titre le nom de l'objet, son auteur et sa date de création et pour corps le texte descriptif. Afin de mettre en avant ces informations textuelles jugées très importantes par les médiateurs (informations contenues dans le titre de l'étiquette)³, le guide humain virtuel les référence dans son discours. Par exemple pour souligner le nom de l'œuvre, ce dernier dit : « Comme vous pouvez le lire en haut de l'étiquette, l'œuvre d'art s'appelle le Confluent de la Nive et de l'Adour ». Cependant, comme nous allons le voir dans la section suivante, dans certains cas l'information textuelle est considérée inaccessible par l'utilisateur. Dans cette optique, il nous est donné grâce à la réalité augmentée de pouvoir afficher un texte virtuel renseignant le nom de l'œuvre et le nom de son auteur, en symétrie avec le discours du guide humain virtuel (figure 3.2.3).

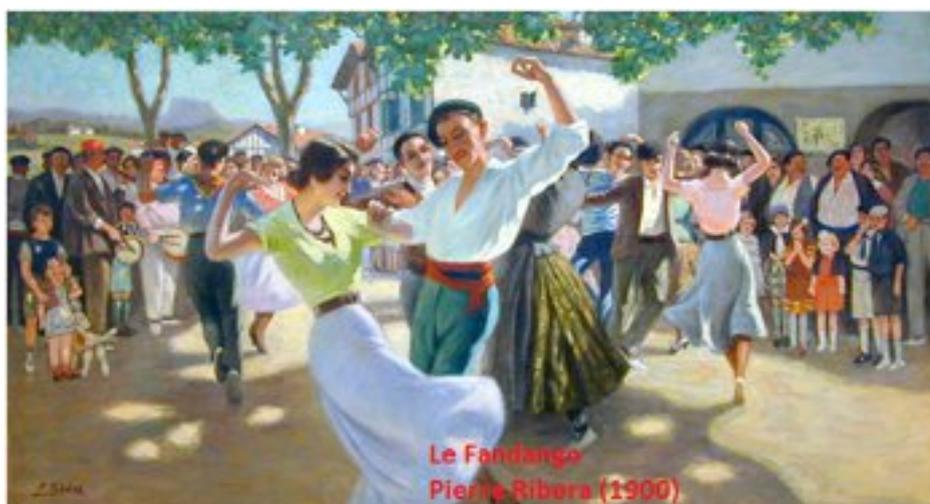


FIGURE 3.2.3 – Le principe de la "Documentation" virtuelle

Ce paradigme d'interaction permettant de se référer à l'information textuelle existante ou de l'afficher en mode virtuel est appelé « Documentation ».

3.2.2.4 Adaptations de l'interaction à l'environnement de l'utilisateur

À moins de monopoliser l'espace, il arrive souvent au musée que l'on doive se reculer d'une œuvre d'art afin de permettre à d'autres visiteurs de pouvoir également en profiter. Par conséquent, les informations textuelles accompagnant cette

3. Ces informations sont mises en évidence dans l'étiquette accompagnant l'objet exposé.

œuvre deviennent illisibles puisqu'on s'en éloigne. Dans ce contexte, il ne nous est plus possible dans le cadre de la « Documentation » de nous référer à l'étiquette accompagnant l'œuvre et nous devons donc passer à l'affichage d'un texte virtuel.

Néanmoins, au moment de forte affluence devant un œuvre exposé, il se trouve que ce dernier soit davantage occulté pour le visiteur. Dans ce contexte, comme le signale Falk [Falk 05], “a crowd of visitors at an important/preferred exhibit causes the visitor to skip that exhibit” (page 771). Lors de telle situation, tous nos paradigmes d'interaction ne peuvent plus être mis en œuvre directement sur l'objet exposé. Pour pallier à ce genre de problème, nous pensons qu'il est intéressant de permettre au visiteur de visualiser une animation mettant en œuvre nos paradigmes d'interaction sur l'écran de son smartphone. Il s'agit dans ce cas d'un mode de visualisation indirecte de l'interaction qui est offert par une transformation de l'interface homme machine de l'application.

Les deux solutions⁴ présentées ci-dessus permettent à nos paradigmes d'interaction de s'adapter à l'environnement de l'utilisateur. De plus, elles permettent de rendre compte du contexte physique du musée comme vu par Falk [Falk 05]. Nous rappelons au lecteur que le contexte physique⁵ fait partie du modèle contextuel d'apprentissage au musée que nous avons détaillé au chapitre précédent.

Par ailleurs, au regard du concept de la plasticité (voir ci-dessous), ajouter du texte virtuel quand l'information textuelle devient inaccessible à l'utilisateur est considéré comme une adaptation à l'environnement de ce dernier. Ce type de réaction aux changements survenant dans l'environnement⁶ de l'utilisateur renvoie à une **évolution**⁷ de l'interface homme machine faisant appel à une stratégie **d'adaptation par ajout** [Calvary 04]. Au sens du même concept, transformer la manière dont est présenté l'interface homme machine est aussi une évolution⁸ de cette dernière mais qui relève d'un **remodelage** [Calvary 04]. Ceci permet dans notre cas de passer d'une visualisation directe (sur l'objet exposé) des paradigmes

4. La première solution désigne l'affichage d'un texte virtuel, la deuxième désigne le passage d'un mode de visualisation direct (sur l'objet exposé) des paradigmes d'interaction à un mode de visualisation indirect (sur l'écran du smartphone).

5. L'affluence étant un facteur décrivant le contexte physique du musée

6. L'environnement au sens de l'équipe de l'ingénierie de l'interaction homme machine du Laboratoire d'Informatique de Grenoble (LIG), est présenté dans [Calvary 04]. Il rejoint le sens de l'environnement physique donné par Falk [Falk 05] pour le cas particulier du contexte physique de musée.

7. L'évolution dans ce cas se caractérise par une intervention sur la cible qui est dans ce cas l'environnement.

8. L'évolution de l'interface homme machine dans ce deuxième cas se base sur un remodelage de cette dernière.

d'interaction à une visualisation indirecte (sur l'écran).



« La plasticité d'une interface dénote sa capacité à s'adapter aux variations du contexte d'usage en termes d'utilisateur, de plate-forme et/ou d'environnement » ([Calvary 04], page 6). L'objectif de l'adaptation est de préserver l'utilité et l'utilisabilité de l'interface.

À l'heure actuelle, les réflexions sur la deuxième partie concernant la transformation de l'interface sont toujours en cours. Ainsi, cette partie fera donc l'objet d'un approfondissement dans de futurs travaux.

3.3 Conclusion

Nous avons vu dans ce chapitre que les théoriciens de l'apprentissage humain [Bruner 60] classifient les modes de représentation de savoir en trois catégories : symbolique, iconique et éactif. Hooper-Greenhill [Hooper Greenhill 94] insiste sur l'importance de combiner ces trois modes au musée afin de se mettre à la portée des différents visiteurs. C'est à la lumière de toutes ces indications, que nous rendons compte dans nos propositions puisant leurs principes de la réalité augmentée, des trois modes de représentation du savoir. En fait, grâce à la réalité augmentée, le mode éactif est représenté par un guide humain virtuel simulant la présence d'un guide humain et appuyé par le paradigme « Sélection ». Le mode symbolique est traduit par le paradigme d'interaction appelé « Documentation ». Le paradigme « Reconstruction » renvoie quant à lui au mode iconique. Nous allons donc voir au chapitre suivant l'implémentation technique de chacune de nos propositions.

Chapitre 4

Systeme M.A.R.T.S : Mobile Augmented Reality Touring System (Conception et Experimentation)

Introduction

Falk [Falk 05] [Falk 00] a qualifié l'apprentissage au musée comme étant une activité libre et totalement orientée par les choix personnels du visiteur. En ce sens, il va falloir que le visiteur puisse bénéficier des informations sur les expositions qui l'intéressent. D'un autre côté, Hooper-Greenhill [Hooper Greenhill 94] insiste sur l'importance de combiner les trois modes de transmission de savoir (vus au chapitre précédent), étant donné que chaque visiteur a une préférence pour l'un et/ou pour l'autre de ces modes. Un système d'aide à la visite doit donc mettre en œuvre différents outils interactifs performants pour combiner différents modes de communication. Grâce à cela, le visiteur pourra tirer profit d'une visite guidée assez riche.

C'est dans cette optique que nous avons conçu notre propre système d'aide nommé M.A.R.T.S, acronyme de « Mobile Augmented Reality Touring System ». Celui-ci est l'implémentation du principe général de fonctionnement d'un système d'assistance à la visite d'un musée que nous avons énoncé au chapitre précédent, tout en y intégrant nos propositions de visualisation et d'interaction.

Nous allons dans ce chapitre présenter le système M.A.R.T.S dans son fonctionnement interne tout en justifiant nos choix de conception et d'implémentation. Qui plus est, durant nos travaux, nous avons été en mesure d'expérimenter en partie nos propositions. Nous en présenterons les premiers résultats.

4.1 Le système M.A.R.T.S

Les systèmes de guide intègrent des fonctionnalités toujours plus riches, nombreuses et complexes. Pour que les utilisateurs puissent pleinement en profiter, leur niveau d'utilisabilité est un élément essentiel. L'utilisabilité correspond à « l'adaptation de la technologie aux caractéristiques de l'utilisateur » [Brangier 03] et se définit par la conjonction de :

- L'efficacité : produire l'effet attendu.
- L'efficience : découvrir les œuvres d'arts et acquérir de nouvelles connaissances avec un minimum d'efforts.
- La satisfaction : niveau de confort ressenti en utilisant le système de guide à la visite de musée.

C'est guidé par ces recommandations que nous avons conçu le système M.A.R.T.S. Nous allons tout d'abord rappeler son principe de fonctionnement avant de présenter son diagramme de conception général.

4.1.1 Rappel de fonctionnement

Pour proposer une solution permettant de créer une visite guidée ayant un but éducatif grâce à la réalité augmentée, nous avons retenu deux aspects fondamentaux :

Simulation de la présence d'un guide humain : les théoriciens de l'apprentissage au musée [Hooper Greenhill 94] insistent sur l'importance de l'apprentissage « face à face » faisant intervenir les guides humains. Étant donné que les recherches en réalité virtuelle et en réalité augmentée [Gerhard 05][Luo 14], ont prouvé la capacité des humains virtuels à simuler la coprésence humaine, nous faisons donc appel à ces derniers pour propulser l'apprentissage au musée.

Les modes d'interaction : ils présentent les paradigmes d'interaction qui permettent d'appuyer le discours du guide humain virtuel et d'améliorer la visualisation des informations promulguées par ce discours. Partant de ces aspects, nous avons alors proposé le fonctionnement d'un tel système (Figure4.1.1) :

L'utilisateur est équipé d'un smartphone lui permettant la visualisation d'un guide humain virtuel intégré à la scène réelle. Le flux vidéo est également augmenté en utilisant nos paradigmes d'interaction. En plus, le flux capturé par le smartphone de l'utilisateur est envoyé à un serveur distant pour être analysé. Le

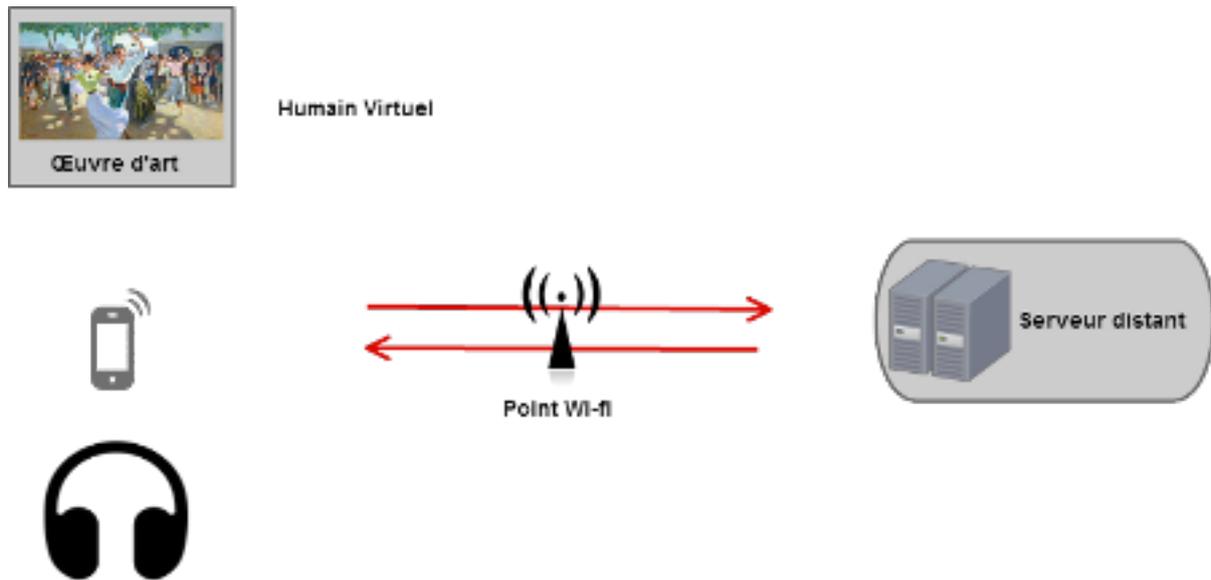


FIGURE 4.1.1 – Principe de Fonctionnement

serveur contient une base de données permettant de reconnaître l'objet capturé par le smartphone.

4.1.2 Diagramme de déploiement

Le diagramme de déploiement ¹ de la figure suivante représente une vue statique de l'infrastructure physique des ressources matérielles qui composent M.A.R.T.S. Ce diagramme montre également la répartition de composants de notre système ainsi que leurs relations.

Le système M.A.R.T.S se découpe en deux parties distinctes, l'une déployée sur le smartphone de l'utilisateur, l'autre doit être installée sur un serveur distant. L'utilisateur sera équipé des éléments suivants :

- **Un smartphone** : le smartphone doit intégrer essentiellement une caméra. Ce choix matériel revient au fait que les visiteurs puissent utiliser leurs propres équipements ² qu'ils savent déjà manipuler. Cela peut bien évidemment créer un confort d'utilisation chez le visiteur. De plus, le smartphone contient aussi l'application mobile qui supporte les paradigmes d'interac-

1. Nous précisons au lecteur que nous n'avons pas utilisé UML pour représenter ce diagramme

2. En se basant sur les chiffres rapporté par la magazine aquitaine numérique numéro 35, mai-juin 2011, 47 pour cent des touristes possèdent des smatphones.

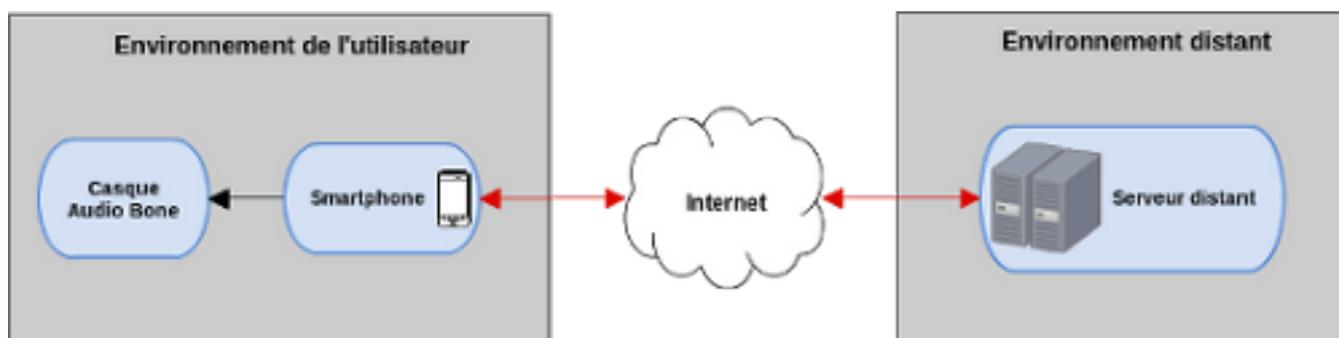


FIGURE 4.1.2 – Design général de M.A.R.T.S

tion, le suivi des points d'intérêts et qui gère la communication internet. Il est à noter que l'application mobile peut être installée avec l'aide du service d'accueil du musée.

- **Un casque Audio Bone**³ : il permet à l'utilisateur d'écouter le discours de l'humain virtuel. Il s'agit d'un casque qui fonctionne par conduction osseuse et donc qui permet de libérer le canal auditif de l'utilisateur. Ce matériel doit être prêté au visiteur par le service d'accueil du musée. L'objectif de l'utilisation d'un tel casque est de ne pas isoler l'utilisateur de son environnement ambiant. En plus, il s'agit encore de permettre à l'utilisateur d'interagir entre autres avec ses accompagnateurs et les autres visiteurs. Ceci permet de tenir compte du volet contexte socio-culturel du modèle contextuel de l'apprentissage au musée de Falk [Falk 05], détaillé au deuxième chapitre de cette thèse.

Sur le serveur distant, nous retrouvons :

- **Une base de données** : la base de données 2D regroupe les objets exposés et les œuvres d'art de musée.
- **Une application** : l'application qui sera installée sur le serveur de musée doit permettre l'analyse des objets 2D contenus dans la base de données. Ceci dans le but de reconnaître l'objet capturé par le smartphone de l'utilisateur. L'application permet aussi de gérer la connexion internet et le transfert de données.

Nous avons fait en sorte que le visiteur soit le moins instrumenté possible pour ne pas perdre de temps à s'équiper et pour ne pas être gêné pendant sa visite. Comme nous allons le voir dans la section suivante, le fonctionnement de M.A.R.T.S sera le plus transparent possible. L'interface utilisateur de M.A.R.T.S sera également

3. <http://www.audioboneheadphones.com/>

la plus simple possible, pour ne pas nécessiter une réelle phase d'apprentissage ou une grande connaissance de l'informatique. Ceci permettra à notre système de s'adapter aux différents types de visiteurs.

En ce qui concerne le service d'accueil de musée, sa mission sera de fournir des casques de type Audio Bone aux visiteurs et les aider à installer l'application sur leurs appareils mobiles. Nous précisons au lecteur que nous avons fait le choix de placer une base de données sur un serveur distant afin de tenir compte des faibles capacités de stockage de certains smartphones⁴.

4.1.3 Choix d'implémentation

En ce qui concerne l'implémentation du prototype expérimental du système M.A.R.T.S, nous avons fait le choix d'une programmation orientée composants. Un composant est un module cohérent d'un programme qui peut être développé séparément et distribué en tant qu'unité indépendante. Il offre une interface par laquelle il peut être connecté, tout en restant inchangé, avec d'autres composants pour former un système de taille plus importante [D'Souza 99]. Opter pour une programmation orientée composants nous est alors apparu comme une solution permettant une grande modularité dans le code, adaptée à la conception d'un prototype expérimental. Pour ce faire, nous avons porté notre choix de développement de composants en nous appuyant sur la bibliothèque Android SDK. Développée en java, elle offre des composants d'interfaces graphiques, d'accès aux données, de connexion réseau. Elle offre également un puissant moyen de communication entre les composants à travers le mécanisme appelé intent⁵.

Pour l'implémentation des algorithmes de traitement d'images, de suivi et de reconnaissance d'objets, nous avons opté pour un développement en C/C++. En fait, les langages C/C++ permettent une exécution rapide du code comparé à Java. Les langages C/C++ ont été utilisés en conjonction avec la bibliothèque de vision par ordinateur OpenCV⁶. Nous précisons au lecteur qu'OpenCV dispose d'une version pour Android et que les composantes de l'application écrites en java et en C/C++ peuvent communiquer grâce au JNI⁷.

4. Avec les avancées technologiques, l'on peut passer d'une architecture client-serveur à une application autonome « self-contained application ».

5. Intent : Les Intents permettent de gérer l'envoi et la réception de messages afin de faire coopérer les composants d'une application ou même des composants de deux applications différentes.

6. Bibliothèque initialement développée par INTEL. Disponible sous licence BSD à l'adresse : <http://opencv.willowgarage.com/wiki/>

7. JNI : est l'acronyme de Java Native Interface est une bibliothèque intégré au JDK qui

Pour ce qui est du traitement 3D, nous avons choisi la bibliothèque OpenGL, qui dispose d'une version pour mobile OpenGL ES et qui est facilement interfaçable avec la bibliothèque Android SDK.

Il est à noter qu'à ce stade de l'avancement de développement, le prototype expérimental du système M.A.R.T.S intègre le guide humain virtuel, les paradigmes de « sélection » et de « documentation ».

4.1.4 Vue d'ensemble de l'architecture interne

Comme nous venons de le voir, l'application de réalité augmentée devant être installée sur le smartphone du visiteur, fait appel à un programme distant permettant de reconnaître ce que voit ce dernier. Nous allons préalablement exposer une vue générale de l'architecture de l'application avant de décrire le rôle de chacun des composants la constituant. Certains d'entre eux, présentant une conception particulière, verront leurs fonctionnalités décrites à la section suivante.

4.1.4.1 Diagramme de relation

Le diagramme de relation permet de mettre en évidence les dépendances et les relations entre les différents éléments constituant l'application M.A.R.T.S. Basés composants, les modules de M.A.R.T.S sont architecturés comme suit (Figure 4.1.3) :

4.1.4.2 Description du rôle des composants

Du côté de l'application de RA installée sur le smartphone de l'utilisateur (ou application utilisateur), on trouve :

- **Interface utilisateur** : il s'agit du composant chargé d'afficher l'image augmentée sur l'écran du smartphone de l'utilisateur. Ce composant sera plus détaillé à la section suivante.
- **Module de RA** : c'est le module de réalité augmentée. Il est chargé de calculer les augmentations à appliquer sur l'objet exposé. Selon l'objet exposé, ce module doit décider du type de l'augmentation à mettre en œuvre : « sélection » et/ou « reconstruction ». La « documentation » et l'intégration du guide humain virtuel pouvant être en général associés aux différents objets

permet de programmer des appels natifs vers des composantes C/C++. Le développement avec JNI sous Android requière l'API NDK d'Android.

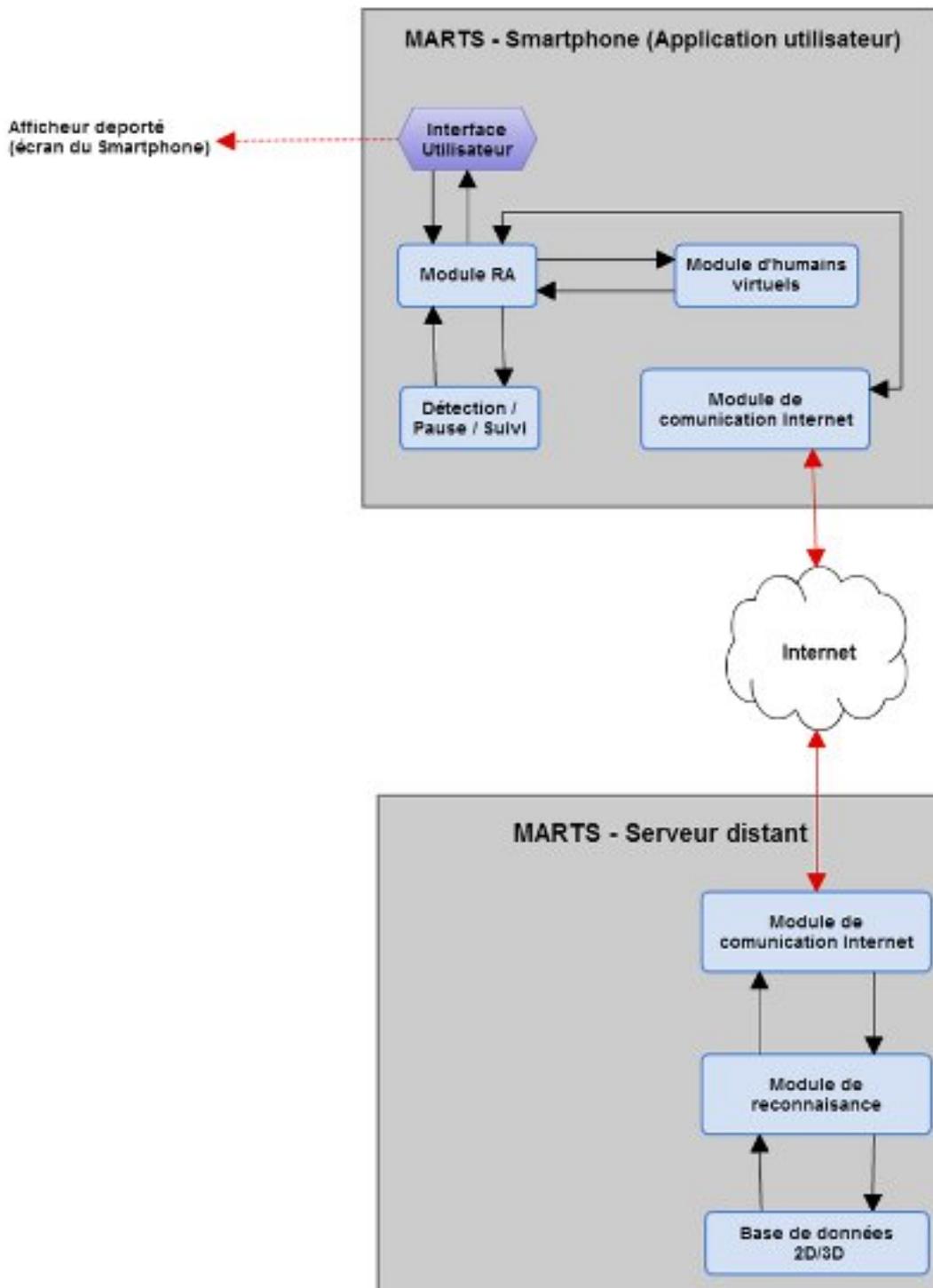


FIGURE 4.1.3 – Vue d'ensemble de l'architecture basée composants du système M.A.R.T.S.

exposés. Le flux vidéo étant ainsi augmenté, il est transmis vers le composant Interface utilisateur en vue d'être affiché. Les procédés algorithmiques permettant de mettre en œuvre les différentes augmentations, sont détaillées à la section suivante.

- **Module d'humains virtuels** : le module d'humains virtuels a été mis en place dans le cadre du projet Aquitaine-Euskadi (voir annexe). Il s'agit d'une bibliothèque d'humains virtuels pour plateforme mobile. Cette bibliothèque comporte un moteur d'animation et un moteur graphique 3D fondé sur Ogre3D⁸. Les mouvements et les gestes de l'humain virtuel sont gérés par l'intermédiaire d'un fichier BML⁹. Ce module se trouve détaillé à la section suivante.
- **Module de communication Internet** : ce module permet à l'application utilisateur de se connecter à l'application distante. Il permet au module de RA de recevoir l'identifiant de l'objet exposé, les images 2D concernant les zones d'intérêt dans l'objet, le script décrivant le discours et l'animation de l'humain virtuel. Il permet également de recevoir le modèle 3D de l'objet exposé dans le cas d'une augmentation de type « reconstruction ». D'un autre côté, ce module permet d'envoyer une image 2D de l'objet capturé par l'interface utilisateur vers le serveur distant. Le transfert de données est assuré par le protocole FTP¹⁰, implémenté par Android SDK.
- **Module de détection/Pose/Suivi** : La fonction de détection sert à retrouver l'objet exposé dans l'image en cours, ou comme dans le cas de « Sélection » à retrouver la(es) zone(s) d'intérêt dans ce dernier. Le calcul de la pose permet dans le cas de « Reconstruction » de calculer la position-orientation (pose) de l'objet exposé afin d'y projeter le modèle 3D. Le suivi, tout en respectant les mouvements de l'utilisateur, permet de recalculer l'augmentation et l'objet réel afin de créer l'illusion de coexistence entre les deux. L'implémentation de ces fonctionnalités se trouve expliquée à la section suivante.

Du côté de l'application mise en place sur le PC distant et appelé par l'application de RA, on trouve :

- **Module de communication internet** : ce module permet le transfert de données de et vers l'application utilisateur. Le fonctionnement de ce composant est identique à celui de l'application de RA. Se référer ci-dessus pour la description de son rôle.
- **Module de reconnaissance** : ce module permet d'identifier l'objet figurant

8. <http://www.ogre3d.org/>

9. <http://blinkenlights.net/project/bml>

10. FTP : File Transfer Protocol

dans l'image capturée par la caméra du smartphone. Il est chargé d'analyser la base de données afin de trouver l'objet en question. Il est implémenté en utilisant le langage C/C++ en conjonction avec la bibliothèque OpenCV. L'algorithme mis en œuvre pour la reconnaissance d'objet est détaillé dans la section suivante.

- **Bases de données 2D/3D** : la base de données regroupe d'un côté les images 2D qui vont servir au module précédent dans l'identification de l'objet exposé. D'un autre côté, la base de données regroupe les modèles 3D d'objets exposés qui vont permettre au module RA de mettre en place les augmentations de type « Reconstruction ».

4.2 M.A.R.T.S : implémentation du prototype expérimental

Nous allons dans cette partie du mémoire décrire l'implémentation de nos propositions théoriques pour soutenir l'apprentissage au musée, que nous avons détaillé au chapitre précédent. Comme nous venons de le voir à la section précédente, les augmentations proposées ne sont pas traitées de la même façon. Ceci est dû à la nature même de ces augmentations. Nous nous sommes donc basés sur des techniques adaptées de la vision par ordinateur afin de mettre en œuvre ces derniers. C'est ce que nous allons voir au niveau de cette section, puis nous terminerons en présentant l'interface graphique qui permettra à l'utilisateur de percevoir la scène augmentée.

4.2.1 Reconnaissance

La fonction reconnaissance se déclenche suite à une requête de l'application utilisateur. Le rôle de la fonction reconnaissance est d'identifier l'objet exposé ayant été capturé par la caméra de l'utilisateur. Pour ce faire, nous proposons une méthode de reconnaissance basée sur celle utilisée par Bay et al. [Bay 08] pour l'évaluation des performances du descripteur SURF. Cette méthode se scinde en trois étapes : la détection de points d'intérêt, l'appariement de points d'intérêt et la sélection de la bonne image de référence (ou image modèle). La dernière étape constitue la phase de reconnaissance proprement dite. Comme nous allons le voir dans cette section, l'apport de notre méthode réside au niveau de la deuxième étape.

Nous précisons au lecteur que Bay et al. [Bay 08] utilisent le nombre de correspondance de points d'intérêt comme un critère pour la reconnaissance d'objet. En ce sens, l'image de référence sélectionnée est celle ayant obtenu le nombre de

correspondances le plus élevé. Il est à noter que Bay et al. [Bay 08] emploient une approche linéaire de recherche du plus proche voisin pour l'appariement de points d'intérêt SURF. En revanche, la recherche linéaire du plus proche voisin est très coûteuse pour les applications temps réel. Par conséquent, plusieurs méthodes se sont intéressées à la recherche approximative du plus proche voisin. Pour ce faire, nous nous sommes appuyés sur la bibliothèque FLANN¹¹ de Muja et al. [Muja 09]. Néanmoins, l'inconvénient de FLANN est la génération de faux positifs (outliers). Les expérimentations que nous avons réalisées avec FLANN montrent que le nombre de correspondances est non significatif et donc il ne peut pas être utilisé comme critère de sélection de l'image de référence. Pour pallier à ce problème, nous proposons donc une méthode de filtrage de correspondances que nous allons présenter plus loin dans cette section. Les tests d'évaluation effectués avec notre méthode montrent la robustesse de cette dernière.

4.2.1.1 Algorithme

Détection des points d'intérêt

Les points caractéristiques sont des points intéressants contenus dans une image. Bien évidemment, ils présentent des propriétés locales importantes au niveau de l'image. En plus, ils sont généralement insensibles aux changements locaux et globaux, tels que l'éclairage, la luminosité et les transformations affines.

En 1988 Harris et al. [Harris 88], introduisent le détecteur de Harris, permettant d'extraire les points d'intérêt locaux. Plus précisément, le détecteur d'Harris permet d'extraire les coins, définis par l'intersection des deux arêtes. Le détecteur d'Harris a l'avantage d'être invariant aux transformations affines. Cependant, son inconvénient majeur réside dans le fait qu'il n'est pas invariant au changement d'échelle. Le détecteur SIFT, introduit par Lowe et al. en 1999 [Lowe 99], est un détecteur invariant au changement d'échelle. Le descripteur SIFT calcule un histogramme de gradients situés dans une fenêtre qui entoure le point clé. Ensuite, il les enregistre dans un vecteur de dimensions 128 ou 64, représentant un descripteur relatif à un point d'intérêt SIFT.

Le détecteur SURF, proposé par Bay et al. [Bay 08], est dérivé de SIFT. En revanche, il se distingue de SIFT par la description et la sélection des points clés. SURF se base sur le déterminant de la matrice Hessienne dans l'extraction des points. En effet, la matrice Hessienne se caractérise par une bonne performance en termes du temps de calcul et de précision. Plus précisément, SURF s'appuie sur

11. FLANN est acronyme de « Fast Approximate Nearest Neighbor Search Library » et est intégré dans OpenCV.

le déterminant de la matrice Hessienne dans le calcul de position et d'échelle d'un point caractéristique. Etant donné un point $X = (x, y)$ appartenant à l'image I , la matrice Hessienne $H(x, \sigma)$ en x à l'échelle σ est définie comme suit :

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (4.2.1)$$

$L_{xx}(x, \sigma)$ représente la convolution de la dérivée seconde de gaussienne $\frac{\delta^2}{\delta x^2} g(\delta)$ avec l'image I , au niveau du point X , et de même pour $L_{xy}(x, \sigma)$ et $L_{yy}(x, \sigma)$. L'extraction des descripteurs SURF s'effectue en deux étapes. La première étape consiste à trouver l'orientation de la région circulaire autour du point d'intérêt. Ensuite, une zone carrée, alignée à l'orientation précédemment calculée, est construite afin d'y extraire le descripteur SURF.

Le détecteur SURF est plus rapide que les autres détecteurs de points caractéristiques, grâce à l'utilisation des images intégrales. Une image intégrale peut être rapidement calculée à partir d'une image initiale en utilisant l'équation suivante :

$$I_{\Sigma}(X) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (4.2.2)$$

En effet, l'image intégrale $I(X)$ en un point $X(x, y)$ est la somme de tous les pixels de l'image initiale I , situés entre ce point et l'origine. Elle permet d'accélérer le calcul des valeurs d'intensité des pixels.

En ce qui nous concerne, nous avons retenu le détecteur SURF. En effet, ce descripteur est principalement reconnu pour sa rapidité de calcul. L'étude comparative de Juan et al. [Juan 09], démontre la supériorité du descripteur SURF par rapport à SIFT et PCA-SIFT d'un point de vue de ces performances en temps d'exécution et de sa robustesse aux changements d'illumination.

Appariement des points d'intérêt

La reconnaissance d'objet en vision par ordinateur, s'appuie essentiellement sur la recherche des correspondances entre deux images d'un même objet ou d'une même scène. Une fois que les caractéristiques visuelles sont extraites d'une image, elles sont appariées avec un ensemble de caractéristiques extraites d'une autre image. Comme nous venons de le voir, les descripteurs des points caractéristiques sont des vecteurs contenant des nombres réels. Bien évidemment, la manière la plus simple de comparer deux points caractéristiques, revient à calculer la distance euclidienne (ou le carré de la distance euclidienne) entre leurs descripteurs asso-

ciées. Etant donné deux vecteurs p et q , leur distance euclidienne est calculée de la manière suivante :

$$dist(p, q) = \sqrt{\sum_{i=1}^{64} (p_i - q_i)^2} \quad (4.2.3)$$

Il est clair que calcul de distance euclidienne serait très lent, dans le cas où la dimension de la base de données d'images est importante. Dans ce cas, il est préférable de choisir les descripteurs SURF de petite taille (vecteurs de dimension 64) au lieu des larges descripteurs SURF (vecteurs de dimension 128).

La recherche linéaire du plus proche voisin est coûteuse pour les applications temps réel. Par conséquent, plusieurs méthodes se sont intéressées à la recherche approximative du plus proche voisin. En fait, c'est ce dernier type de recherche que nous avons utilisé pour le calcul de correspondances de points SURF. Dans cette optique, nous nous sommes appuyés sur le travail de Muja et al. [Muja 09], intitulé «Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration». Muja et al. [Muja 09] proposent une bibliothèque d'algorithmes appelée FLANN (Fast Library for Approximate Nearest Neighbors), que nous avons employée dans notre publication intitulée « Mobile Augmented Reality Applications to Discover New Environments » [Ghouaïel 13b]. En effet, FLANN contient une collection d'algorithmes permettant de résoudre le problème de recherche approximative de plus proche voisin. Cette collection, fait appel, entre autres, à l'algorithme hiérarchique des K-moyennes (hierarchical k-means tree) [Arai 07] et à l'algorithme des arbres-KD randomisés multiples (multiple randomized kd-trees) [Duch 98]. FLANN permet de sélectionner l'algorithme adéquat, en fonction de l'ensemble des données et en fonction d'autres facteurs tels que le temps de calcul et la précision.

Comme le met en évidence les tests d'évaluation présentés à la section suivante, FLANN génère énormément des faux positifs « outliers ». Par conséquent, le nombre de correspondances n'est pas significatif et ne permet pas de sélectionner la bonne image modèle. En ce sens, nous avons trouvé que dans la majorité des cas, le nombre de correspondances qui est le plus élevé n'appartient pas à cette dernière.

Nous proposons donc d'effectuer une étape de filtrage suite à l'exécution de l'algorithme de recherche approximative du plus proche voisin. Le filtrage consiste à sélectionner les bonnes correspondances. Dans cette optique, deux points d'intérêts sont considérés comme appariés si la distance euclidienne entre leurs descripteurs est strictement inférieure à un seuil donné. Le seuil que nous employons est la médiane des distances euclidiennes minimales, obtenues suite à une première exécution de FLANN. Ainsi, l'appariement de points d'intérêt mettant en œuvre notre

principe de filtrage se résume en ces étapes :

Étape initiale

1. Appliquer l'algorithme de recherche approximative du plus proche voisin sélectionné par la bibliothèque FLANN.

Étape de filtrage

1. Pour chaque image modèle, enregistrer la distance euclidienne minimale obtenue.
2. Calculer la médiane des distances minimales.
3. Appliquer une deuxième fois l'algorithme de recherche approximative de plus proche voisin. Cette étape ne concerne que les images modèles ayant une distance euclidienne minimale inférieure à la médiane, précédemment calculée.
4. Pour chaque image modèle, enregistrer le nombre de correspondances de points SURF.

Sélection de l'image de référence

Nous supposons que suite à l'étape de filtrage, le nombre de correspondances devient significatif. L'image de référence liée à l'objet recherché se rapporte donc au nombre maximal de correspondances.

4.2.1.2 Évaluation et Résultats

Afin d'estimer les performances de notre méthode de filtrage, nous avons mené une étude expérimentale. Cette étude consiste à trouver dix images tests dans une base constitué de 50 images modèles en se basant sur l'algorithme cité ci-dessus. Nous avons donc comparé notre méthode de filtrage à l'utilisation de FLANN seule pour l'appariement de points d'intérêt. La figure 4.2.1 montre les résultats que nous avons obtenus pour les deux méthodes :

En se basant sur les résultats schématisés dans ce tableau, nous pouvons donc conclure que notre méthode de filtrage pour l'appariement des points d'intérêt permet de déterminer l'objet recherché dans 60% des cas. Cependant, FLANN utilisée toute seule ne permet en aucun cas de déterminer l'objet apparaissant dans l'image en cours. De plus, les résultats montrent que notre méthode permet de réduire de façon évidente le nombre de faux positifs (outliers). Afin d'améliorer la précision de notre algorithme de reconnaissance présenté ci-dessus, nous allons en plus de la méthode de filtrage intégrer une phase de détection d'objet tel que

Image Test	FLANN		Méthode de Filtrage	
	Image modèle trouvée ?	Nombre maximal de correspondances	Image modèle trouvée ?	Nombre maximal de correspondances
1	non	3552	oui	154
2	non	3401	oui	138
3	non	2234	oui	36
4	non	1104	non	251
5	non	1083	oui	48
6	non	1803	non	164
7	non	2456	oui	25
8	non	2507	oui	117
9	non	647	non	143
10	non	1216	non	106

FIGURE 4.2.1 – Comparaison avec FLANN

décrite par Augereau et al. [Augereau 13]. Cette amélioration fera donc l'objet d'un détail dans des travaux futurs.

4.2.2 Le suivi

Dans le cas d'une augmentation de type « Documentation », un recalage parfait entre l'objet exposé et le texte virtuel n'est pas absolument nécessaire. En effet, il doit être évident pour l'utilisateur que le texte virtuel est associé à l'objet en question. Cependant, dans le cas d'une augmentation de type « Reconstruction » ou « Sélection », toute déviation du virtuel de la zone réelle ciblée (ou augmentée), peut engendrer des problèmes de perception chez l'utilisateur. Par exemple, dans le cas de « Reconstruction » un recalage robuste est exigé afin de créer l'illusion de coexistence du virtuel et du réel chez le visiteur. Pour le cas de la « Sélection », toute déviation du quadrilatère de la zone d'intérêt ciblé, peut amener une ambiguïté de désignation. En ce sens, afin que les augmentations apparaissant devant le regard du visiteur restent recalées en accord avec les mouvements ou déplacements qu'il serait susceptible d'effectuer, nous devons faire appel à une méthode suivie.

Comme nous avons pu le constater au premier chapitre de ce mémoire, la littérature en vision par ordinateur est riche de techniques adaptées au suivi. En

ce qui nous concerne, nous avons choisi l'algorithme PTAM¹² [Klein 07], basé sur des points d'intérêt. Notre choix revient essentiellement au fait que cet algorithme assure un recalage temps réel robuste sur une plateforme mobile.

4.2.3 Guide humain virtuel

Présentation

La bibliothèque mobile d'agent conversationnel 3D a été mise en place dans le cadre du projet Aquitaine-Euskadi (voir annexe B). Le moteur graphique 3D (avatar engine) relatif à cette bibliothèque, a été construit au dessus du moteur graphique Ogre3D [Ogre 00]. En effet, notre bibliothèque a été conçue pour être facilement intégré à n'importe quel SDK de réalité augmentée mobile, indépendamment de son fonctionnement interne. La figure 4.2.2 montre l'architecture générale de la bibliothèque mobile d'agent conversationnel.

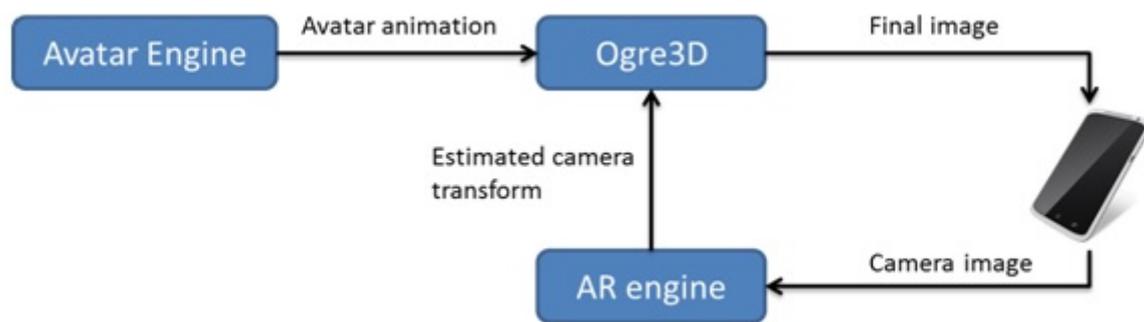


FIGURE 4.2.2 – Flux de données

Le flux de données schématisé par la figure 10, montre que la seule information requise par le moteur graphique 3D est la matrice de transformation monde-caméra. Cette matrice est utilisée afin de projeter l'avatar dans la bonne pose (position et orientation), créant ainsi l'illusion de la coexistence entre l'humain virtuel et le monde réel. Pour ce faire, nous avons implémenté une version mobile de l'algorithme de suivi PTAM [Klein 07]. Il s'agit d'un algorithme de suivi, basé sur les points caractéristiques (features points). PTAM est largement détaillé dans le premier chapitre de ce mémoire.

12. PTAM est largement détaillé au premier chapitre de ce mémoire.

Structure de la bibliothèque mobile d'agent conversationnel

Le moteur d'animation 3D (animation engine) est le module de la bibliothèque qui permet de contrôler les mouvements de l'humain virtuel. En effet, nous avons adopté le moteur d'animation proposé par Vicomtech¹³, notre partenaire du projet Aquitaine-Euskadi. Le moteur d'animation est chargé de générer l'animation de l'humain virtuel. Ce module est développé en se basant sur Open Scene Graph [openscenegraph 98]. La structure générale du moteur d'animation est représentée à la figure 4.2.3 :

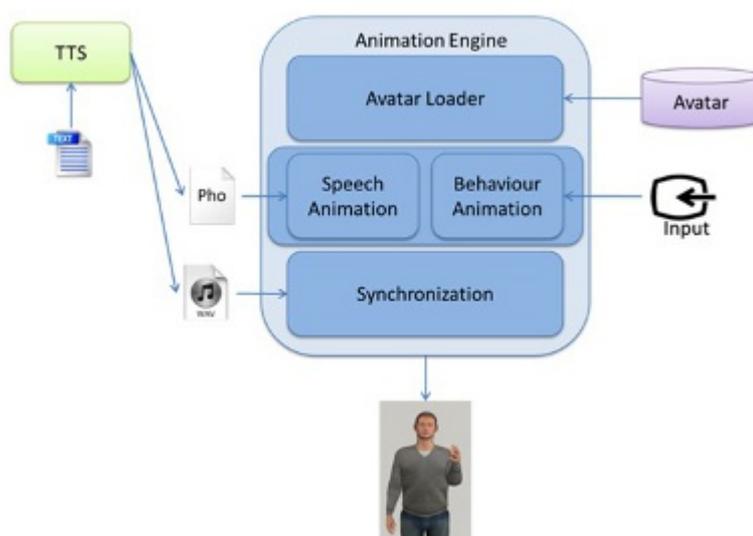


FIGURE 4.2.3 – Structure du moteur d'avatar

Comme le met en évidence la figure 11, l'animation de l'avatar est contrôlée par deux principaux modules : le module vocal (Speech Animation) et le module comportemental (Animation Behaviour).

D'une part, le module vocal gère les expressions faciales en se basant sur le texte à prononcer. En effet, il exécute l'animation faciale quand l'avatar parle. Cette dernière est générée en utilisant les techniques de morphing. Plusieurs expressions sont adoptées, nous citons par exemple : face neutre, clin d'œil, yeux fermés. En outre, l'animation faciale permet de traduire les émotions d'Ekman¹⁴. Elle est donnée par un fichier *.pho. Notons aussi que la voix de l'avatar est générée par un moteur de

13. Vicomtech : centre de recherche appliquée pour l'informatique graphique interactive et multimédia : www.vicomtech.org

14. Les principales émotions d'Ekman sont : la joie, la tristesse, la colère, la peur, le dégoût, le mépris et la surprise

synthèse vocale TTS.

D'autre part, le module comportemental permet de gérer les différents mouvements et gestes de l'humain virtuel. Il se base sur des règles prédéfinies pour le comportement naturel afin de générer le comportement corporel de l'avatar. Les éléments du comportement corporel (mouvements et gestes) peuvent être présentés au moteur d'animation par l'intermédiaire d'un fichier BML.

La sortie finale du moteur d'animation est gérée par le module de synchronisation. Ce module reçoit et combine les animations calculées par les différents modules. En fait, il est directement relié au moteur graphique 3D, et donc responsable du rendu dynamique de l'avatar.



Nous invitons le lecteur, souhaitant avoir un complément d'information sur le moteur d'animation de notre bibliothèque mobile d'agent conversationnel 3D, de consulter l'article de Carretero et al. [del Puy Carretero 12].

4.2.4 Paradigmes d'interaction

Nous rappelons au lecteur que les paradigmes d'interaction « Documentation », « Sélection » et « Reconstruction » servent à appuyer le discours de l'humain virtuel. L'implémentation de ces derniers se base essentiellement sur des techniques appropriées de la vision par ordinateur.

4.2.4.1 Sélection

Comme expliqué de manière extensive au chapitre précédent, la « Sélection » est un acte de désignation qui cible une zone d'intérêt sur un objet exposé. Pour implémenter cette augmentation, nous nous sommes basés sur la méthode SURF [Bay 08] d'extraction de points d'intérêt en conjonction avec le modèle (image) 2D de la zone d'intérêt. Il est à noter que SURF a été choisi principalement pour sa rapidité reconnue en comparaison avec SIFT [Juan 09]. Pour la détection de la zone d'intérêt, nous proposons la méthode suivante inspirée du travail de Augereau et al. [Augereau 13] :

1. Les points d'intérêts sont extraits de l'image en cours et du modèle en utilisant la méthode SURF.

2. Les points d'intérêts sont appariés en utilisant la méthode FLANN¹⁵.
3. Si suffisamment¹⁶ de points d'intérêts ont été appariés avec succès entre les deux images, la zone d'intérêt recherchée est considérée comme trouvée. Nous pouvons donc passer à la détection puis la sélection (cadrage) de cette zone sur l'objet.
4. Ensuite, on procède au calcul de l'homographie [Agarwal 05] afin de déterminer la transformation affine entre le modèle et l'image en cours. L'homographie utilise les correspondances de points calculés auparavant. Il est à noter qu'au niveau de cette étape, nous faisons appel à l'algorithme RANSAC [Fischler 81] afin d'éliminer les points aberrants.
5. Les quatre coins de la zone d'intérêt sont calculés en appliquant la transformation affine (homographie) trouvée à l'étape précédente, sur les extrémités du modèle traduites par les coordonnées images suivantes¹⁷ : $(0, 0)$; $(0, \text{height})$; $(\text{width}, \text{height})$; $(\text{width}, 0)$.
6. Nous faisons appel aux fonctionnalités de la bibliothèque OpenGL afin de dessiner le quadrilatère reliant les quatre coins et délimitant la zone d'intérêt.

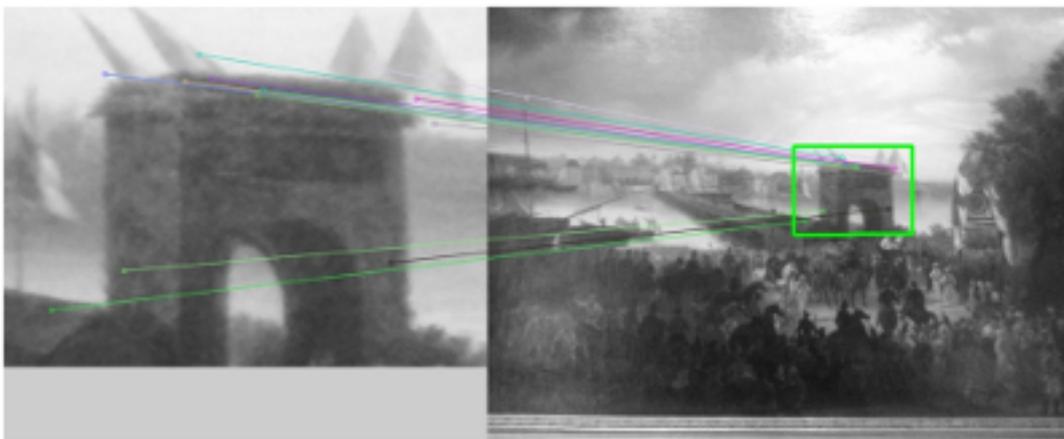


FIGURE 4.2.4 – Mise en œuvre de la « Sélection »

15. À ce niveau, nous nous contentons d'utiliser FLANN sans faire appel à notre méthode de filtrage expliquée plus haut car un seul modèle est utilisé.

16. Le nombre minimal de points est fixé à 4; il s'agit du nombre requis par l'homographie.

17. Height et Width désignent respectivement la hauteur et la largeur du modèle.

4.2.4.2 Reconstruction

Nous rappelons au lecteur que le but de ce paradigme d'interaction est de montrer au visiteur l'objet exposé dans son état original en utilisant le modèle CAO de ce dernier. Dans le cadre des travaux de cette thèse, nous n'avons pas pu achever l'implémentation du paradigme « Reconstruction ». En effet, l'implémentation du principe de cette augmentation nécessite le modèle CAO de l'objet à reconstruire en réalité augmentée. Ne disposant pas encore de tels modèles, nous avons commencé à réfléchir aux grandes étapes permettant la mise en œuvre de la « Reconstruction ». Pour ce faire, nous nous sommes inspirés de la méthode de Bottecchia [Bottecchia 10] :

1. À la phase d'initialisation, nous utilisons la méthode manuelle de Platonov [Platonov 06] afin de construire un modèle 3D de points d'intérêt. Ce modèle nous permettra d'obtenir une correspondance 2D-3D de ces points d'intérêt.
2. Au cours de l'utilisation de l'application, la correspondance 2D-3D de points d'intérêt calculée auparavant servira à retrouver l'objet dans l'image en cours.
3. Nous calculons la pose (Position-Orientation) de l'objet exposé, localisé dans l'image à l'étape précédente, en utilisant l'algorithme POSIT [DeMenthon 95] détaillé au premier chapitre de ce mémoire. Ensuite, nous nous appuyons sur les fonctionnalités d'OpenGL afin de projeter le modèle CAO suivant la pose ici obtenue.
4. Afin d'assurer un suivi en temps réel robuste de l'objet 3D projeté sur l'objet réel, nous faisons appel à la méthode de suivi décrite par l'algorithme PTAM [Klein 07]. Étant donné que le développement de cette partie est en cours, la description détaillée de notre méthode figurera donc dans de futurs travaux.

4.2.4.3 Documentation

Le paradigme d'interaction « Documentation » ne nécessite pas un traitement complexe comme les autres paradigmes d'interaction. L'affichage du texte virtuel se fait grâce à la bibliothèque OpenGL.

4.2.4.4 Interface utilisateur

Nous avons veillé à ce que l'interface utilisateur de M.A.R.T.S soit la plus simple possible, ne nécessitant pas de phase d'apprentissage ou une grande connaissance de l'informatique. Ceci permettra à notre système de s'adapter aux différents types de visiteurs.

Nous précisons au lecteur que l'interface utilisateur de M.A.R.T.S a été développée en utilisant Android SDK fournissant tous les widgets nécessaires à son élaboration. Pour respecter un principe d'utilisation simple, tournée vers l'utilisateur final au sens de [Cooper 95], celle-ci assure une transparence de nos paradigmes technologiques d'interaction. Dans cette optique, notre interface contient deux boutons : « Découvrir » et « Arrêter ». Le premier bouton permet d'intégrer un guide humain virtuel à la scène réelle. Il permet également de déclencher tous les types de paradigmes appuyant le discours de ce dernier. Le deuxième, quant à lui, sert à arrêter l'animation à tout moment.

Par ailleurs, l'interface de M.A.R.T.S n'impose pas au visiteur un circuit particulier au sein du musée, ni les œuvres d'art à regarder. En fait, comme nous l'avons vu au deuxième chapitre de cette thèse, tenir compte des propres choix du visiteur et de son intérêt pour tel ou tel œuvre d'art s'inscrit sous le cadre du volet contexte personnel du modèle de Falk [Falk 05]. En plus, la présentation d'un objet exposé se fait suite à la demande du visiteur en cliquant sur le bouton « Découvrir » et s'arrête aussi à la demande de celui-ci en cliquant sur le bouton « Arrêter ». Ceci est dans le but de respecter le principe de l'apprentissage libre au musée « free-choice learning » de Falk [Falk 05].

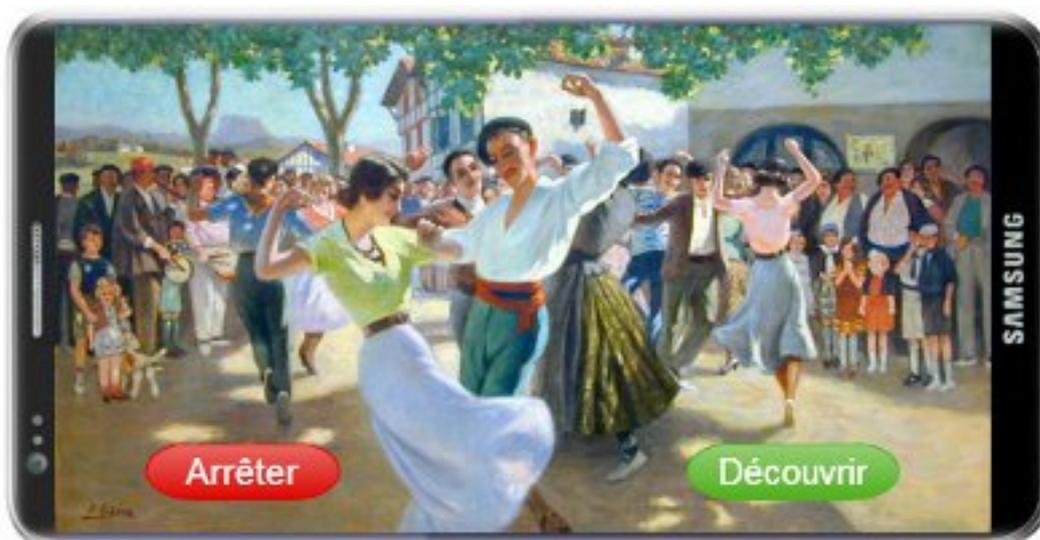


FIGURE 4.2.5 – Interface utilisateur de M.A.R.T.S

4.2.4.5 Adaptations de l'interaction à l'environnement de l'utilisateur

Nous rappelons au lecteur que dans le cas de « Documentation », si le guide humain virtuel ne peut pas se référer au titre de l'étiquette (quand l'utilisateur ne peut pas trop s'approcher de cette dernière, par exemple au moment de forte affluence) pour désigner le nom de l'œuvre d'art et son auteur, on doit afficher un texte virtuel se rapportant à ces informations. Dans ce cas, nous nous appuyons sur la distance¹⁸ de l'utilisateur par rapport à l'objet exposé.

En vision par ordinateur, la technique de stéréovision est largement utilisée [Brown 03] [Hager 95][Kytö 11] afin de calculer ce genre de distance. Cette technique se base sur la géométrie épipolaire [Zhang 96] et requière deux caméras avec un déplacement horizontal connue à l'avance. Ceci constitue l'inconvénient majeur de cette méthode, étant donné que les smartphones en possession des visiteurs ne sont pas, à l'heure actuelle, équipés de caméra stéréoscopique. Dans cette optique, nous avons trouvé la méthode décrite par [Holzmann 12] très intéressante. En effet, leur méthode se base sur la technique de parallaxe de mouvement¹⁹ et utilise une caméra monoculaire pour créer une vision stéréoscopique. En outre, elle a l'avantage de fonctionner sous une plateforme mobile. Cette méthode suppose que l'utilisateur effectue un déplacement latéral de son smartphone pour créer deux vues différentes de l'objet exposé. Étant donné que le déplacement n'est pas connu à l'avance, Holzmann et al. [Holzmann 12] utilisent le système inertiel afin de calculer le déplacement de la caméra. Leur méthode semi-automatique peut se résumer en ces étapes :

1. L'utilisateur effectue un mouvement latéral de son smartphone qui est enregistré en utilisant le gyroscope et l'accéléromètre intégrés à ce dernier. Ceci permet de créer deux images de deux points vues différents du même objet. La distance entre la position initiale et la position finale de la caméra s'appelle « baseline » et elle est notée b ²⁰. Cette dernière est calculée en se basant sur l'accélération et la vitesse angulaire de la caméra provenant respectivement de l'accéléromètre et du gyroscope, comme décrit dans [Seifert 07]. Par la suite, depuis son smartphone, l'utilisateur peut sélectionner l'objet en question.
2. Les points d'intérêt SURF sont extraits (une région de 70×70 pixels) de

18. Calculer la distance de l'utilisateur par rapport à l'objet exposé revient à calculer la distance de la caméra à l'objet exposé.

19. Indice monoculaire de profondeur qui se réfère à un déplacement apparent de la position d'un objet vu de différents points de la vue.

20. La ligne de base « baseline » dénote plus précisément la distance entre la position initiale et finale du centre optique de la caméra.

l'objet sélectionné et sont appariés. Ensuite, on calcule la disparité D qui se rapporte au déplacement horizontal des points d'intérêt appariés. La distance Z séparant la caméra de l'objet se mesure en utilisant la formule de triangulation suivante, avec f qui dénote la distance focale de la caméra :

$$Z = \frac{f \cdot b}{D}$$

Nous précisons au lecteur que la distance limite, au-delà de laquelle le titre de l'étiquette n'est plus lisible pour l'utilisateur, doit être définie empiriquement, car elle varie en fonction du contexte physique de chaque musée.

Comme nous l'avons vu au chapitre précédent, quand l'objet exposé se trouve occulté pour des raisons de fortes affluences par exemple, nous mettons en œuvre une transformation de l'interface homme machine de MARTS. Cette transformation consiste à basculer l'affichage des paradigmes d'interaction d'un mode direct de la réalité augmentée (affichage sur la vidéo de l'objet exposé) en un mode indirect (affichage sur une image de l'objet sur l'écran du smartphone). Il est à noter que cette adaptation de l'interface homme machine dépasse le spectre de cette thèse et fera l'objet de futurs travaux.

4.3 Évaluation expérimentale

Dans le cadre de ces travaux de thèse, nous avons pu développer une partie du système M.A.R.T.S, notamment le paradigme de « Sélection », le paradigme de « Documentation » et le module d'humains virtuels (projet Aquitaine-Euskadi en annexe B). L'interface logicielle utilisateur permet donc d'intégrer ces paradigmes d'interaction. Pour tester en partie la pertinence de nos propositions, nous avons donc expérimenté le principe de la simulation de la présence d'un guide humain par le biais des humains virtuels. En plus, nous avons également expérimenté le principe de références ostensives (paradigme de « Sélection ») représentant une stratégie importante de la communication, ainsi que le paradigme de « Documentation ».

4.3.1 Ce que l'on cherche à évaluer

Dans le cadre d'une visite guidée au musée, nous souhaitons comparer différents moyens de communication entre eux. Pour ce faire, nous procurons au visiteur découvrant le musée : les étiquettes (textes descriptifs), un audio-guide et notre système M.A.R.T.S. Les audio-guides étant un moyen largement utilisé par les musées, nous avons voulu donc tester la pertinence d'une communication audio seule

face à un système comme M.A.R.T.S. De plus, nous avons voulu tester l'intérêt de notre système M.A.R.T.S vis-à-vis des étiquettes seules, mode symbolique de la communication.

De façon plus précise, nous nous sommes principalement intéressés dans notre étude à répondre à la question suivante : « **Est-ce que les différents moyens de communication permettent au visiteur d'acquérir le même niveau de connaissances ?** »

En effet, quantifier les connaissances acquises suite à la visite au musée représente un aspect fondamental pour une application à vocation éducative. Ceci est mis en avant dans les travaux de Falk et al. [Falk 05]. En fin de test, nous avons également questionné les utilisateurs sur leurs impressions en vue d'établir un classement des systèmes.

4.3.2 Description de l'évaluation

4.3.2.1 Descriptions des objets exposés et des moyens de communication

L'étude expérimentale a eu lieu au musée basque et de l'histoire de Bayonne²¹. Pour notre évaluation, nous avons utilisé six objets exposés. Pour chaque objet exposé, les descriptions ont été établies avec l'aide des conservateurs et médiateurs de musée basque²². Les descriptions fournies pour chaque objet ont les mêmes longueurs.

En ce qui concerne les moyens de communication utilisés pour les tests, nous avons donc voulu comparer la pertinence de M.A.R.T.S face aux deux moyens le plus largement utilisés qui sont l'audio-guide et l'étiquette. Nous avons donc testé trois configurations :

- Étiquette : un texte descriptif de l'œuvre d'art.
- Audio-guide : de la réalité augmentée en mode audio, l'utilisateur est équipé d'un casque audio lui permettant d'écouter des enregistrements vocaux générés automatiquement par un module TTS²³.
- M.A.R.T.S : c'est l'utilisation du système M.A.R.T.S avec les casques Audio Bone et avec un Smartphone de type Galaxy SII.

21. <http://www.musee-basque.com/>

22. Notamment de madame Maider Etchepare Jaureguy, responsable du service des publics et du développement culturel de musée basque et de l'histoire de Bayonne.

23. Text To Speech

4.3.2.2 Sujets et plan de recherche

Sujets

Dix sept visiteurs ont participé à l'étude; 9 hommes et 8 femmes avec une moyenne d'âge de 40 ans. Ils n'ont aucun lien de parenté entre eux. De plus, ils ont visité le musée basque pour la première fois, étant tous fortement intéressés par la culture basque et l'histoire de la ville de Bayonne. Ceci fait partie du principe de l'homogénéité de l'échantillon recommandé par Falk et al. [Falk 05]²⁴.

Plan de recherche

Chaque participant a expérimenté les trois moyens de communication (Étiquette, Audio-guide, M.A.R.T.S), l'ordre d'utilisation étant aléatoirement décidé. Pour chacun des trois modes, deux objets exposés, tirés aléatoirement, sont présentés à l'utilisateur parmi les six proposées. Les différents tirages aléatoires ont pour finalité de neutraliser (limiter) les variables parasites comme par exemple ici le phénomène d'accoutumance (ou d'habituation) induit par l'apprentissage.

Les participants n'ont pas été préalablement familiarisés aux nouveaux modes de communication. De par leur usages courants (porter des casques audio, utiliser le Smart- phone . . .) aucun apprentissage n'a donc été nécessaire.

L'unique consigne que les participants aient reçue était celle d'utiliser le moyen de médiation approprié au niveau de chaque objet exposé. Aucune limite de temps n'a été imposée. Pour finir, nous leur avons demandé de remplir un questionnaire jugeant différents critères en vue d'établir un classement.

4.4 Analyse et interprétations

Dans ce qui suit, nous allons présenter les résultats en deux parties. En premier lieu, nous allons examiner le nombre des réponses correctes en fonction des systèmes utilisés. Par la suite, nous allons examiner le résultat du questionnaire d'évaluation. Les analyses statistiques ont été établies à l'aide du logiciel SPSS (Statistical Package for the Social Sciences) de la société IBM. Pour finir, nous discuterons les résultats de ces deux parties.

24. Selon Falk, la connaissance a priori et l'intérêt personnel influencent les acquis des visiteurs suite à la visite de musée. Ce dernier insiste sur le fait de choisir un échantillon homogène vis-à-vis ces deux critères afin de garantir la cohérence des résultats.

4.4.1 Méthode d'analyse des résultats

4.4.1.1 Nombre de réponses correctes

Nous avons étudié l'influence de 3 facteurs (Étiquette, Audio-guide, M.A.R.T.S) sur la variable dépendante²⁵ quantitative qu'est le nombre de réponses correctes concernant les objets exposés. Pour cela, nous avons donc soumis les sujets à chacun des niveaux de la variable indépendante²⁶ (moyen de communication) conformément au plan de recherche précédemment décrit. Nous sommes donc dans un plan de test à mesures répétées.

Afin d'inférer une relation entre les moyens de médiation et le nombre de réponses justes, nous avons utilisé une analyse de la variance (ANOVA) à mesures répétées en ayant choisi un niveau de risque *alpha*²⁷ de 0.05. Notre variable indépendante qualitative X est donc le moyen de communication (à 3 niveaux de valeur), la variable dépendante quantitative Y étant le nombre de réponses justes. Il est à noter que Y varie de 0 à 2, étant donné que pour chaque objet exposé une seule question est posée et que chaque moyen de médiation est testé avec deux objets exposés.

Notre hypothèse statistique H_0 (hypothèse nulle²⁸) est : « **Le nombre de réponse justes est égal pour tous les modes de médiation (rÉtiquette = rAudio-guide = rM.A.R.T.S)** ». L'hypothèse bilatérale²⁹ H_1 correspondant à notre objectif de recherche est donc : « **Est-il vrai de dire que le nombre de réponses justes varie en fonction des modes de médiation utilisés ?** ».

4.4.1.2 Les données du questionnaire

À la fin des tests, chaque visiteur s'est vu remettre un questionnaire en 9 points sur la façon dont a été ressentie l'expérience. Pour chaque question, nous leur avons demandé d'évaluer les trois modes de médiation sur une échelle ordinale de 0 (faible) à 14 (fort).

25. La réponse des sujets aux différentes conditions expérimentales.

26. Facteur représentant les causes postulées des variations de la variable dépendante.

27. Alpha est le risque numérique (probabilité) de commettre une erreur statistique. Elle est défini en termes de probabilité, et son seuil de signification en sciences humaines est de 0.05.

28. Hypothèse nulle : qui postule qu'il n'y a pas de différence entre les moyennes des trois groupes (ou des trois mesures).

29. Hypothèse bilatérale : qui postule qu'il existe une différence entre les moyennes des trois groupes (ou des trois mesures).

Dans le but d'estimer l'expérience d'apprentissage, nous nous sommes basés sur les questions relatives au modèle GLO [Fuchs 07][Hooper-Greenhill 03]. Ce modèle a été inventé par Hooper-Greenhill et al. [Hooper-Greenhill 03], et est largement détaillé au chapitre 2 de ce mémoire. Nous rappelons au lecteur que nos propositions³⁰ dans le cadre de cette thèse se basent, pour une grande partie d'entre elles, sur les travaux de Hooper-Greenhill et al. [Hooper Greenhill 94].

Pour chaque moyen de médiation testé, nous avons posé les questions suivantes :

- **Valeurs et attitudes** « Comment votre perception de l'art et de l'histoire du pays-basque a-t-elle changé ? »
- **Connaissances et Compréhension** « Est-ce que vous avez le sentiment d'avoir appris quelque chose de nouveau ? »
- **Amusement, Inspiration et Créativité** « L'exposition vous a-t-elle surprise ? »
- **Activité, Comportement et Progression** « Est-ce que vous envisagez effectuer des recherches à propos de l'art et de l'histoire du pays-basque ? »
- **Compétences** « Vous sentez-vous capables de reconnaître certains symboles et motifs de l'art du pays-basque ? »

Pour compléter le questionnaire, nous avons sélectionné les questions que l'on retrouve le plus souvent dans la littérature [Dünser 08] ainsi que l'échelle de notation utilisée (de 0 = faible à 14 = fort) et que nous avons jugé adaptée à notre expérience. Ceci est dans le but de mesurer l'appréciation de chaque moyen de médiation. Les questions ont été les suivantes :

- « Est-il facile de faire le lien entre les indications du moyen de médiation et l'oeuvre d'art ? »
- « L'acquisition des informations est-elle facile ou difficile ? (effort ressenti) »
- « Quel confort d'utilisation avez-vous ressenti ? »
- « Quel niveau de stress avez-vous ressenti pour chaque mode de médiation ? »

4.4.2 Résultats et interprétations

4.4.2.1 Nombre de réponses correctes

La figure 4.4.1 récapitule les statistiques descriptives des tests utilisateurs. Il y est indiqué pour chaque mode le nombre moyen de réponses correctes et l'écart-type (standard deviation en anglais, abrégé Std. Deviation). La variable F correspond à la variable du test de Fisher-Snedecor et donc au résultat de l'ANOVA.

30. Nous invitons le lecteur à se référer au chapitre précédent.

Quant à la variable p^{31} , elle représente l'erreur alpha, variable permettant de confirmer ou d'infirmar l'hypothèse bilatérale H1.

Mode	Mean	N	Std. Deviation
Etiquette	1,35	17	,606
Audio-guide	1,41	17	,507
M.A.R.T.S	1,53	17	,514
Total	1,43	51	,539

FIGURE 4.4.1 – Table de moyennes

ANOVA

Nombre

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,275	2	,137	,463	,632
Within Groups	14,235	48	,297		
Total	14,510	50			

FIGURE 4.4.2 – Effet inter-sujet sur le nombre de réponses correctes

L'analyse des données indique que les participants, en utilisant le mode de médiation Etiquette, ont obtenu un nombre moyen (nombre moyen de réponses correctes) de 1.35 ($\sigma = 0.606$), un nombre moyen de 1.41 ($\sigma = 0.507$) pour Audio-guide, comparativement à 1.53 ($\sigma = 0.514$) pour M.A.R.T.S. Nous observons donc un gain en faveur de M.A.R.T.S en terme de nombre moyen de réponses correctes, d'environ 6% en comparaison avec Audio-guide et de 9% en comparaison avec Etiquette.

Le résultat de l'ANOVA à mesures répétées ($F(2,48) = 0.463$, $p = 0.632$) va nous permettre de statuer quant à la dépendance réelle du nombre de réponses correctes en fonction des modes de médiation. La valeur de F obtenue est de 0.463 soit une valeur inférieure à la valeur critique de 3.49 donnée par la table de Fisher-Snedecor. L'hypothèse nulle H_0 doit donc être conservée. La probabilité de commettre l'erreur alpha sur le résultat de F est de $p = 0.632$, soit supérieure

31. Appelé Sig dans la table générée par l'outil SPSS.

au seuil de 5%. La différence entre les trois modes de médiation n'est alors pas statistiquement significative, bien qu'il apparaisse un gain en faveur de M.A.R.T.S.

On peut donc en conclure que statistiquement le mode de médiation n'influence pas sur la rétention du visiteur des informations qui lui sont présentées. Nous commenterons ce résultat par la suite.

4.4.2.2 Les données du questionnaire

	Etiquette		Audio-guide		M.A.R.T.S	
	Mean	Std. deviation	Mean	Std. deviation	Mean	Std. deviation
Q1	6,47	3,002	8,24	2,840	9,82	2,789
Q2	8,12	2,934	9,24	3,231	11,47	2,095
Q3	6,29	3,424	8,76	2,587	11	1,936
Q4	7,41	3,043	7,65	2,234	8,29	2,285
Q5	7,41	3,261	8,35	2,714	8,59	3,537
Q6	5,76	2,562	8,35	1,693	11,65	1,539
Q7	7,88	3,100	9,94	2,633	10,82	2,038
Q8	6,47	3,970	10,59	1,770	9,35	2,060
Q9	3,47	3,923	3,88	3,238	3,82	2,789

FIGURE 4.4.3 – Notes moyennes pour chaque mode en fonction des questions.

À la question 1 (Q1) : « Comment votre perception de l'art et de l'histoire du pays-basque a-t-elle changé ? » (0=Très défavorable à 14=Très favorable), nous obtenons le classement Etiquette<Audio-guide<M.A.R.T.S. En revanche, le degré d'accord entre les sujets n'est que de 33.7%. Les rangs moyens nous indiquent que le désaccord porte sur Etiquette et Audio-guide. L'examen des moyennes nous indique cependant que les sujets accordent l'amélioration de leur perception de l'art et de l'histoire basque à M.A.R.T.S et à l'Audio-guide.

La question 2 (Q2) porte essentiellement sur le sentiment d'avoir appris quelque chose de nouveau. Suite aux réponses des visiteurs, nous obtenons le classement Etiquette<Audio-guide<M.A.R.T.S. Cependant, le degré d'accord entre les sujets est de 12.6 %. En regardant les rangs moyens, on remarque que le désaccord porte surtout sur Etiquette et Audio-guide. Les moyennes permettent de voir une légère supériorité de l'Audio-guide par rapport à l'Etiquette et que les trois systèmes permettent au visiteur d'apprendre de sa visite. Toutefois, tous ont jugé M.A.R.T.S

boles et motifs de l'art du pays-basque? » (0=Pas du tout à 14=Tout à fait), nous obtenons le classement Etiquette<Audio-guide<M.A.R.T.S. La valeur d'accordance (W) étant très proche de 43%, ne permet donc de rien conclure quant à l'accord entre les sujets. En se penchant sur les rangs et les moyennes, nous pouvons déduire que les visiteurs estiment acquérir le même degré de compétences avec l'un ou l'autre des trois systèmes expérimentés.

A la question 6 (Q6) : « Est-il facile de faire le lien entre les indications du moyen de médiation et l'œuvre d'art ? » (0=Très difficile à 14=Très facile), nous obtenons le classement Etiquette<Audio-guide<M.A.R.T.S. avec un degré d'accord de 73.7%. En se penchant sur les rangs et les moyennes, on déduit une large supériorité de M.A.R.T.S et de l'Audio-guide. Avec ces deux systèmes, le visiteur peut se placer devant l'œuvre d'art et donc pourra plus facilement faire le lien entre les commentaires audio et celui-ci. D'un autre côté, la différence entre M.A.R.T.S et Audio-guide renvoie surtout à l'efficacité de la désignation assurée par notre paradigme d'interaction « Sélection ». Ce résultat illustre parfaitement l'utilité des modes de communication basés sur l'image.

La question 7 (Q7) concerne la facilité de l'acquisition des informations communiquées par l'outil expérimenté. Nous avons demandé aux participants : « L'acquisition des informations est-elle facile ou difficile ? » (0=Très difficile à 14=Très facile). Nous obtenons donc le classement Etiquette<Audio-guide<M.A.R.T.S avec un degré d'accord de 5%. Les rangs moyens et les moyennes montrent l'infériorité de l'Etiquette en comparaison avec Audio-guide et M.A.R.T.S. Le désaccord porte essentiellement sur ces deux derniers outils étant donné que les deux se basent essentiellement sur les commentaires audio. Ce résultat montre que les sujets préfèrent écouter les descriptions relatives à l'œuvre d'art plutôt que de les lire.

La question 8 (Q8) traite plutôt du degré du confort ressenti avec chaque système. À la question : « Quel confort d'utilisation avez-vous ressenti ? » (0=Très inconfortable 14=Très confortable), nous obtenons le classement Etiquette<M.A.R.T.S<Audio-guide. Cependant le degré d'accord est seulement de 17.2%. En examinant de plus près les rangs et les moyennes, nous voyons que le désaccord porte surtout sur M.A.R.T.S et Audio-guide. Le confort d'utilisation ressenti n'est pas à remettre en cause avec ces deux derniers. Toutefois, tous ont jugé Etiquette très peu confortable à utiliser. Cela est certainement à mettre en relation avec le fait que les utilisateurs se trouvent parfois obligés de se déplacer entre l'Etiquette et l'œuvre d'art afin de pouvoir faire le lien entre les descriptions et celui-ci.

Enfin, à la question 9 (Q9) : « Quel niveau de stress avez-vous ressenti pour

chaque mode de médiation ? » (0=Très faible à 14=Très fort), le classement est M.A.R.T.S <Audio-guide<Etiquette avec un degré d'accord de 23%. Nous remarquons d'après les données statistiques que le désaccord porte sur l'Etiquette et l'Audio-guide qui, sans pouvoir les départager, ont été jugé comme n'induisant pas un niveau de stress important. Il est aussi à noter que les trois systèmes ont été considéré très peu stressants.

4.4.2.3 Discussion des résultats obtenus

Les premiers résultats semblent montrer que statistiquement parlant, il n'y a pas de différence significative en termes de nombre de réponses correctes concernant les objets exposées, et ce malgré un gain apparent en faveur de M.A.R.T.S (9% en comparaison avec l'Etiquette et 6% en comparaison avec l'Audio-guide). Ce résultat peut être mis en balance avec le nombre de participants impliqués dans l'expérimentation. Plus l'effectif (ou le nombre de degrés de liberté) est petit, plus la variance sans biais est incertaine. Néanmoins, les sujets ont été sélectionnés pour leur représentativité et ces premiers résultats nous permettent de dégager une première tendance. Ceci est conforme avec les résultats des questions 5 et 2 qui montrent que les visiteurs ont l'impression d'acquérir de nouvelles connaissances et compétences de la même manière avec l'un ou l'autre des trois systèmes.

Toutefois, nous avons observé une disposition favorable des sujets à utiliser l'Audio-guide et M.A.R.T.S. Ceci se manifeste surtout par les réponses des sujets en ce qui concerne la facilité d'acquisition d'informations (question 7). En effet, dans le cas de ces deux systèmes, les utilisateurs ont fortement apprécié le fait d'écouter les descriptions relatives à l'œuvre d'art au lieu de les lire (cas de l'Etiquette). Néanmoins, une différence entre les deux persiste sur les classements se rapportant surtout à la question 6. En fait, dans la majorité des cas, nous avons eu des remarques sur le fait qu'il est plus facile avec M.A.R.T.S de repérer les éléments de l'œuvre d'art référencés par les commentaires audio. Ceci est à mettre en rapport directement avec notre paradigme d'interaction « Sélection » permettant de référencer les zones d'intérêt sur l'œuvre d'art et donc de faire facilement le lien avec la réalité. Dans le cas de l'Audio-guide, il faut alors faire l'effort de bien chercher sur l'œuvre d'art afin de faire le lien avec les commentaires audio, ce qui a souvent été difficile pour les visiteurs. Il est aussi à noter que l'utilisation de la RA grâce à nos paradigmes « Documentation » et « Sélection » a suscité des sentiments de surprise (question 3) chez la majorité des participants qui se sont montrés très impressionnés par celle-ci. Ces derniers ont jugé l'exposition être plus saisissante grâce à la RA.

Par ailleurs, nous avons relevé beaucoup de remarques concernant M.A.R.T.S

et se rapportant à l'ergonomie de l'Audio Bone. La non-immersivité de ce type de casque a été beaucoup apprécié par les utilisateurs conservant ainsi la possibilité de converser avec leurs compagnons. En revanche, malgré un poids égal à 117 grammes, Samsung SII semble être lourd pour les visiteurs. Ces derniers ont exprimé vouloir le garder dans leur poche (comme ce qu'ils ont fait avec l'Audio-guide) pendant qu'ils écoutaient le discours du guide humain virtuel, et ne le pointait vers l'objet exposé que lorsqu'il s'agit d'afficher les paradigmes « Sélection » et « Documentation ». Nous pouvons donc conclure que pour améliorer le confort d'utilisation ressentie avec M.A.R.T.S, la simulation de coprésence du guide humain peut être réduite à l'aspect conversationnel. Ce constat est conforme aux résultats de la question 8 dans laquelle les sujets ont jugé l'Audio-guide plus confortable à utiliser que notre système M.A.R.T.S.

Conclusion

Dans ce chapitre, nous avons présenté le système M.A.R.T.S. Ce système est une implémentation matérielle et logicielle de la proposition de fonctionnement d'un système d'assistance à la visite au musée que nous avons énoncée au chapitre précédent. Nous proposons donc ici un système de médiation respectant les principes du modèle contextuel de Falk [Falk 05] pour l'apprentissage au musée.

Durant nos travaux, nous avons été en mesure d'étudier la pertinence de M.A.R.T.S en tant qu'interface permettant de décrire les objets exposés au visiteur. Plus précisément, nous avons pu expérimenter le principe de la simulation de la présence d'un guide humain et de nos paradigmes « Sélection » et « Documentation ». Ceci est dans le but d'estimer leurs implications sur l'expérience d'apprentissage au musée.

À travers des tests utilisateurs, nous avons constaté qu'employer les paradigmes « Sélection » et « Documentation » pouvait faire gagner environ 9% en termes de rétention d'informations en comparaison avec l'Étiquette et 6% en comparaison avec l'Audio-guide. Nous avons aussi prouvé que ces paradigmes permettent de faciliter l'acquisition d'informations. En outre, il était plus simple au visiteur d'établir le lien entre les descriptions et l'œuvre exposée grâce au mécanisme de désignation basée image « Sélection ». Nous devons cependant poursuivre nos expérimentations sur un échantillon plus important d'utilisateurs pour que le résultat soit statistiquement plus significatif.

Partie 2

« The art of progress is to preserve order amid change and to preserve change amid order. »"

(Alfred North Whitehead)

Chapitre 5

M.A.R.T.S : Vers un Système de RAM pour la visite guidée à l'extérieur

Introduction

A l'instar de la ville de Bayonne, les villes et pays d'arts et d'histoires sont riches en monuments historiques. Les immeubles en constituent une grande partie¹. De ce fait, plusieurs sorties touristiques sont organisées à l'extérieur afin de découvrir ce genre de patrimoine historique. Dans cette optique, nous nous intéresserons dans ce chapitre à comment créer une visite guidée permettant d'inciter le touriste à apprendre de son activité.

De plus, à l'extérieur se pose le problème de mobilité; le touriste doit pouvoir atteindre un point d'intérêt dans un environnement méconnu. Comment le touriste peut être guidé de telle sorte que sa visite soit la moins gênante possible? Quel mode utiliser pour la navigation afin que le touriste puisse garder toute son attention sur l'environnement qui l'entoure et donc laisser plus de place pour la découverte et l'apprentissage?

1. Les immeubles protégés au titre des monuments historique sont de l'ordre de 44 236 monuments, selon les statistiques de 2012 : <http://www.data.gouv.fr/fr/dataset/liste-des-immeubles-protoges-au-titre-des-monuments-historiques>

5.1 Visite guidée à l'extérieur

A travers la découverte des lieux, la rencontre des monuments historiques² fait apparaître une dimension de l'apprentissage qui semble relever d'une confrontation à la réalité. Cette dimension d'apprentissage immédiat est évidente et est appréciée par les touristes [Brogère 12]. Ceci s'inscrit dans le cadre d'une expérience d'apprentissage tel que mise en avant par Van Winkle et al. [Van Winkle 12]. Cependant, à l'extérieur, s'ajoute le problème de la mobilité. En effet, le touriste doit se déplacer de sa localisation vers le point d'intérêt, qui est dans notre cas un immeuble patrimonial. Comme on peut le relever dans les travaux de Reitmayr et al. [Reitmayr 04], un système numérique de visite guidée à l'extérieur doit accomplir deux fonctions principales : la recherche d'itinéraire³ et l'apport d'informations concernant le point d'intérêt.

Comme le fait remarquer Van Winkle et al. [Van Winkle 12], l'apprentissage dans le tourisme se fait grâce à des opportunités planifiées et d'autres non planifiées. Ainsi, afin de permettre à l'utilisateur de se focaliser sur l'environnement extérieur et donc laisser plus de place à l'apprentissage, nous suggérons de rendre transparente la fonction de navigation. Ceci veut dire que l'utilisateur se laissera guider par les instructions de navigation tout en gardant son attention sur son environnement. Les sens visuel et auditif étant les sens dominants pour l'individu⁴[Nagel 04], nous proposons donc d'utiliser la modalité haptique pour transmettre à l'utilisateur des informations de direction lui permettant d'atteindre sa destination. De plus, contrairement aux autres sens, la modalité haptique n'est pas vite surchargée et peut être fortement exploitée pour des usages autres que la navigation. L'implémentation de la fonction de navigation sera donc détaillée plus loin dans ce chapitre.

5.2 La navigation

Comme expliqué dans la première section de ce chapitre, nous avons eu recours à la modalité haptique afin de ne pas obstruer les sens auditif et visuel. Ces derniers considérés comme les sens dominants de l'individu [Nagel 04] seront donc exploités afin d'apporter de nouvelles connaissances au touriste. Dans cette optique, nous avons mis en place un système de RA haptique, nommé HaptiNav [Ghouaiel 13a],

2. Nous nous intéressons dans cette partie de mémoire aux immeubles patrimoniaux.

3. Ce terme est une traduction de l'expression anglaise « wayfinding ».

4. Nagel affirme que la vue et l'ouïe constituent les sens les plus élevés de l'être humain. Ces deux sens sont séparés l'un de l'autre et des autres sens. En revanche, ils existent des transitions possibles entre le reste des sens, appelés « sens faibles ».

que nous détaillons dans les sections suivantes. HaptiNav illustre le concept d'augmentation haptique AH, référencé dans la thèse de Bayart [Bayart 07]. En fait, il utilise la modalité haptique pour transmettre à l'utilisateur des informations de direction.

5.2.1 Principe

Conception de l'interface haptique

La figure 5.2.1 met en avant trois prototypes possibles pour notre système vibro-tactile HaptiNav, d'aide à la navigation piétonne. Le premier prototype montre la tablette tactile, un système conçu pour être posé sur la paume de la main. Yang et al. [Yang 10] proposent une tablette tactile composée de 12 panneaux dont chacun contient un moteur vibreur « tactor ». Leur tablette tactile contient 12 vibreurs disposés suivant une matrice de 4 lignes et de colonnes. Afin de transmettre à l'utilisateur une information spatio-directionnelle, le système T-mobile [Yang 10] combine trois vibreurs. Par exemple, ce système actionne les vibreurs de la première ligne pour indiquer la direction du nord. L'inconvénient majeur de ce prototype qu'il occupe l'une des mains de l'utilisateur, l'empêchant donc de se déplacer mains libres. Le second prototype se base sur un seul vibreur, en l'occurrence, le vibreur intégré au téléphone mobile. Le système Pocket Navigator citePielot :2010 [PocketNavigator 10] utilise ce prototype, il code la direction que l'utilisateur doit suivre en différents modèles de vibration appelés tactons. L'approche présentée dans Pocket Navigator utilise trois tactons différents pour indiquer à l'utilisateur de continuer tout droit, tourner à droite, tourner à gauche ou s'arrêter. Comme le montre la figure 5.2.2, le système traduit l'action de continuer tout droit par deux courtes vibrations, l'action de tourner à gauche par une longue vibration suivie par une courte, l'action de tourner à droite par une courte vibration suivie par une longue et l'action de faire un demi-tour par trois courtes impulsions.

Nous avons exclu ce dernier prototype, mono-vibreux, puisque Sonja et al. [Salisbury 04] soulignent que Pocket Navigator engendre beaucoup d'erreurs de navigation. En outre, il induit une charge mentale importante, responsable de la diminution de la vitesse de marche de l'utilisateur. Nous avons choisi de développer le prototype de la ceinture vibro-tactile essentiellement en raison de la simplicité d'usage, de la fiabilité et de du confort de navigation. Il est à noter que, contrairement à ActiveBelt [Tsukada 04], notre système contient uniquement quatre vibreurs. Ces derniers sont disposés comme le montre la figure 5.2.1, un de chaque côté de la ceinture, un devant et un derrière. Les vibreurs de devant, de gauche, de droite et de derrière indiquent respectivement qu'il faut aller tout droit, aller à gauche, aller à droite et rebrousser chemin. Quand le système déclenche deux vibreurs, c'est la

	<p>La tablette vibro-tactile</p>
	<p>Le vibreur du téléphone mobile</p>
	<p>La ceinture vibro-Tactile</p>

FIGURE 5.2.1 – Prototypes possibles

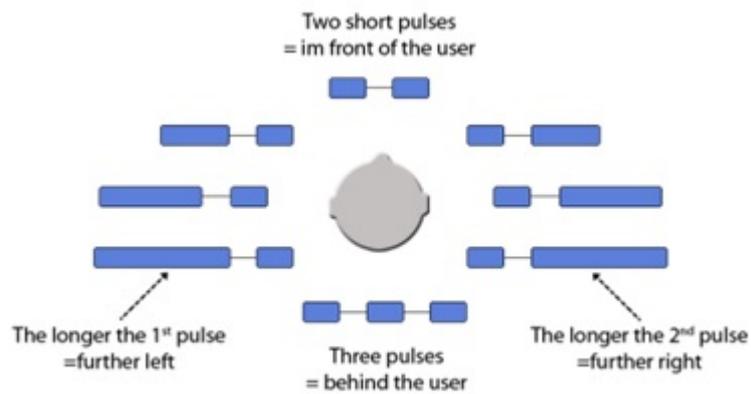


FIGURE 5.2.2 – Schéma tactile de Pocket Navigator citePielot :2010

direction entre les deux qu'il faut prendre.

Algorithme



Nous conseillons au lecteur souhaitant obtenir un complément d'information sur les fichiers KML, de consulter l'annexe A.

Voici un exemple de requête google maps : <http://maps.google.com/maps?f=d&hl=en&saddr=4000,1000&daddr=4020,1030&ie=UTF8&om=0&output=kml>

Le paramètre « saddr » contient les coordonnées GPS de la position présente de l'utilisateur et le paramètre « daddr » contient les coordonnées GPS de la destination.

Le changement d'orientation OT , signifie l'angle avec lequel il faut tourner pour passer de l'orientation courante à l'orientation désirée. Il est obtenu en calculant la différence entre OD et OC .

$$OT = OD - OC \quad (5.2.1)$$

Puisque $0 \leq OD < 360$ et $0 \leq OC < 360$, nous avons donc $-360 < OT < 360$. Par la suite, deux cas peuvent se présenter ; le premier cas est quand l'utilisateur doit se déplacer dans le sens des aiguilles d'une montre ($0 \leq OT < 360$), le deuxième cas est quand l'utilisateur doit se déplacer dans le sens contraire aux aiguilles d'une montre ($-360 < OT < 0$). Ces deux cas sont illustrés par la figure 5.2.3.

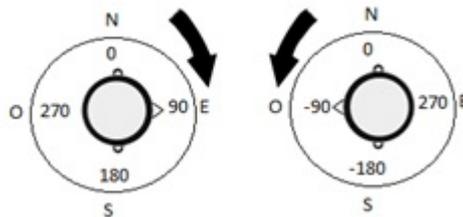


FIGURE 5.2.3 – Changement d'orientation

Le signe de la différence d'orientation, OT , détermine le sens de déplacement et donc le vibreur à déclencher.

5.2.2 Implémentation

Comme le montre la figure 5.2.4, le prototype matériel de notre système Hapti-Nav est constitué d'un Smartphone de type Galaxy S2 fonctionnant sous Android, d'une carte électronique de type Arduino BT [ArduinoBT 11] et d'une simple ceinture sur laquelle on a fixé 4 vibreurs. Arduino BT est une carte Arduino avec le module bluetooth intégré, permettant ainsi la communication sans fil avec le Smartphone. Pour implémenter la partie applicative du système, déployée sur le mobile et sur la carte, nous avons utilisé les SDKs Android et Arduino. Nous nous sommes servis d'Arduino et d'Android pour développer respectivement l'application installée sur la carte électronique et celle installée sur le Smartphone. L'application de la carte électronique permet de contrôler les vibreurs et le microcontrôleur. Cependant, l'application du Smartphone permet de localiser l'utilisateur et calculer son trajet. Pour faire communiquer les deux applications, nous nous sommes appuyés sur l'interface logicielle Amarino, développée dans le cadre du projet « Android meet Arduino » [Amarino 10]. Ce projet a vu le jour au sein de l'université de Klagenfurt en Autriche.

Le schéma ci-dessous met en relief la structure de la partie applicative de notre système. Comme expliqué auparavant, la partie applicative du système est scindée en trois modules à savoir le module Smartphone, le module Carte et le module de communication Smartphone-Carte.

Une fois que l'utilisateur fait entrer l'adresse de sa destination, l'application mobile interroge le serveur google maps sur l'itinéraire à suivre. Ce dernier retourne l'itinéraire requis sous forme de fichier KML. Par la suite, l'application mobile permet de localiser et de guider, en permanence, l'utilisateur. En effet, il s'agit au début de calculer sa position GPS afin de trouver le point repère le plus proche.



Nous remarquons que l'itinéraire décrit par le fichier KML comporte plusieurs points repères. Chaque point repère est caractérisé par sa position GPS et par son orientation.

Ensuite, l'application mobile calcule la différence d'orientation entre l'utilisateur et le point repère en question. En plus, elle envoie l'angle, représentant la différence d'orientation, à l'application téléchargée sur la carte électronique. L'application mobile communique l'angle calculé à la carte, via Bluetooth 2.0, en utilisant le module Smartphone-Carte basé sur la bibliothèque logicielle Amarino [Amarino 10]. Enfin, l'application détermine le(s) vibreur(s) correspondant(s) à la

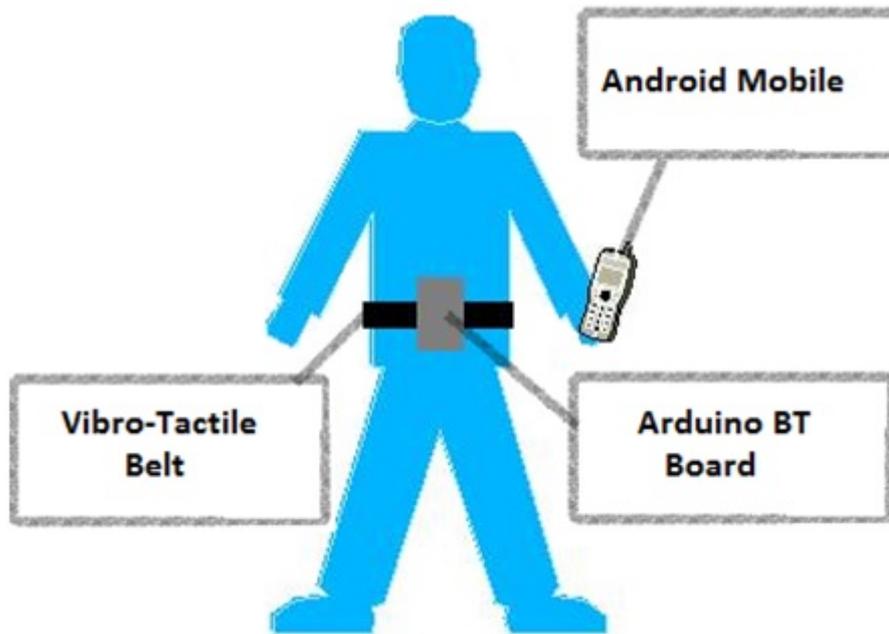


FIGURE 5.2.4 – Prototype matériel d’HaptiNav



FIGURE 5.2.5 – Ceinture vibro-tactile



FIGURE 5.2.6 – Structure de la partie applicative de système

direction transmise, permettant de l(es) activer. Cette boucle continue jusqu'à la fin de l'application.



FIGURE 5.2.7 – HaptiNav

5.2.3 Etude Expérimentale

Nous pouvons comparer notre système, décrit dans les sections précédentes, à d'autres existants. Nous avons omis volontairement le travail de Lin et al. [Lin 08]

vu que le système y proposé ne permet de communiquer que deux directions, gauche et droite. Nous avons choisi de comparer notre système avec la version mobile de l'application Google Maps, en mode audio, très connue du grand public. De plus, nous avons comparé HaptiNav à l'application Pocket Navigator [Pielot 10] [PocketNavigator 10], fonctionnant en mode vibro-tactile.

Protocole

Dans la présente étude expérimentale, nous avons été amenés à conduire trois expérimentations, dans lesquelles nous avons fait varier le mode de navigation. Les expérimentations en question, ont eu lieu à la technopole Izarbel, lieu dans lequel se situe l'ESTIA. En effet, la première consiste à utiliser l'application Google Maps, lancée en mode audio. La deuxième consiste à naviguer avec le mode vibro-tactile en utilisant HaptiNav. La dernière utilise le système Pocket Navigator, fonctionnant également en mode vibro-tactile. Au cours des trois expérimentations, nous avons gardé le même trajet, illustré par la figure 5.2.8, étant donné que chaque sujet ne participe qu'à une seule expérimentation.



FIGURE 5.2.8 – Trajet emprunté lors des expérimentations

Lors d'un test, le sujet est suivi par un expérimentateur, chargé de gérer le tableau de bord. Il est interdit pour le participant, de regarder l'écran du Smartphone et de poser des questions à l'expérimentateur. Le sujet doit marcher avec sa vitesse habituelle. Aucune phase d'apprentissage ne précède la phase du test.

Participants

36 personnes non rémunérés, dont 18 femmes et 18 hommes ont participé à la l'étude expérimentale. Ce sont en général, des étudiants, des stagiaires ou des employés de l'école supérieure ESTIA. Leurs âges varient de 22 ans à 39 ans (moyenne =30.5). Ils sont divisés en trois groupes, chaque groupe est formé de 12 personnes. Chacun des groupes participe respectivement à la première, deuxième et troisième expérimentation. Il est à noter qu'aucun des sujets n'est impliqué dans les travaux de recherche présentés dans ce mémoire. Tous les participants ont déjà utilisé un plan géographique, et 23 sur 24 parmi eux, ont l'habitude d'utiliser les systèmes de navigation électroniques, tels que Tom-tom. Il existe deux sujets qui ne sont jamais rendu sur le lieu de l'étude expérimentale, néanmoins, les autres le connaissent déjà. En revanche, le fait de connaître ce lieu ne présente aucune importance, étant donné que les sujets ne peuvent connaître la destination avant la fin de l'expérience.

Etude Quantitative

Les résultats de l'étude expérimentale montrent que chacune des trois techniques de navigation peut acceptablement être utilisée pour la navigation piétonne. La figure 5.2.9 illustre le nombre d'erreurs de navigation ainsi que le nombre d'événements de désorientation qui ont eu lieu. Nous précisons au lecteur que l'expérimentateur enregistre une erreur de navigation, quand le sujet ne suit pas la direction indiquée par le système de navigation, au niveau d'un point repère. Un point repère est schématisé par une flèche bleue, au niveau de la figure 5.2.8. En outre, l'expérimentateur note un événement de désorientation, quand le sujet s'arrête plus de 10 secondes, ou quand il annonce à l'expérimentateur qu'il est confus. Suite à une erreur de navigation ou à un événement de désorientation, le sujet est redirigé vers la bonne direction.

Le diagramme schématisé par la figure 5.2.9, montre le nombre d'erreurs de navigation et le nombre d'événements liés à la désorientation des trois groupes de participants, par rapport aux trois systèmes étudiés. Nous pouvons en déduire qu'HaptiNav et Google Maps, ont des nombres d'erreurs très proches, respectivement 29 et 27. En revanche, les désorientations ont eu lieu presque deux fois plus souvent avec Google Maps. De plus, les événements de désorientation se sont produits presque trois fois plus souvent avec Pocket Navigator en comparaison avec HaptiNav.

Il est important de souligner que certaines erreurs de navigation survenues avec HaptiNav et Pocket Navigator, se rapportent à une mauvaise analyse mentale du signal vibro-tactile. En plus, nous remarquons que les trois systèmes sont soumis aux imprécisions engendrées par le système GPS et par la centrale inertielle.

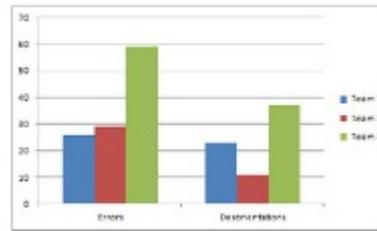


FIGURE 5.2.9 – Erreurs et désorientations

La figure 5.2.10 met en avant la comparaison des moyennes d’erreurs entre les trois groupes. Le plus petit nombre d’erreurs enregistré pour HaptiNav est 0. Cependant, le plus petit nombre d’erreurs enregistré pour Pocket Navigator est 4.

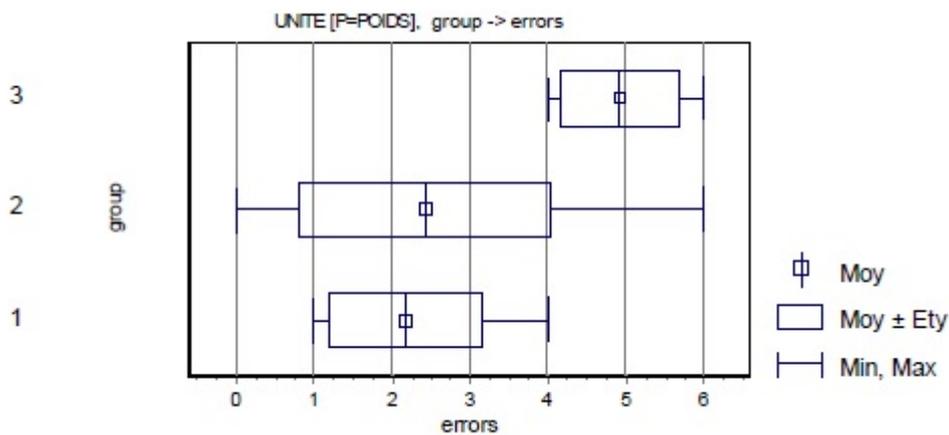


FIGURE 5.2.10 – Diagramme de moyenne d’erreurs

Étant donné que le test d’ANOVA donne une valeur de p^5 égale à 0.6641 ($p = 0.6641 > 0.5$), cela nous permet de conclure qu’il n’existe pas de différence entre les moyennes d’erreurs de la première et de la deuxième expérience. Nous pouvons donc en déduire que les performances de HaptiNav sont proches de celles de Google Maps Audio. En plus, le test d’ANOVA donne une valeur de p égale à 0,0001 ($p = 0,0001 < 0,5$) pour la deuxième et la troisième expérience, ce qui nous permet de déduire qu’ HaptiNav est meilleur que Pocket Navigator, dans le

5. valeur de probabilité comprise entre 0 et 1

sens où il permet de réduire les erreurs de navigation.

La figure 5.2.11 suivante met en relief la comparaison des moyennes des événements de désorientation entre les trois groupes étudiés. Puisque la valeur de p ($p = 0.0014 < 0.05$) est en dessous du seuil significatif de 0.05, cela permet de conclure qu'il existe une différence significative entre HaptiNav et Google Maps. En conséquence, nous pouvons affirmer qu' HaptiNav est plus performant que Google Maps en termes des événements de désorientation enregistrés au cours de la navigation. En outre, le test de l'ANOVA met en évidence une différence significative entre Google Maps et Pocket Navigator ($p = 0.0028 < 0.05$). Ainsi, il est évident que le système Pocket Navigator induit plus de désorientations que le système Google Maps.

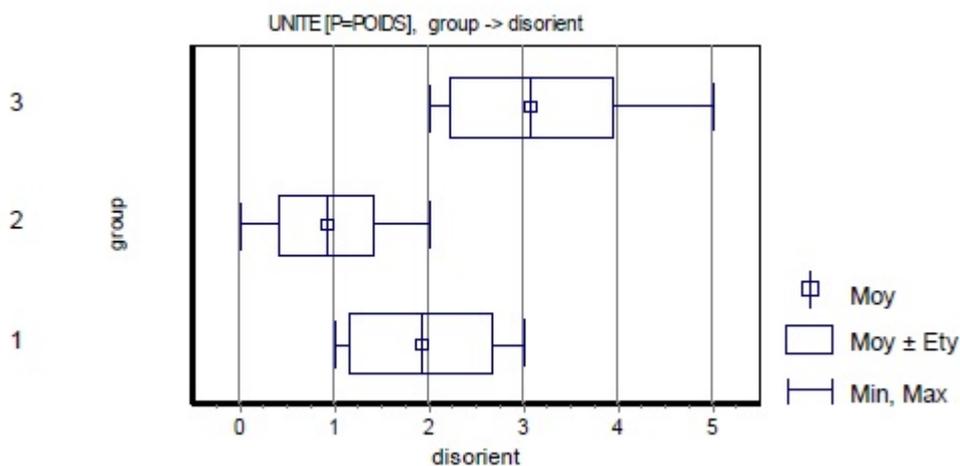


FIGURE 5.2.11 – Diagramme de nombre des désorientations

Etude qualitative

Dans la présente étude qualitative, nous avons été amenés à utiliser le test Nasa TLX «Task Load Index » 1.0 [Center 86], permettant d'estimer d'une manière subjective la charge induite par l'expérimentation. Le test Nasa TLX se base sur six facteurs pour mesurer la charge : la charge mentale « mental demand », la charge physique « physical demand », la charge temporelle « temporal demand », la performance, l'effort et la frustration.



Nous conseillons au lecteur souhaitant obtenir un complément d'information sur le test Nasa TLX, de consulter la référence suivante : [Center 86].

Chaque sujet est invité à évaluer, en attribuant une note de 0 à 20, le système étudié par rapport à chaque facteur. Aucune différence significative n'est détectée entre HaptiNav et Google Maps Audio en termes de charge physique, de charge temporelle et de performance ($p = 0,68 > 0,05$). Néanmoins, l'analyse statistique de la charge mentale, de la frustration et de l'effort, fournit des valeurs de p inférieures à 0,5, pour le système HaptiNav, quand ce dernier est comparé à Google Maps et à Pocket Navigator. Ces résultats permettent de mettre en évidence l'avantage de HaptiNav par rapport aux deux autres systèmes. En outre, les résultats obtenus montrent que la charge mentale notée pour HaptiNav est en moyenne égale à 4,5 points, ce qui explique la valeur basse de la frustration induite par ce système. Nous précisons au lecteur que la charge mentale relative à HaptiNav, permet de traduire l'effort mental que met le participant pour interpréter le signal transmis par ce système. Contrairement à HaptiNav, Pocket Navigator est classé avec le plus haut niveau de frustration, qui est de 17,84 points.

Nous avons posé la question suivante au deuxième groupe de personnes ayant utilisé HaptiNav : ***Accepteriez-vous d'utiliser notre système pour trouver votre destination dans une ville que vous visitez pour la première fois?*** 41% des participants ont répondu par non, ce qui signifie que 5 sujets sur 12 interrogées, refusent de porter notre interface haptique lors de la visite d'une ville pour la première fois. Ces participants expliquent qu'ils ne souhaitent pas porter une ceinture en ville parce qu'ils ne veulent pas se faire remarquer par les autres piétons. Ils suggèrent que ce système pourrait être plus petit. Nous avons posé la question suivante au troisième groupe ayant participé à l'expérience de navigation avec Pocket Navigator) : ***Accepteriez-vous d'utiliser le système Pocket Navigator pour trouver votre destination dans une ville que vous visitez pour la première fois?*** Plus de 80% précisent qu'ils n'acceptent pas parce Pocket Navigator est très imprécis. Nous avons posé la question suivante au premier groupe ayant à l'expérience de navigation avec Google Maps Audio : ***Accepteriez-vous d'utiliser Google Maps Audio pour trouver votre destination dans une ville que vous visitez pour la première fois?*** Plus de 80% ont dit qu'ils sont en mesure d'utiliser Google Maps Audio à moins qu'il soit activé, en plus, en mode graphique.

Discussion

Les résultats obtenus confirment que le système HaptiNav peut être utilisé pour la navigation, en mode haptique, mains libres et yeux libres. Nous considérons que l'amélioration de la qualité de navigation, dépend de l'amélioration de l'intensité des vibrations dans le deuxième prototype de système HaptiNav. En effet, les participants remarquent que les vibrations sont difficiles à distinguer lorsque la ceinture du système est portée au-dessus de vêtements épais. Ce fait explique l'occurrence de certaines erreurs de navigation qui ont eu lieu avec HaptiNav. Vu que certains participants disent qu'ils refusent de porter HaptiNav à cause de la ceinture, nous prévoyons donc de la remplacer par un bracelet qui peut être porté autour du poignet, dans le deuxième prototype de notre système.

5.3 M.A.R.T.S en environnement urbain

Contrairement au contexte particulier du musée, on ne peut pas trouver actuellement en littérature de modèle contextuel décrivant l'expérience d'apprentissage à l'extérieur. Néanmoins, comme nous l'avons vu à la section précédente plusieurs auteurs mettent en exergue l'existence d'une telle expérience.

Comme nous l'avons vu dans le troisième chapitre de ce mémoire, la représentation de l'information peut prendre trois formes à savoir : le mode iconique, symbolique et éactif [Bruner 60]. Selon ce même auteur, l'individu tend à privilégier pour l'apprentissage, l'un et/ou l'autre des trois modes selon ses choix, ses préférences, ses habitudes et encore selon la situation dans laquelle il se trouve. La coexistence de ces trois modes de représentation de l'information s'avère donc primordiale afin de rendre compte des différents styles d'apprentissage de touristes. Partant de ce constat, nous avons conçu notre système M.A.R.T.S qui a été expérimenté dans le cadre de visite guidée au musée. De ce fait, si l'on procède à un raisonnement par analogie, à l'extérieur, un immeuble patrimonial peut être assimilé à une œuvre d'art. Tout comme les œuvres d'arts au musée, les immeubles historiques possèdent leurs caractéristiques historiques, civils, architecturales et subissent au cours des âges plusieurs dégradations. Dans cette optique, nos propositions précédemment évoquées, peuvent lors d'une visite guidée à l'extérieur, se présenter comme suit : :

- **Guide humain virtuel** : l'idée d'une exploration guidée, par l'intermédiaire d'un guide touristique par exemple, semble être intéressante pour tout l'engagement qu'elle éveille chez le touriste [Brougère 12]. Ce moyen permet de mobiliser les touristes et donc leur offre la possibilité d'acquérir de nouvelles

connaissances. Dans cette optique, en l'absence d'un guide touristique (personne humaine), employer un guide humain virtuel pour simuler la présence de celui-ci, nous paraît être fortement utile. A l'instar d'un guide touristique, un guide humain virtuel pourrait propulser l'apprentissage mobilisé par l'activité touristique. Afin d'appuyer le discours du guide humain virtuel, nous proposons de faire appel à nos paradigmes d'interaction présentées à la première partie de cette thèse, à savoir : la « Documentation », la « Reconstruction » et la « Sélection ».

- **La « Documentation »** : il nous est donné grâce à la réalité augmentée de pouvoir afficher un texte virtuel renseignant par exemple le nom du bâtiment, la date de sa création, le nom de l'architecte, le style architectural. . .
- **La « Reconstruction »** : grâce à la réalité augmentée, nous avons la possibilité de reconstruire en 3D un bâtiment dégradé. En effet, nous pouvons ajouter des éléments virtuels au visiteur de manière à faire apparaître à quoi ressemblait le bâtiment dans son état original.
- **La « Sélection »** : ce paradigme d'interaction pourrait être utilisé afin de montrer au touriste des parties spécifiques du bâtiment renvoyant à un certain ordre architectural ou incarnant certains symboles.

Nous faisons remarquer au lecteur que cette partie du mémoire fait l'objet de nos actuels travaux. Nous avons toutefois commencé à réfléchir sur certains aspects de la mise en œuvre de M.A.R.T.S à l'extérieur. Il est à noter que les immeubles historiques présentent des spécificités géométriques très différentes de celles des œuvres d'art au musée. Par la suite, la méthode de reconnaissance 2D intégrée au composant reconnaissance du système M.A.R.T.S, ne peut être utilisée pour les bâtiments. De ce fait, nous avons été amenés à travailler sur la reconnaissance automatique des bâtiments. Ceci est présenté en détail dans la section suivante.

5.4 Reconnaissance du patrimoine architectural

Comme nous l'avons pu voir au chapitre précédent, les points d'intérêts sont sensibles aux faux appariements. Ce phénomène s'accroît d'ailleurs en présence d'occultation qui est une situation courante en milieu urbain. Par ailleurs, les bâtiments présentent des structures régulières constituées généralement de segments de droites. De ces faits, plusieurs travaux en vision par ordinateur se sont orientés vers une reconnaissance de bâtiments basée sur les segments de lignes [Li 09][Li 02][Trinh 07].

5.4.1 Algorithme

L'algorithme de reconnaissance de bâtiments que nous présentons dans cette section, s'appuie également sur les segments de lignes droites. Cet algorithme se scinde en trois étapes : l'alignement, la représentation et la mesure de similarité. Ces étapes sont expliquées aux sections suivantes.

Au début, l'algorithme procède à l'alignement de l'image test avec l'image modèle. Ensuite, il calcule les points de fuite dans l'image rectifiée. Si le nombre des points de fuite détectés est nul ou égal à 1, celle-ci est rejetée⁶. Dans le cas contraire, l'algorithme procède à la construction des modèles de lignes (section 5.4.5) relatives à l'image alignée (rectifiée) et à l'image modèle. Enfin, l'algorithme estime le degré de similarité des ces dernières via la mesure de Hausdorff [Huttenlocher 93]. L'algorithme 5.1 décrit en détail les étapes de notre méthode.

5.4.2 Géométrie des points et des lignes de fuite

Dans le monde réel, deux droites sont dites parallèles, si elles n'ont aucun point commun ou si elles sont confondues. Elles se rencontrent au niveau d'un point, imaginaire, jamais atteint, situé à l'infini. Dans une image, les représentations de deux droites parallèles se croisent, d'une façon finie ou infinie, au niveau d'un seul point, appelé point de fuite.

En dessin, les artistes emploient les points de fuite dans leurs dessins, afin de leur aider à dessiner en perspective. En vision par ordinateur, les points de fuite contiennent des informations importantes, pouvant être utilisés par exemple, pour la calibration de la caméra et le suivi du mouvement. Dans cette optique, plusieurs travaux [Elloumi 12] [Pflugfelder 05] [Caprile 90] ont proposé des techniques s'appuyant sur les propriétés des points de fuite, dans le but de déterminer les paramètres intrinsèques et extrinsèques de la caméra. La géométrie relative aux points de fuite est illustrée par la figure 5.4.1.

Les trois points de fuite (v_x, v_y, v_z) représentent dans le plan image ($Z = 1$) les intersections suivant les trois axes X , Y et Z d'un système de coordonnées cartésiennes en trois dimensions, dont le centre du repère est le point C .

De même qu'un point de fuite est l'ensemble de convergence d'un ensemble de droites parallèles, une ligne de fuite est la droite de convergence d'un ensemble de plans parallèles. On peut aussi voir une ligne de fuite d'un plan, comme l'ensemble des points de fuite issus de l'ensemble des droites parallèles appartenant à ce même

6. ce cas décrit un échec du processus d'alignement

Algorithme 5.1 Principe de l'Algorithme

Entrée

$M_{1..N}$: Ensemble des images modèles contenues dans la base
 T : image test

Sortie

L : Modèle de ligne de l'image
 V : nombre de points de fuite
 μ : Mesure de Hausdorff
 K : indice de l'image modèle correspondante au bâtiment

Algorithme

Pour i de 1 à N faire

Aligner T par rapport à $M(i)$

$V \leftarrow$ Calculer les points de fuite de T

Si ($V < 2$)

Exit

Sinon

$L_T \leftarrow$ Générer le modèle de ligne de l'image aligné T

$L(i) \leftarrow$ Générer le modèle de ligne de $M(i)$

$\mu \leftarrow$ Mesure de Hausdorff ($M(i), T$)

Si $\mu \leq \delta$

$K \leftarrow i$

Fin Si

Fin Si

Fin Pour

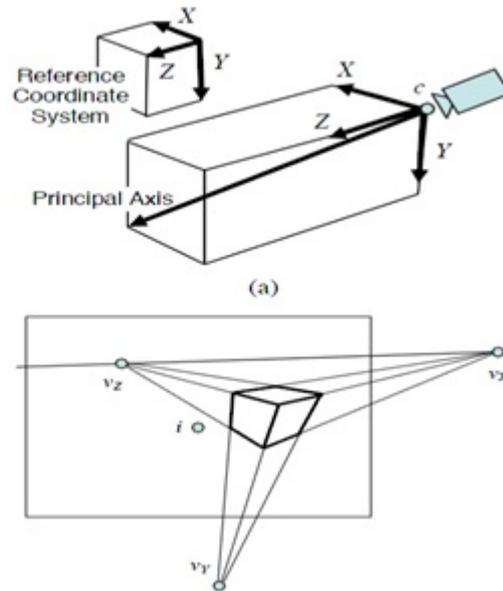


FIGURE 5.4.1 – Géométrie des points de fuite [Kim 06]

plan. Par conséquent, deux plans parallèles possèdent la même ligne de fuite.

D'une façon plus précise, On peut définir une ligne de fuite l_v d'un plan π , comme étant la droite passant par plusieurs points de fuite, issus de plusieurs groupes de droites parallèles. Soient v_x le point de fuite de deux droites parallèles appartenant à π et v_y le point de fuite de deux autres droites parallèles (entres elles, mais pas avec les deux première droites) appartenant elles aussi à π . Les points v_x et v_y sont tous les deux les projetés des points à l'infini situés sur la droite à l'infini issue de l'intersection de π et de π_∞ . La ligne de fuite l_v passe donc par ces deux points et on a alors $l_v = v_x \times v_y$ [Lai 09]. La figure 5.4.2, illustre la ligne l_v définie par une paire de points de fuite. Cette ligne correspond à l'horizon.

5.4.3 Détection des points de fuite

Le calcul de points de fuite a fait l'objet de plusieurs travaux. Les méthodes proposées diffèrent selon la formalisation mathématique des points de fuite et donc selon le choix de l'espace de travail : plan image [Cantoni 01] [Rother 00], plan projectif [Nieto 11] [Nieto 10] ou sphère de Gauss [Barnard 83] [Collins 90] [Kosecka 02] [Lutton 94]. Contrairement au plan image, la sphère de Gauss et le plan projectif permettent de traiter simultanément les points de fuite finis et les

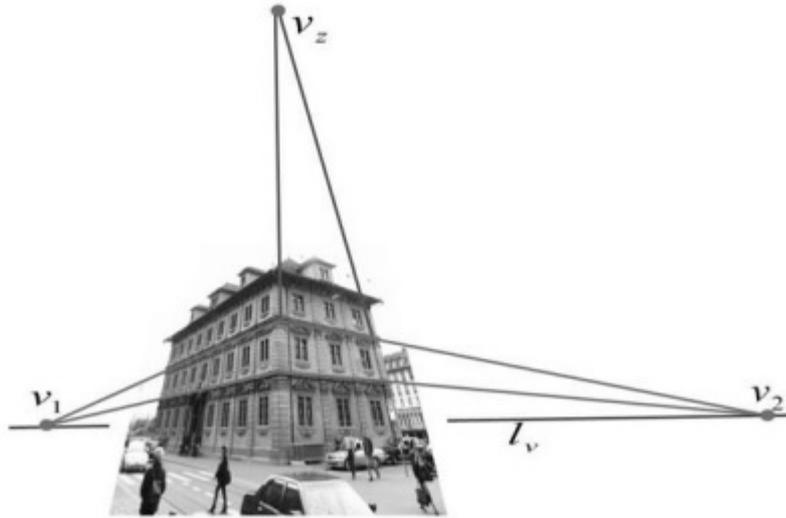


FIGURE 5.4.2 – Géométrie des lignes de fuite [Lai 09]

points de fuite infinis. La méthode basée sur la sphère de Gauss, a été introduite en 1983 par Barnard [Inkila 05]. L'avantage de cette méthode est de ramener dans un espace fini tous les points de fuite. Lutton et al. [Lutton 94] adoptent une approche basée sur la transformée de Hough [Hough 62] afin de calculer les points de fuite. Les méthodes de calcul de points de fuite existantes, varient aussi selon l'approche d'estimation adoptée. Plusieurs travaux [Tardif 09] [Antone 00] emploient l'algorithme Espérance-Maximisation (EM) [Dempster 77] pour regrouper et classifier les lignes qui convergent vers un ou plusieurs points de fuite. Récemment, plusieurs auteurs [Elloumi 12] [Kalantari 06] [Nieto 10] adoptent des méthodes d'estimation robustes basées sur RANSAC [Fischler 81]. En général, les méthodes basées sur RANSAC estiment les points de fuite candidats, en calculant les intersections de segments de lignes. Ensuite, il s'agit d'ajouter itérativement les segments de lignes parallèles compatibles avec ces points candidats.

Méthode de détection de points de fuite adoptée

Nous nous sommes appuyés sur l'approche proposée par Nieto et al. [Nieto 10] pour calculer les points de fuites d'un bâtiment. Nieto et al. [Nieto 10] proposent une méthode permettant d'estimer simultanément plusieurs points de fuite. Leur méthode peut fonctionner en temps réel sur des séquences vidéo. La robustesse de leur méthode revient d'abord au fait qu'il utilisent l'estimateur MSAC (M-estimator SAmple and Consensus) [Torr 97].



Nous conseillons au lecteur souhaitant avoir un complément d'information sur MSAC de consulter le travail de TORR et al. [Torr 97].

En plus, ils mettent en œuvre une nouvelle approche d'estimation d'erreur qui mesure la distance entre un point de fuite et une orientation donnée. La fonction de calcul d'erreur proposée est exprimée par l'équation (4).

$$d(l, v) = \left| \frac{-l_2 s_1 + l_1 s_2}{\sqrt{l_1^2 + l_2^2} \sqrt{s_1^2 + s_2^2}} \right| \quad (5.4.1)$$

En fait, elle est indépendante de la position du point de fuite candidat. Elle peut encore manipuler les points de fuite infinis. En effet, cette fonction consiste à calculer le sinus de la différence entre l'orientation d'un segment de ligne et la direction d'un point de fuite candidat. Nieto et al. [Nieto 10] considèrent la distance entre un point de fuite v et une ligne l , comme étant la valeur absolue du sinus de l'angle entre l et une ligne s . La ligne s joint le point de fuite v à un point référence appartenant à l .

La ligne $s = (s_1, s_2, s_3)$ est calculée par : $s = v \times \frac{1}{2}(a + b)$ et la ligne l par : $l = a \times b$. Les points a et b représentent les extrémités d'un segment de ligne passant par l . La figure 5.4.3 en explique la géométrie en question.

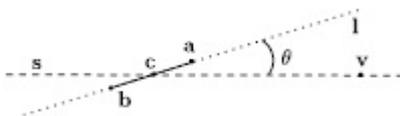


FIGURE 5.4.3 – Géométrie relative à l'erreur

Au niveau de l'équation (4), la valeur absolue est utilisée car l'on souhaite calculer la valeur de la déviation relative entre les lignes sans tenir compte du signe. L'angle entre les deux lignes est limité entre 0 et π . L'usage de la fonction sinus est justifié par le fait que pour des faibles valeurs de déviation de l'angle θ , $\sin\theta$ est quasiment égal à θ , ce qui est le cas des segments de lignes retenus.

L'estimation des points de fuite se base sur un processus itératif, mis en œuvre par MSAC [Torr 97]. À chaque itération, étant donnée un ensemble complet d'observations $L = \{l_i\}_{i=1}^N$, un seul point de fuite candidat est calculé. Chaque segment de ligne l_i est supposé appartenir à l'une des deux classes étiquetées par C_l , avec

$C_l = \{ON, OFF\}$. Les deux valeurs que peuvent prendre C_l indiquent respectivement que, le segment de ligne est orienté vers le point de fuite candidat ou non. Le modèle de vraisemblance est défini par $p(L | V) = \prod_{i=1}^N p(l_i | v)$; avec $p(l_i | v) = \sum_{C_l} p(C_l) p(l_i | v, C_l)$. $p(l_i | v, C_l = ON)$ représente la fonction de densité modélisée par une distribution normale :

$$p(l_i | v, C_l = ON) \propto \left(-\frac{1}{2\sigma^2} d^2(l_i, v)\right) \quad (5.4.2)$$

Dans l'équation (5), $d(l_i, v)$ représente la fonction de calcul de distance donnée par l'équation (4). Après l'étape d'estimation d'un point de fuite, les segments retenus « inliers », sont éliminés de l'ensemble de données afin de calculer le point de fuite suivant, s'il existe.

Pour résumer MSAC procède itérativement en deux étapes :

1. Au début, un échantillon minimal de données est choisi aléatoirement afin de générer un point de fuite candidat. Dans ce cas, l'échantillon de données est composé de deux segments de lignes en plus du point de fuite candidat v^* , calculée par : $v^* = l_i \times l_j$
2. Ensuite, l'ensemble des consensus $s(v^*)$ est estimé. En fait, il s'agit de calculer l'ensemble des segments de lignes qui sont compatibles avec le consensus v^* . Par la suite, les segments de ligne qui donnent naissance au point de fuite v^* , sont sélectionnés. Ainsi, le calcul des consensus $s(v^*)$, s'effectue de la façon suivante : $s(v^*) \triangleq \{l_i \in L : d^2(l_i, v) \leq \delta\}$; δ représente un certain seuil régi par les statistiques des données.

5.4.4 Alignement

L'alignement de point de vue est une phase primordiale pour l'algorithme présenté dans cette section. Au niveau de cette phase, l'image est transformée et est mise à l'échelle en comparaison avec le modèle. En d'autres termes, il s'agit d'aligner l'image test avec le modèle. Cela revient à transformer l'image test afin d'éliminer les invariances d'échelle, de rotation, de translation et de point de vue, par rapport au modèle. L'approche que nous adoptons consiste à retrouver l'homographie qui transforme l'image test en l'image modèle.

La méthode d'alignement que nous utilisons, s'appuie sur les points caractéristiques SURF. Effectivement, les points SURF sont détectés dans le modèle et dans l'image test. Les descripteurs associés aux points clés sont mis en correspondance en utilisant notre premier algorithme, présenté au début de ce chapitre. Les points

clés appariés sont utilisés ensuite afin de calculer l'homographie planaire qui transforme les points de l'image test en points de l'image modèle.

L'homographie planaire ou l'homographie 2D [Agarwal 05] est une relation qui relie les points de deux vues différentes. En effet, les points image se trouvant sur le plan d'une vue, peuvent être convertis en points image correspondants, se trouvant sur le plan d'une autre vue, par l'intermédiaire d'une homographie 2D. En fait, l'homographie 2D consiste en une transformation projective. Elle dépend des paramètres intrinsèques et extrinsèques de la caméra. La matrice d'homographie projective G , est exprimé en fonction de la rotation R , de la translation t et de la normale au plan image n , à l'aide de l'équation suivante :

$$G = y K (R + tn^T) K^{-1} \quad (5.4.3)$$

Avec K : la matrice de calibration de la caméra.

L'homographie dans l'espace euclidien, peut être calculée à partir de la matrice d'homographie projective, en utilisant une estimation de la matrice de calibration K noté \widehat{K}

$$\widehat{K} = \widehat{K}^{-1} G \widehat{K} \quad (5.4.4)$$

Supposons maintenant que la matrice de calibration est inconnue, supposons donc que $\widehat{K} = K$, ce qui donne :

$$\widehat{H} = y(R + tn^T) \quad (5.4.5)$$

Soient un point p appartenant à l'image test et un point q appartenant à l'image modèle. A ce niveau, nous considérons que l'image test et l'image modèle représentent deux vues différentes d'une même image. Soient β_1 et β_2 deux facteurs d'échelle. En utilisant les coordonnées homogènes de q et de p , qui sont respectivement (x_1, y_1, w_1) et (x_2, y_2, w_2) , nous avons :

$$\beta_2 p = H \beta_1 q \quad (5.4.6)$$

L'équation (10) signifie que p est égale à Hq , les deux multiplié par deux certaines valeur d'échelle, respectivement, β_2 et β_1 . Par conséquent, $p \sim Hq$ est une transformation directe entre les points de deux plans d'images. Donc, en coordonnées homogènes, nous avons :

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \quad (5.4.7)$$

En utilisant les coordonnées non homogènes $x'_2 = \frac{x_2}{w_2}$ et $y'_2 = \frac{y_2}{w_2}$, nous avons :

$$x_2' = \frac{H_{11}x_1 + H_{12}y_1 + H_{13}w_1}{H_{31}x_1 + H_{32}y_1 + H_{33}w_1} \quad (5.4.8)$$

$$x_2' = \frac{H_{21}x_1 + H_{22}y_1 + H_{23}w_1}{H_{31}x_1 + H_{32}y_1 + H_{33}w_1} \quad (5.4.9)$$

Pour estimer l'homographie entre deux images planaire d'une même scène, en utilisant les équations (12) et (13), il suffit de fixer w_1 à 1 [Agarwal 05] :

$$x_2' = (H_{31}x_1 + H_{32}y_1 + H_{33}) = H_{11}x_1 + H_{12}y_1 + H_{13} \quad (5.4.10)$$

$$y_2'(H_{31}x_1 + H_{32}y_1 + H_{33}) = H_{21}x_1 + H_{22}y_1 + H_{23} \quad (5.4.11)$$

Etant donnée que que les coefficients de H sont linéaires, l'on peut résoudre H en réajustant les équations (14) et (15), on obtient donc :

$$a_z^T H = 0 \quad (5.4.12)$$

$$a_y^T H = 0 \quad (5.4.13)$$

Avec :

$$H = (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})$$

$$a_z = (-x_1, -y_1, -1, 0, 0, 0, x_2'x_1, x_2'y_1, x_2')$$

$$a_y = (0, 0, 0, -x_1, -y_1, -1, y_2'x_1, y_2'y_1, y_2')$$

Etant donné deux ensembles de points qui correspondent à deux images, on peut former ce système d'équations linéaires :

$$AH = 0 \quad (5.4.14)$$

Avec :

$$A = \begin{pmatrix} a_{z1}^T \\ a_{y1}^T \\ \vdots \\ a_{zn}^T \\ a_{yn}^T \end{pmatrix}$$

L'équation 5.4.14 est résolue en utilisant l'algorithme RANSAC [ref]. La boucle RANSAC se déroule de la manière suivante :

1. Sélectionner aléatoirement 4 correspondances de points SURF.

2. Calculer la matrice d'homographie H en utilisant la méthode des moindres carrés linéaires homogènes, décrite dans [Inkila 05].
3. Calculer les points « inliers » respectant la condition suivante : $\|p, Hq\| < \varepsilon^7$.
4. Recommencer la boucle à partir de l'étape 2 en utilisant la nouvelle liste de points « inliers ».

L'image test est rectifiée en se basant sur l'homographie retrouvée. Tous les pixels de l'image test sont déformés en utilisant la matrice H . Chaque pixel de l'image test est considéré comme un vecteur 2D et il est transformé en utilisant l'équation (10).

La figure 5.4.4 montre les résultats obtenus, relatives à l'alignement de point de vue d'une image test par rapport à une image modèle. La méthode d'alignement adoptée, est celle décrite dans cette section.



FIGURE 5.4.4 – Alignement de point de vue. De la droite vers la gauche : l'image test, l'image modèle, l'image rectifiée

5.4.5 Représentation

Au niveau de cette étape de notre algorithme, nous proposons de représenter un bâtiment par l'intermédiaire des lignes de Hough ayant servi au calcul des trois points de fuite. Les trois points de fuite correspondent aux trois directions orthogonales X , Y et Z , appelés aussi les directions principales du monde de Manhattan. Nous rappelons au lecteur que nous nous sommes inspirés de la méthode de Nieto et al. [Nieto 10], afin de calculer les points de fuite. Cette méthode est détaillée de manière extensive au niveau de la section 5.4.3. Représenter un bâtiment par un modèle de lignes est intuitif. Bien évidemment, ceci se justifie par la structure régulière des bâtiments, comportant des segments de lignes droites. Le processus de construction d'un modèle de lignes d'un bâtiment, peut se résumer par les étapes suivantes :

7. Epsilon représente le seuil de rejet de RANSAC. Nous avons fixé ce seuil à 3.0 qui est la valeur par défaut définie par OpenCV.

- Extraction des contours par le filtre de Canny [Canny 86].
- Détection de lignes de Hough [Hough 62].
- Regroupement des lignes de Hough suivant leurs directions.

En effet, l'image est tout d'abord réduite à une image de contours. Les contours sont considérés comme la façon la plus compacte, permettant de représenter les formes des objets contenus dans une image. Ils correspondent à des zones de fortes variations locales d'intensité des pixels d'une image. Un filtre de détection de contours permet de repérer les emplacements des pixels qui font partie des contours. Le détecteur de Canny [Canny 86] est utilisé pour extraire les contours. Par la suite, les lignes de Hough [Hough 62] sont extraites de ces contours. La transformée de Hough [Hough 62] est une technique générale qui permet d'identifier les emplacements et les orientations des courbes et des lignes dans une image. La transformée de Hough calcule les valeurs des paramètres qui caractérisent le modèle d'une ligne, à savoir la pente \widehat{m} et l'intercepte \widehat{c} . L'idée clé de la transformée de Hough, consiste à identifier les ensembles de points colinéaires dans une image. Un ensemble de points qui se trouvent sur une ligne droite peut être définie par une relation f , de telle sorte que :

$$f((\widehat{m}, \widehat{c}), (x, y)) = y - \widehat{m}x - \widehat{c} = 0 \quad (5.4.15)$$



Pour avoir plus de détails sur la transformée de Hough, nous invitons le lecteur à consulter l'article de Hough [Hough 62].

Enfin, les lignes sont regroupées selon les trois directions orthogonales X, Y et Z. Le regroupement « clustering » s'effectue par l'intermédiaire de l'algorithme MSAC [Torr 97]. Ce dernier permet de classer les lignes de Hough, en se basant sur le point de fuite généré. Les lignes de Hough n'ayant pas servi au calcul d'un point de fuite sont éliminés. La figure 5.4.5 montre un exemple de modèle ligne obtenu par le processus de représentation, décrit à la présente section.

5.4.6 Mesure de Similarité

Pour mesurer le degré de ressemblance entre l'image actuelle et une image référence, nous utilisons la mesure de Hausdorff [Huttenlocher 93]. La distance de Hausdorff indique à quel niveau un point de l'image modèle est spatialement proche d'un point de l'image test. En revanche, elle ne permet pas de déterminer les correspondances entre les points du modèle et les points de l'image. Effectivement, la distance de Hausdorff permet de calculer la ressemblance entre le modèle



FIGURE 5.4.5 – Construction de modèle de lignes d'un bâtiment

et l'image test superposés l'un sur l'autre. La distance de Hausdorff est une distance min-max définie comme suit. Etant donnée deux ensembles finis de points A et B tels que $A = \{a_1, \dots, a_n\}$ et $B = \{b_1, \dots, b_n\}$, la distance de Hausdorff est calculé par l'équation suivante :

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (5.4.16)$$

avec :

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (5.4.17)$$

La fonction $h(A, B)$ est appelée la distance de Hausdorff dirigée « directed » de A vers B . Elle mesure la distance de a au plus proche voisin dans B en utilisant la norme $\|\cdot\|$. Cette norme peut être par exemple la distance euclidienne. De la même manière, $h(B, A)$ est la distance dirigée de Hausdorff de B vers A . En effet, $h(A, B)$ classe chaque point de A en se basant sur sa distance au plus proche point appartenant à B , et donc utilise le point ayant le plus grand classement (the most mismatched point of A) pour calculer la distance de Hausdorff. Intuitivement, si $h(A, B) = d$ alors chaque point de A doit être situé à une certaine distance de chaque point de B , et il existe aussi un certain point de A qui est exactement distant d'une valeur d du plus proche point de B (the most mismatched point). La distance de Hausdorff $H(A, B)$ est le maximum entre $h(A, B)$ et $h(B, A)$. Ainsi, cette métrique mesure le degré de discordance ou dissemblance entre deux ensembles A et B .

5.4.7 Expérience et Résultats

Pour mesurer la performance de notre méthode de reconnaissance de bâtiments, nous avons comparé 10 images tests à 50 images modèles. Les images ont été prises de la base de données ZuBud décrite en détail dans [Shao 03]. Cette dernière se compose désormais de 201 bâtiments avec cinq images par bâtiment. Les images ont été acquises avec une des variations de points de vue, et avec différentes conditions météorologiques et d'éclairage. Notre méthode permet de trouver le bâtiment

recherché avec un pourcentage de succès estimé à 80%. Il est à remarquer que les 20% d'échec ont eu lieu notamment avec les images ayant de grandes variations de points de vue.

Conclusion

Se trouver en face d'un monument méconnu laisse naître le besoin de découvrir et d'appréhender ses différents aspects : historiques, civils, architecturaux, etc. Afin d'aider le touriste à appréhender l'histoire d'un immeuble patrimonial et ses différents aspects, nous proposons d'adapter notre système M.A.R.T.S pour une visite en extérieur.

A cet effet, nous avons commencé à réfléchir sur comment étendre nos propositions relatives à la visite au musée pour un fonctionnement à l'extérieur. Nous avons également pensé à faire évoluer le composant reconnaissance de notre système M.A.R.T.S en y intégrant une méthode de reconnaissance plus adaptée aux aspects géométriques des bâtiments. Ce chapitre constitue donc un début de travail sur la visite guidée à l'extérieur, sujet que nous envisageons de plus approfondir dans de futurs travaux.

Conclusion et perspectives

Rappel du Contexte

La culture et le tourisme, étroitement liés, sont de nature à renforcer l'attractivité et la compétitivité de régions et de pays. En 2010, les visites liées au patrimoine culturel ont atteint 47% de l'ensemble des activités pratiquées par les touristes du pays basque, et se classent ainsi au deuxième rang. Etant donné que le tourisme constitue un secteur essentiel pour l'économie française, les professionnels du tourisme et les musées ont engagé depuis quelques années une profonde réflexion pour répondre aux nouvelles attentes de leurs visiteurs à travers l'exploitation de plusieurs stratégies. En ce sens, mettre le numérique au service des visiteurs, représente une stratégie par laquelle on vise à renforcer l'attractivité des destinations.

Au fil de l'évolution technologique, la recherche de distinction et d'attractivité ont amené les musées à identifier, tester et intégrer de nouveaux outils dans l'aide à la visite, de l'étiquette en passant par l'audio-guide jusqu'aux systèmes informatiques. Grâce à ladite évolution technologique, la décennie écoulée a vu se « matérialiser » un concept, datant pourtant des années 60, qui risque fort de changer le mode de médiation au musée : la Réalité Augmentée.

Alors ? Nouvel outil miracle en perspective ? En quoi cela consiste-t-il ? Aurait-on enfin une visite attractive ? Quelle forme pourrait prendre un système d'aide à la visite intégrant ce nouveau concept ? Comment faire en sorte d'impliquer le visiteur dans une expérience riche d'apprentissage au musée ? C'est à ces questions auxquelles nous avons apporté des éléments de réponse tout au long de ce mémoire.

Synthèse du document

Nous avons commencé ce mémoire en définissant ce qu'est la réalité augmentée. Pour cela, nous avons retracé l'évolution de ses définitions et nous nous sommes rendu compte qu'elles sont toujours soumises à discussions et interprétations, bien que certaines fassent références. Nous avons alors proposé notre propre définition,

avec pour objectif d'établir une compréhension commune de ce concept lors de la lecture de ce document. Nous avons ensuite présenté les fondements technologiques qui constituent un système de réalité augmentée, cela afin d'en comprendre les rouages et donc les possibilités nous étant offertes dans la conception de notre propre système.

Nous avons consacré en partie le chapitre suivant à l'étude des attentes des visiteurs face au musée. Nous avons donc identifié le besoin d'apprendre et de découvrir comme étant la quête principale du visiteur. En ce sens, nous nous sommes intéressé à l'expérience d'apprentissage au musée tel que définie par le modèle contextuel de Falk [Falk 05]. Nous avons alors passé en revue les outils actuels de la médiation au musée. Il est apparu que ceux-ci sont jugés peu pertinents pour faire vivre au visiteur une expérience riche d'apprentissage. Nous nous sommes alors tournés vers les avantages des systèmes d'aide à la visite basés sur l'utilisation de la réalité augmentée. Les manques constatés nous ont alors permis d'établir notre problématique.

Ensuite, nous avons vu que les théoriciens de l'apprentissage humain [Bruner 60] classifient les modes de représentation de savoir en trois catégories : symbolique, iconique et éactif. Hooper- Greenhill [Hooper Greenhill 94] insiste sur l'importance de combiner ces trois modes au musée afin de se mettre à la portée des différents visiteurs. Il en est ressorti que les systèmes actuels sont peu attentifs à retranscrire correctement ces modes. C'est à la lumière de toutes ces indications, que nous en rendons compte dans nos propositions puisant leurs principes dans la réalité augmentée.

Enfin, nous avons présenté une implémentation de nos propositions sous la forme d'un système de visite guidée que nous avons nommé « M.A.R.T.S ». Durant son développement, nous avons été en mesure de tester l'influence de la simulation de coprésence d'un guide humain, en conjonction avec la désignation basée image « Sélection » et des informations textuelles « Documentation » sur l'expérience d'apprentissage au musée. Les premiers résultats montrent les avantages induits par la prise en compte de nos propositions. De plus, l'accueil positif des utilisateurs réservé à notre système M.A.R.T.S nous a conduit à émettre, dans le dernier chapitre de ce mémoire, nos réflexions sur comment étendre celui-ci à une visite guidée à l'extérieur.

Résumé de nos contributions et résultats

La question à laquelle nous avons voulu répondre durant nos travaux est partie intégrante de la formulation de la problématique que nous avons énoncée, à savoir :

« **Comment présenter les informations concernant l'oeuvre d'art de telle sorte à impliquer le visiteur dans une expérience d'apprentissage riche au musée ?** »

Pour cela, nous avons identifié et mis en avant les principaux aspects régissant la visite guidée au musée :

Simulation de la coprésence d'un guide humain : les théoriciens de l'apprentissage au musée [Hooper Greenhill 94] insistent sur l'importance de l'apprentissage « face à face » faisant intervenir les guides humains. Étant donné que les recherches en réalité virtuelle et en réalité augmentée [Gerhard 05][Luo 14], ont prouvé la capacité des humains virtuels à simuler la coprésence humaine, nous avons fait donc appel à ces derniers pour soutenir l'apprentissage au musée.

Les modes d'interaction : ils présentent les paradigmes d'interaction qui permettent d'appuyer le discours du guide humain virtuel et d'améliorer la visualisation des informations promulguées par ce discours.

Partant de ces aspects, nous avons alors proposé le fonctionnement général d'un système de visite guidée au musée intégrant ces propositions :

- **Sélection** : une interaction représentative du geste déictique.
- **Documentation** : une interaction renvoyant au mode symbolique de transmission du savoir.
- **Reconstruction** : une interaction renvoyant au mode iconique, elle utilise des animations 2D/3D adéquates afin de montrer à quoi ressemblait l'aspect original de l'oeuvre d'art.

Pour respecter les exigences qu'impose l'apprentissage libre au musée [Falk 00], nous avons traduit cela en contraintes :

1. Définir un parcours libre afin de tenir compte des propres choix du visiteur et de son intérêt pour tel ou tel oeuvre d'art (volet personnel du modèle contextuel de Falk [Falk 05]).
2. La non-immersivité de l'interface utilisateur afin de laisser le visiteur interagir avec ses compagnons et les autres visiteurs (volet social du modèle contextuel)

de Falk [Falk 05]).

3. Adapter l'interface utilisateur à l'environnement physique du musée afin de tenir compte des événements comme l'affluence (volet physique du modèle contextuel de Falk [Falk 05]).

En nous basant sur le principe de fonctionnement que nous venons d'énoncer, nous avons conçu et implémenté notre propre système que nous avons baptisé M.A.R.T.S, pour Mobile Augmented Reality Touring System. Le prototype matériel se base sur une vision déportée sur l'écran du Smartphone et un casque Audio Bone. Ceci est dans le but d'assurer la non-immersivité de notre système. Lors de l'élaboration de celui-ci, nous avons également présenté une méthode de reconnaissance 2D rapide permettant d'identifier d'une manière relativement précise les œuvres d'art. Nous avons également présenté une méthode simple pour la désignation basée image « Sélection ».

Durant nos travaux, nous avons été en mesure d'expérimenter le principe de la simulation de la coprésence d'un guide humain en conjonction avec le geste diététique « Sélection » et le paradigme « Documentation ». À travers des tests utilisateurs, nous avons constaté que l'application de ces principes pouvait faire gagner environ 9% en termes de rétention d'informations en comparaison avec l'Etiquette et 6% en comparaison avec l'Audio-guide. Mais au-delà de ce résultat, nous avons également pu observer de multiples bénéfices comme une amélioration qualitative de l'expérience d'apprentissage. D'un point de vue cognitif, nous avons remarqué une facilité d'acquisition de l'information ou encore une simple détection sur l'œuvre d'art des éléments référencés par les descriptions. D'un point de vue émotionnel, la visite a été considérée plus saisissante.

Ces premiers résultats tendent à montrer l'importance d'intégrer dans les systèmes de visite guidée les trois modes de transmission de savoir ainsi que les mécanismes de la communication non verbale (geste diététique), et nous encouragent donc à poursuivre dans cette voie.

Compte tenu des résultats promoteurs que nous avons obtenu avec M.A.R.T.S, nous avons commencé à réfléchir sur son extension pour un fonctionnement à l'extérieur. Pour ce faire, nous avons proposé une méthode précise pour la reconnaissance des bâtiments se basant sur les points de fuite. De plus, afin de ne pas perturber l'expérience d'apprentissage avec M.A.R.T.S nous avons proposé une solution pour une navigation transparente à l'extérieur, utilisant la modalité haptique.

Limitations et Perspectives

Durant le quatrième chapitre, nous avons pu voir que le développement du système M.A.R.T.S n'est pas entièrement finalisé. En effet, l'implémentation des principes de « Reconstruction » nécessite d'explorer plus profondément le domaine de la réalité augmentée qui s'intéresse au suivi robuste temps réel d'objets 3D. Comme nous l'avons indiqué, pour ce faire, nous nous sommes orientés vers une méthode de suivi basée sur une connaissance a priori du modèle (2D ou 3D). Pour cela nous avons commencé à développer notre propre méthode de suivi en nous inspirant des travaux de [Bottecchia 10].

À l'issue de ce développement, il sera alors possible d'expérimenter en totalité les principes de nos paradigmes d'interaction. De plus, comme nous l'avons vu au troisième chapitre de ce mémoire, adapter la visualisation de l'interaction à l'environnement physique s'avère primordial aux moments d'affluence. En ce sens, pour tenir compte de la distance de l'utilisateur par rapport à l'œuvre d'art, nous avons implémenté la partie qui traite de l'ajout d'un texte virtuel. Cependant, nous allons continuer à travailler sur la transformation de l'interface utilisateur. En fait, dans le cas d'occultation partielle ou totale de l'œuvre d'art, on doit passer d'un affichage direct de nos paradigmes d'interaction (mode de réalité augmentée) à un affichage indirect (animations mettant en œuvre ces paradigmes).

Les premiers résultats obtenus durant notre expérimentation montrent que notre système pourrait parfaitement améliorer l'expérience d'apprentissage au musée, et cela, grâce à nos propositions puisant leurs principes de la réalité augmentée. Cependant, bien que nos résultats expriment un potentiel favorable en termes d'assistance à la visite du musée, nous devons poursuivre nos expérimentations sur un échantillon de tests plus important pour que les tendances, parfois fortes, soient statistiquement significatives.

Par ailleurs, compte tenu des résultats favorables que nous avons eu avec M.A.R.T.S au musée, nous allons continuer le développement de celui-ci dans le cadre d'un futur projet de collaboration avec le musée basque et de l'histoire de Bayonne. De plus, nous avons commencé à réfléchir sur comment étendre M.A.R.T.S pour une visite guidée à l'extérieur. Pour ce faire, nous avons travaillé au cinquième chapitre sur la reconnaissance des bâtiments, partie majeure du patrimoine historique et architecturale. Dans cette optique, nous avons proposé notre méthode de reconnaissance des bâtiments basée sur les points de fuite. Testée avec la base de données ZuBud [Shao 03], notre méthode permet de trouver le bâtiment recherché avec un pourcentage de succès estimé à 80%. Avec la même base de données, la méthode de Trinh et al. [Trinh 07] effectue un taux de performance de 94.78%.

En revanche, contrairement à notre méthode (calcule trois points de fuite), leur méthode se base sur le calcul de plusieurs points de fuite (jusqu'à six points de fuite). Afin de mettre en avant le compromis temps de calcul/précision assurée par notre méthode, nous allons la comparer dans de futurs travaux à la méthode de Trinh et al. [Trinh 07].

Annexes

« The world isn't interested in
the storms you encountered,
but whether or not you
bought in the ship. »"

(Raul Armesto)

Annexe A

Fichier KML

KML (Keyhole Markup Language) est un format de fichier et de grammaire XML pour la modélisation et le stockage de caractéristiques géographiques comme les points, les lignes, les images, les polygones et les modèles pour l’affichage dans Google Earth, dans Google Maps et dans d’autres applications.

Un fichier KML est traité par Google Earth de la même manière que les fichiers HTML et XML sont traités par les navigateurs Web. Comme un fichier HTML, un fichier KML a une structure basée sur des balises avec des noms et des attributs utilisés pour un affichage spécifique.

La figure suivante montre un exemple d’un fichier KML, décrivant un itinéraire entre le quai des corsaires et la rue des prébendes. Nous avons simplifié le fichier, de manière à le rendre plus lisible.

La balise englobante du fichier est `< kml >` . Les balises `< Placemark >` traduisent les points repères de l’itinéraire. La position et l’orientation du point repère est décrite par les balises : `< longitude >`, `< latitude >`, `< tilt >` et `< heading >` .

```
<?xml version="1.0" encoding="UTF-8"?>
- <kml xmlns="http://earth.google.com/kml/2.0">
  - <Document>
    <name>Quai des Corsaires to Rue des Prébendes</name>
    - <Style id="roadStyle">
      - <LineStyle>
        <color>7fcf0064</color>
        <width>6</width>
      </LineStyle>
    </Style>
    <Snippet/>
    - <Placemark>
      <name>Turn right onto Pont Marengo</name>
      - <description>
        <![CDATA[go 63&#160;m]]>
      </description>
      <StyleMap/>
      - <Point>
        <coordinates>-1.473940,43.491020,0</coordinates>
      </Point>
      - <LookAt>
        <longitude>-1.473940</longitude>
        <latitude>43.491020</latitude>
        <range>100.000000</range>
        <tilt>45.000000</tilt>
        <heading>198.703964</heading>
      </LookAt>
    </Placemark>
  </Document>
</kml>
```

FIGURE A.1 – Structure d'un fichier KML

Annexe B

Projet Aquitaine Euskadi

Titre du projet

Applications de réalité augmentée mobile pour le tourisme : animation multi-média et visite guidée par un agent conversationnel virtuel.

Partenaire chef de file du projet

- **Nom de l'organisme** ESTIA
- **Statut** Etablissement d'Enseignement Supérieur et de Recherche
- **Adresse** Technopole Izarbel, 64210 BIDART
- **Téléphone** +(33)559438400
- **Fax** +(33)559438405
- **E-mail** j.guiresse@estia.fr
- **Nom et qualité du représentant et responsable légal (juridique)**
Jean-Roch Guiresse, Directeur
- **Nom et qualité du représentant et responsable financier** Juhane Dascon, service Direction-Gestion
- **Nom et qualité du responsable du projet** Jean-Marc Cieutat, Enseignant-Chercheur

Autre(s) Partenaire(s)

- **Nom de l'organisme** Fundación Centro de Tecnologías de Interacción Visual y Comunicaciones Vicomtech
- **Statut** Centre de Recherche Appliquée

- **Adresse** Mikeletegi Pasealekua 57, Parque Tecnológico 20009 Donostia-San Sebastián
- **Téléphone** +(34)943309230
- **Fax** +(34)943309393
- **E-mail** jflorez@vicomtech.org
- **Nom et qualité du représentant et responsable légal (juridique)**
Julián Flórez, Directeur
- **Nom et qualité du représentant et responsable financier** Eburne Loyarte, Directeur Financier
- **Nom et qualité du responsable du projet** Jairo R. Sánchez, Chercheur

Resumé du projet

Le projet porte sur la mutualisation des compétences en réalité augmentée de l'ESTIA et de VICOMTECH pour concevoir un environnement commun de développement d'applications de réalité augmentée mobile dédiées au tourisme. L'environnement doit permettre, d'une part, aux offices de tourisme de définir des itinéraires touristiques comportant des sites touristiques à découvrir et d'y associer des contenus numériques et permettre, d'autre part, à un touriste d'accéder à ces contenus en fonction de ses choix et de sa géolocalisation. Plusieurs niveaux de fonctionnalités sont alors prévus comme l'activation de contenus multimédia (commentaires audio d'un site touristique dans la langue maternelle du visiteur, ...), l'ajout d'indices virtuels (texte, annotation, 3D, ...) dans un flux vidéo, et enfin la visite guidée d'un site touristique par un agent conversationnel virtuel. Le dernier niveau de 5 fonctionnalités a un caractère plus exploratoire mais offre de nombreuses perspectives, proposant une visite plus interactive et personnalisée.

Objectifs du projet

Le premier objectif du projet est de mettre à disposition des professionnels du tourisme un nouvel outil capable d'aider un touriste à avoir une meilleure connaissance des lieux qui l'entourent et de la culture locale, et de lui procurer les informations utiles à son séjour. Le tourisme est un secteur d'activité très important de notre économie.

Pour y parvenir, l'ESTIA et VICOMTECH mettront leurs compétences en commun pour concevoir un environnement de développement d'applications de réalité augmentée mobiles. Trois niveaux de fonctionnalités sont prévus qui présentent un degré d'innovation toujours plus élevé.

Pour tester l'outil, nous travaillerons en collaboration étroite avec les offices de tourisme respectifs des villes de San-Sébastien et de Biarritz. Les premières illustrations porteront sur les sites touristiques et les infrastructures de ces deux villes, à la fois prestigieuses et résolument tournées vers le tourisme. Les contenus multimédia seront respectivement traduits en français à l'intention des touristes français qui visiteront San-Sébastien et en espagnol à l'intention des touristes espagnols qui visiteront Biarritz. Il s'agit aussi de traduire les éléments d'une culture à une autre. Des protocoles de tests seront d'ailleurs mis en place pour tester la pertinence des outils proposés et faire évoluer leur qualité.

À l'issue du programme, nous prévoyons des communications scientifiques nationales, dans chaque pays, sur les résultats scientifiques et technologiques relatifs à l'environnement proposé, et dans au moins une conférence internationale où un article commun pourra être présenté. Mais le projet traite d'un sujet qui avant tout est très médiatique. Dans ces conditions, et pour faire connaître l'outil, nous ciblerons les médias de communication traditionnels (journaux, télévision, internet).

Zone concernée (localisation du projet)

La CCI de Bayonne Pays-Basque anime le cluster tourisme « Goazen » qui réunit l'ensemble des professionnels du tourisme du pays-basque français. Cela montre l'importance du sujet comme étant très stratégique à notre région, et c'est une chance supplémentaire pour l'ESTIA de pouvoir dans ce cadre collaborer directement avec les différentes branches du secteur du tourisme.

Les liens sont bien sûr très forts entre le pays-basque français et le pays basque espagnol pour lequel le tourisme est aussi un secteur économique très stratégique. Faire connaître la culture basque et les patrimoines respectifs sont des éléments de motivation du projet.

Le choix des sites pilote, en les villes de Biarritz et de San-Sébastien, a été bien sûr fait en tenant compte des implantations respectives de l'ESTIA et de VICOMTECH.

Caractère interregional du projet

Il se trouve que l'ESTIA et VICOMTECH ont respectivement des compétences scientifiques et technologiques reconnues en réalité virtuelle et réalité augmentée qui seront fort utiles au projet présenté comme étant ambitieux. La collaboration entre les deux partenaires au projet serait un moyen de faire émerger un pôle d'excellence sur ces thèmes au niveau européen.

La constitution de clusters est un moyen de regrouper les professionnels d'un secteur et de les faire travailler conjointement avec les établissements de recherche. Ce projet en est la possibilité.

Dates du projet

Le projet a débuté le 1er janvier 2013 pour se terminer le 31 décembre 2014.

Annexe C

SDK de Réalité Augmentée Mobile

Dans le cadre du projet Aquitaine-Euskadi, nous avons été amenés à concevoir un SDK de RA mobile. En effet, la technologie de la réalité augmentée a vu le jour depuis plus d'une décennie, mais c'est avant tout l'essor de l'informatique mobile qui aura fait connaître la RA auprès du grand public. L'intérêt de ce nouveau marché n'aura pas échappé aux industriels, avec l'apparition de plus en plus de kits de développements d'applications de RA ces derniers mois, qui viennent compléter les applications déjà connus comme Metaio, Layar, Wikitude ou encore Vuforia ; à ce propos le lecteur pourra trouver un banc comparatif des kits existants à l'adresse : <http://socialcompare.com/fr/comparison/augmented-reality-SDKs>.

Toutefois, force est de constater que les solutions proposées sont toutes très similaires, pour les avoir testées et comparées. Les augmentations sont exclusivement visuelles, pour des applications au final peu interactives. Les solutions proposées reposent sur le même modèle technique, à savoir une architecture client-serveur où l'éditeur propose des conditions tarifaires étroitement dépendantes du nombre de connexions faites au serveur. Le calcul de la pose de la caméra s'effectue au moyen de la même technologie, à savoir la détection et la comparaison de la distribution de points invariants dans une image (SIFT, SURF, ...), parfois complété par la reconnaissance de QRCode.

En tenant compte de l'explosion du marché des Smartphones, il est judicieux de penser qu'une méthodologie de conception et un ensemble d'outils permettant de créer rapidement des applications de RAM (Réalité Augmentée Mobile) est chose utile, afin de faciliter le développement d'applications de RA. Dans cette optique, nous proposons notre SDK de réalité augmentée mobile, abrégé SDK MAR. Nous avons choisi de respecter la spécification standard ARML 2.0 [ARML 12], dans le but de proposer un SDK standard et flexible de développement d'application de

RAM.

Architecture logicielle

Architecture Globale

Afin de mettre en place l'architecture logicielle du SDK MAR, nous nous sommes inspirés de la constitution logicielle d'un système de RA, décrite par Bimber et al [Bimber 05]. Celle-ci nous a permise d'identifier les différents composants logiciels indispensables pour le SDK MAR.

L'architecture logicielle du SDK MAR se base sur six composants principaux, à savoir : la caméra, la base de données, la localisation, la reconnaissance, le suivi et le rendu. La figure C.1 met en relief la configuration globale du SDK MAR. En effet, il s'agit d'une architecture orientée composant. Chaque composant est chargé de fonctions spécifiques.

Dépendances entre composants

Le diagramme UML, schématisé par la figure C.2, montre les principales dépendances entre les six composants constituant le SDK MAR. Bien évidemment, les composants Rendu et Suivi dépendent du composant Camera. En plus, le composant Reconnaissance dépend à la fois des modules Camera, Localisation et Base de données. En outre, le composant Suivi dépend des modules Localisation et Base de données. Dans un but de clarté, ces deux derniers liens n'ont pas été schématisés dans la figure C.2.

Dans ce qui suit, nous allons détailler les différents composants logiciels et les liens associés. Nous allons également préciser, pour chaque module, les appartenances des classes du modèle d'objet.

Composant Reconnaissance : ce composant fournit un ensemble de méthodes pour effectuer la reconnaissance d'objets dans une scène réelle. Il permet aussi de s'affranchir de marqueurs type QRCode puisque la reconnaissance s'effectue sur des images prétraitées contenues dans la base de données. Deux méthodes de reconnaissance sont mises en place, la première en 2D est basée sur les points caractéristiques SURF, permettant ainsi de reconnaître des objets tels que les œuvres d'art au musée (notamment les objets texturés). La deuxième se base sur les lignes caractéristiques. Elle est essentiellement adaptée à la reconnaissance des bâtiments. Ces deux méthodes sont détaillées d'une manière extensive aux chapitres 4 et 5 de

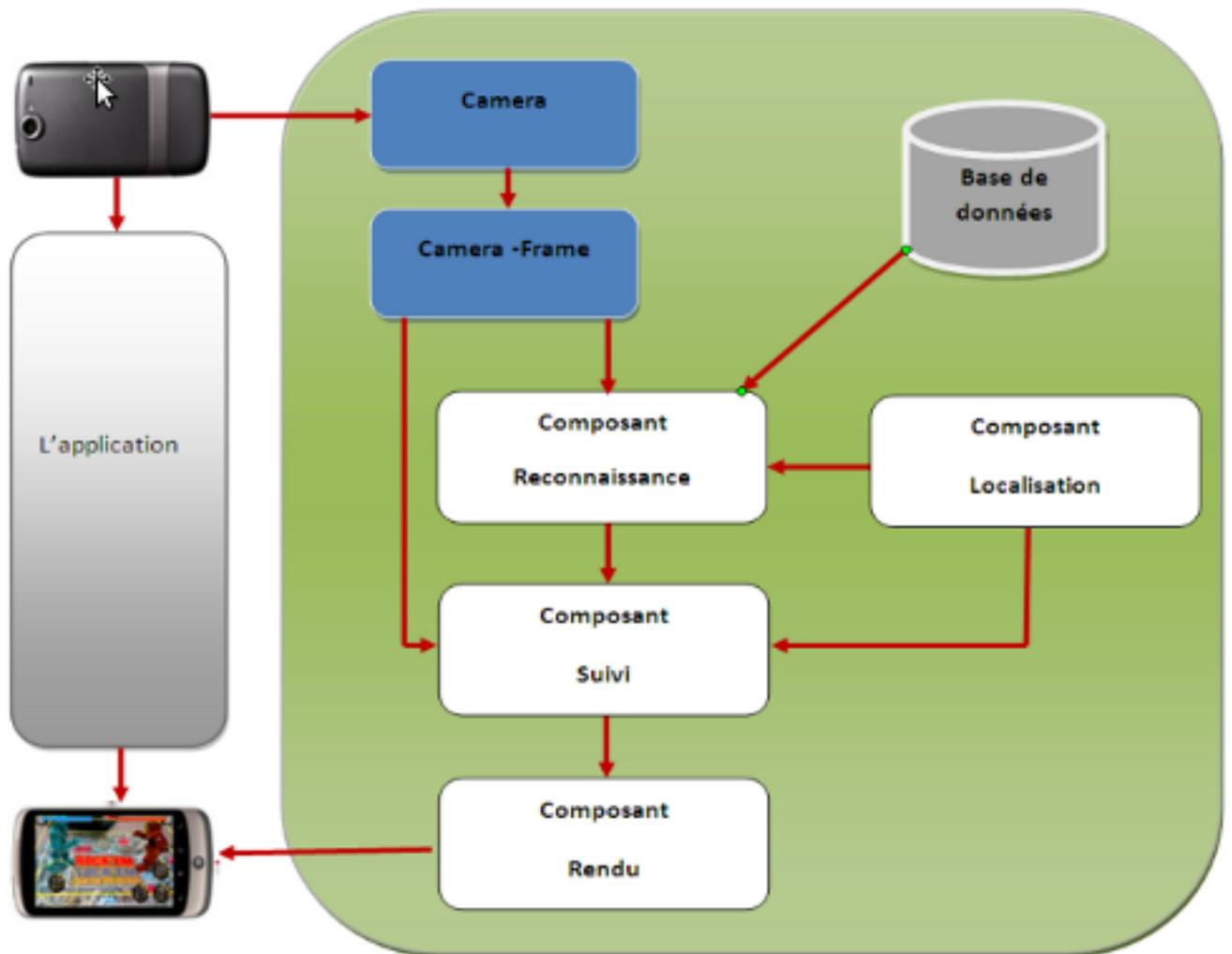


FIGURE C.1 – . Architecture logicielle globale du SDK MAR

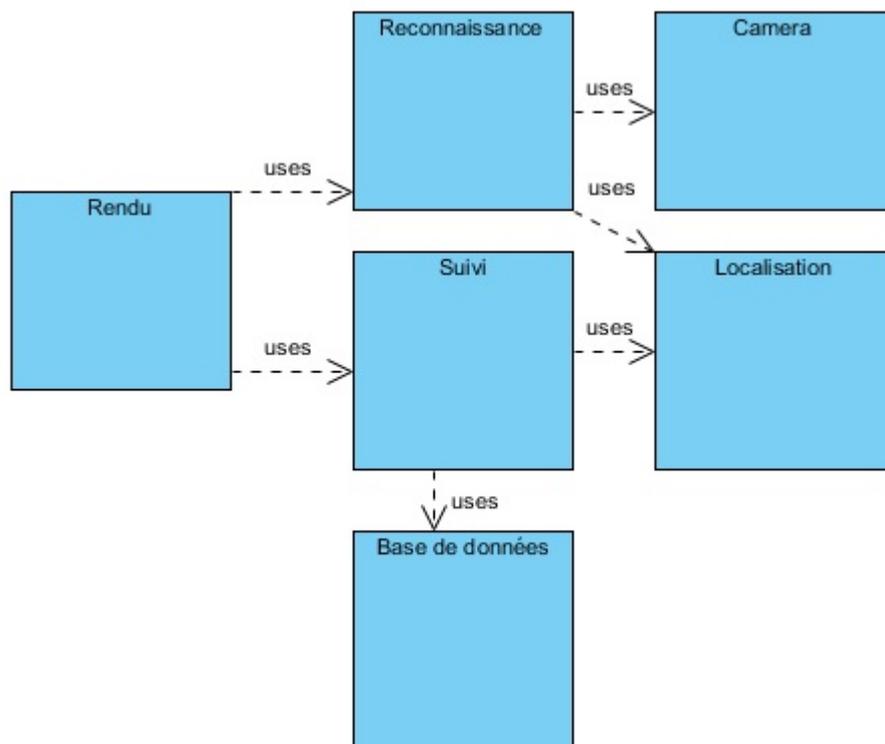


FIGURE C.2 – Liens entre les composants du SDK

ce mémoire. Il est à noter que le composant Reconnaissance dépend principalement de la classe Feature et de la classe Trackable d'ARML.

Composant Rendu : pour le rendu graphique, ce composant utilise l'API Ogre 3D [Ogre 00] qui est un moteur 3D supportant OpenGL ES, version mobile d'OpenGL. Cette librairie fournit un accès simplifié aux paramètres habituels d'une scène virtuelle comme la camera, les lumières ou les objets 3D. D'un autre côté, ce composant fait appel à notre bibliothèque mobile d'agents conversationnels 3D. Cette bibliothèque est décrite en détail dans le chapitre 4 de ce mémoire. En plus du rendu graphique, ce composant assure d'autres types de retour tels que le retour haptique et le retour audio. Nous précisons au lecteur que ce composant contient essentiellement les classes suivantes d'ARML : Asset, Plasticity, VisualAsset, AudioAsset, Model, visualAsset2D.

Composant Base de données : le composant Base de données permet à l'utilisateur de créer une ou plusieurs bases de données. La base de données permet de stocker les images représentant les points d'intérêts (bâtiments, statues) ou les œuvres d'art. Ces images représentent des images de référence utilisées lors de la reconnaissance ou des marqueurs utilisés pour le suivi. Ce composant est essentiellement dépendant de la classe Feature et de la classe Trackable d'ARML.

Composant Camera : le composant Camera s'occupe de fournir les services de gestion vidéo et du contrôle de la caméra. Ce composant contient les classes qui permettent de contrôler la caméra et les classes qui permettant d'obtenir les images de cette même caméra.

Composant Suivi : le composant Suivi met en œuvre les algorithmes de vision par ordinateur permettant de détecter et de suivre les objets du monde réel. Deux types d'algorithme ont été implémentés dans ce module. Le premier s'appelle PTAM, il se base sur les points caractéristiques (features points). PTAM est détaillé dans le premier chapitre de ce mémoire. Le second se base sur les points de fuite, dont la détection est détaillée dans le cinquième chapitre de ce mémoire. Il est à noter que le composant Suivi, inclut principalement les classes suivantes : Tracker, Trackable, TrackerConfig.

Composant Localisation : le rôle du composant de localisation consiste à localiser l'utilisateur. Ce composant utilise les capteurs intégrés au Smartphone tel que le GPS et la centrale inertielle. Grâce au GPS, un filtrage est effectué au niveau du module de reconnaissance : inutile de chercher un monument de Bordeaux alors que l'on se trouve à Biarritz. Nous précisons que le composant Localisation dépend

essentiellement des classes suivantes : ARAnchor, Anchor, Geometry, GMLGeometry.

Remarque

L'architecture logicielle proposée permet de garantir une certaine extensibilité pour le SDK MAR. Effectivement, chaque composant peut inclure différents méthodes pour la mise en œuvre d'une fonction donnée.

Bibliographie

- [Agarwal 05] A. Agarwal, C. V. Jawahar & P. J. Narayanan. *A Survey of Planar Homography Estimation Techniques*. 2005.
- [Ahriz Roula 10] Iness Ahriz Roula. *Application des techniques d'apprentissage à la géolocalisation par radio fingerprint*. These de doctorat, 2010.
- [Alessi 00] Norman E. Alessi & Milton P. Huang. *Evolution of the Virtual Human : From Term to Potential Application in Psychiatry*. Cyberpsy., Behavior, and Soc. Networking, vol. 3, 2000.
- [Alibert 06] David Alibert, Régis Bigot & Georges Hatchuel. *Aller au musée : un vrai plaisir, mais une attente de convivialité et de pédagogie*. Centre de Recherche pour l'Etude et l'Observation des Conditions de Vie, 2006.
- [Amarino 10] Amarino. <http://www.amarino-toolkit.net/>. 2010.
- [Anastassova 07] M. Anastassova, J.-M. Burkhardt, C. Megard & P. Ehanno. *L'ergonomie de la réalité augmentée pour l'apprentissage : une revue*. Le travail humain, vol. 70, 2007.
- [Anderson 83] John Robert Anderson. *The Architecture of Cognition*. Harvard University Press, 1983.
- [Antone 00] Matthew E. Antone & Seth J. Teller. *Automatic Recovery of Relative Camera Rotations for Urban Scenes*. CVPR, 2000.
- [Arai 07] Kohei Arai & Ali Ridho Barakbah. *Hierarchical K-means : an algorithm for centroids initialization for K-means*. Reports of the Faculty of Science and Engineering, Saga University, vol. 36, 2007.

- [ArduinoBT 11] ArduinoBT. <http://arduino.cc/en/Main/ArduinoBoardBT?from=Main.ArduinoBoardBluetooth/>. 2011.
- [ARML 12] ARML. <http://openarml.org/wikitude4.html>. 2012.
- [Augereau 13] Olivier Augereau, Nicholas Journet & Jean-Philippe Domenger. *Semi-structured document image matching and recognition*. Document Recognition and Retrieval, 2013.
- [Azuma 97] Ronald T. Azuma. *A survey of augmented reality*. Presence : Teleoperators and Virtual Environments, vol. 6, 1997.
- [Azuma 01] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier & Blair MacIntyre. *Recent Advances in Augmented Reality*. IEEE Comput. Graph. Appl., vol. 21, 2001.
- [Bachmann 99] E. R. Bachmann, I. Duman, U. Y. Usta, R. B. McGhee, X. P. Yun & M. J. Zyda. *Orientation tracking for humans and robots using inertial sensors*. Computational Intelligence in Robotics and Automation, 1999. CIRA '99. Proceedings. IEEE International Symposium on, 1999.
- [Barnard 83] Stephen T. Barnard. *Interpreting perspective images*. Artif. Intell., vol. 21, 1983.
- [Bationo Tillon 10] Anne Bationo Tillon, Eric Marchand, Jean Laneurit, Fabien Servant, Isabelle Marchal & Pascal Houlier. *A day at the museum : An augmented fine-art exhibit*. IEEE International Symposium on Mixed and Augmented Reality 2010 Arts, Media, and Humanities Proceedings, vol. 9, 2010.
- [Bationo Tillon 11] Anne Bationo Tillon, Isabelle Marchal & Pascal Houlier. *Mobile augmented reality in the museum : Can a lace-like technology take you closer to works of art?* IEEE International Symposium On Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH)., 2011.
- [Bay 08] Herbert Bay, Andreas Ess, Tinne Tuytelaars & Luc Van Gool. *Speeded-Up Robust Features (SURF)*. Comput. Vis. Image Underst., vol. 110, 2008.

- [Bayart 07] Benjamin Bayart. *Realité augmentée haptique : théorie et applications*. Robotique, Université d'Evry, 2007.
- [Beaty 97] Liz Beaty, Graham Gibbs & Alistair Morgan. *Learning orientations and study contracts*. The experience of learning, 1997.
- [Bimber 05] Oliver Bimber & Ramesh Raskar. *Spatial Augmented Reality : Merging Real and Virtual Worlds*. A. K. Peters, Ltd., 2005.
- [Bolt 80] Richard A. Bolt. *Put-that-there : Voice and gesture at the graphics interface*. SIGGRAPH Comput. Graph., vol. 14, 1980.
- [Bottecchia 10] Sébastien Bottecchia. *Système T.A.C. : Télé-Assistance Collaborative. Réalité augmentée et NTIC au service des opérateurs et des experts dans le cadre d'une tâche de maintenance industrielle supervisée*. thèse de l'université de Toulouse, 2010.
- [Bourgeon-Renault 07] Dominique Bourgeon-Renault, Anne Gombault, Marine Le Gall-Ely, Christine Petr & Caroline Urbain. *Gratuité des musées et valeur perçue par les publics*. La lettre de l'OCIM, 2007.
- [Brangier 03] Eric Brangier & Javier Barcenilla. *Concevoir un produit facile à utiliser : adapter les technologies à l'homme*. Editions d'Organisation, 2003.
- [Brougère 12] Gilles Brougère. *Pratiques touristiques et apprentissages*. Mondes du tourisme, vol. 5, 2012.
- [Brown 03] Myron Z Brown, Darius Burschka & Gregory D Hager. *Advances in Computational Stereo*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, 2003.
- [Bruner 60] Jerome S Bruner. *The Process of Education*. Harvard University Press, 1960.
- [Burigana 13] Paola Burigana & Michaël Caucat. *Enquête sur les attentes des visiteurs au jardin botanique de Montréal*. La lettre de l'OCIM, 2013.
- [Calvary 04] Gaele Calvary, Alexandre Demeure, Joelle Coutaz & Olfa Daassi. *Adaptation des Interfaces Homme-Machine à leur contexte d'usage Plasticité des IHM*. Revue d'Intelligence Artificielle, vol. 18, 2004.

- [Canny 86] John Canny. *A Computational Approach to Edge Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, 1986.
- [Cantoni 01] Virginio Cantoni, Luca Lombardi, Marco Porta & Nicolas Sicard. *Vanishing point detection : Representation analysis and new approaches*. Proceedings of International Conference on Image Processing, 2001.
- [Caprile 90] B. Caprile & V. Torre. *Using vanishing points for camera calibration*. Int. J. Comput. Vision, vol. 4, 1990.
- [Casanueva 01] Juan S Casanueva & Edwin H Blake. *The Effects of Avatars on Co-presence in a Collaborative Virtual Environment*. Annual Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT2001). Pretoria, South, 2001.
- [Caudell 92] Thomas P. Caudell & David W. Mizell. *Augmented reality : an application of heads-up display technology to manual manufacturing processes*. System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on, vol. ii, 1992.
- [Cavallo 93] Pina Boggi Cavallo & Antonio Iannaccone. *Représentations sociales et construction des connaissances*. Papers on Social Representations, vol. 2, 1993.
- [Center 86] NASA Ames Research Center. *NASA TASK LOAD INDEX (TLX) v.1.0*. 1986.
- [Chamberland 91] Estelle Chamberland. *Les thèmes de la contextualisation chez les visiteurs de musée*. Canadian Journal of Education, vol. 21, 1991.
- [Chen 12] Jiajian Chen, Greg Turk & Blair MacIntyre. *A non-photorealistic rendering framework with temporal coherence for augmented reality*. ISMAR, 2012.
- [Clair 11] Aaron St Clair, Ross Mead & Maja J Mataric. *Investigating the effects of visual saliency on deictic gesture production by a humanoid robot*. RO-MAN, IEEE, 2011.
- [CNT 10] Conseil national du tourisme CNT. *Le poids économique et social du tourisme*. 2010.
- [Cohen 79] Erik Cohen. *A Phenomenology of Tourist Experiences*. Sociology, vol. 13, 1979.

-
- [Collins 90] Robert T. Collins & Richard S. Weiss. *Vanishing point calculation as a statistical inference on the unit sphere*. ICCV, 1990.
- [Cooper 95] Alan Cooper. *The myth of metaphor*. Visual Basic Programmer's Journal, 1995.
- [Daengbuppha 09] Jaruwan Daengbuppha. *Modelling visitor experience : a case study from World Heritage Sites*. PhD Thesis, Bournemouth University, 2009.
- [Damala 08] Areti Damala, Pierre Cubaud, Anne Bationo, Pascal Houlier & Isabelle Marchal. *Bridging the gap between the digital and the physical : design and evaluation of a mobile augmented reality guide for the museum visit*. Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts, 2008.
- [del Puy Carretero 12] María del Puy Carretero, Aitor Ardanza, Sara García, Helen Díez, David Oyarzun & Nuria Ruiz. *Improving Gestural Communication in Virtual Characters*. Proceedings of the 7th international conference on Articulated Motion and Deformable Objects, vol. 7, 2012.
- [DeMenthon 95] Daniel F. DeMenthon & Larry S. Davis. *Model-Based Object Pose in 25 Lines of Code*. International Journal of Computer Vision, vol. 15, 1995.
- [Dempster 77] A. P. Dempster, N. M. Laird & D. B. Rubin. *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society, Series B, vol. 39, 1977.
- [Deshayes 02] Sophie Deshayes. *Audioguides et Musées*. La lettre de l'OCIM, 2002.
- [D'Souza 99] Desmond Francis D'Souza & Alan Cameron Wills. *Objects, components, and frameworks with UML : the catalysis approach*. Addison-Wesley Professional, 1999.
- [Duch 98] Amalia Duch, Vladimir Estivill-Castro & Conrado Martinez. *Randomized K-Dimensional Binary Search Trees*. Proceedings of the 9th International Symposium on Algorithms and Computation, 1998.
- [Dufresne-Tassé 96] Colette Dufresne-Tassé. *Le questionnement de l'adulte au musée et les obstacles à sa progression*. Revue canadienne de l'éducation, vol. 21, 1996.

- [Dünser 08] Andreas Dünser, Raphaël Grasset & Mark Billinghurst. *A survey of evaluation techniques used in augmented reality studies*. ACM SIGGRAPH ASIA 2008 courses, 2008.
- [Elloumi 12] Wael Elloumi, Sylvie Treuillet & Rémy Leconge. *Tracking Orthogonal Vanishing Points in Video Sequences for a Reliable Camera Orientation in Manhattan World*. 5th International Congress on Image and Signal Processing (CISP 2012), 2012.
- [Falk 92] J.H. Falk & L.D. Dierking. *The museum experience*. Whalesback Books, 1992.
- [Falk 98] John H Falk, Theano Moussouri & Douglas Coulson. *The Effect of Visitors Agendas on Museum Learning*. Curator : The Museum Journal, vol. 41, 1998.
- [Falk 00] John Howard Falk & Lynn Diane Dierking. *Learning from Museums : Visitor Experiences and the Making of Meaning*. AltaMira Press, 2000.
- [Falk 05] John Falk & Martin Storcksdieck. *Using the contextual model of learning to understand visitor learning from a science center exhibition*. Science Education, vol. 89, 2005.
- [Fischler 81] Martin A. Fischler & Robert C. Bolles. *Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography*. Commun. ACM, vol. 24, 1981.
- [Friedman 77] Jerome H Friedman, Jon Louis Bentley & Raphael Ari Finkel. *An Algorithm for Finding Best Matches in Logarithmic Expected Time*. ACM Transactions on Mathematical Software (TOMS), vol. 3, 1977.
- [Fuchs 01] Philippe Fuchs, Guillaume Moreau & Jean Paul Papin. *Le Traité de la Réalité Virtuelle*. Presses de l'Ecole des Mines, 2001.
- [Fuchs 07] Jenni Fuchs. *Generic learning outcomes as a strategic tool for evaluating learning impact*. TICOM-CECA Conference, Vienna, 2007.
- [Fuchs 09] Phillippe Fuchs, Guillaume Moreau & Stéphane Donikian. *Les humains virtuels*. Le traité de la réalité virtuelle, vol. 5, 2009.

-
- [Fukiage 12] Taiki Fukiage, Takeshi Oishi & Katsushi Ikeuchi. *Reduction of contradictory partial occlusion in mixed reality by using characteristics of transparency perception*. ISMAR, 2012.
- [Gammon 03] Ben Gammon. *Assessing learning in museum environment A practical guide for museum evaluators*. London : science museum report, 2003.
- [Gerhard 05] Michael Gerhard, David Moore & Dave Hobbs. *Close Encounters of the Virtual Kind : Agents Simulating Co-presence*. Applied Artificial Intelligence : An International Journal, vol. 19, 2005.
- [Ghouaiel 13a] Nehla Ghouaiel, Jean-Marc Cieutat & Jean-Pierre Jessel. *Haptic System for Eyes Free and Hands Free Pedestrian Navigation*. The Sixth International Conference on Advances in Computer-Human Interactions (ACHI), vol. 6, 2013.
- [Ghouaiel 13b] Nehla Ghouaiel, Jean-Marc Cieutat & Jean-Pierre Jessel. *Mobile Augmented Reality Applications to Discover New Environments*. Science and Information (SAI) Conference, vol. 1, 2013.
- [Goodman 87] Nelson Goodman. *Of Mind and Other Matters*. Harvard University Press, 1987.
- [Graham 13] Jo Graham. *Evidencing the impact of the GLOs 2008 – 2013*. The University of Leicester’s School of Museum Studies, report, 2013.
- [Grauman 11] Kristen Grauman & Bastian Leibe. *Visual Object Recognition*. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2011.
- [Gustafsson 02] Fredrik Gustafsson, Fredrik Gunnarsson, Niclas Bergman, Urban Forssell, Jonas Jansson, Rickard Karlsson & Per-Johan Nordlund. *Particle filters for positioning, navigation, and tracking*. IEEE Transactions on Signal Processing, vol. 50, 2002.
- [Hager 95] Gregory D Hager, Wen Chung Chang & A S Morse. *Robot Hand-Eye Coordination Based on Stereo Vision*. IEEE Control Systems Magazine, vol. 15, 1995.
- [Haid 04] Markus Haid & Jan Breitenbach. *Low cost inertial orientation tracking with Kalman filter*. Applied Mathematics and Computation, vol. 153, 2004.

- [Haralick 94] Robert M Haralick, Chung-Nan Lee, Karsten Ottenberg & Michael Nölle. *Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem*. International Journal of Computer Vision, vol. 13, 1994.
- [Harris 88] C. Harris & M. Stephens. *A Combined Corner and Edge Detector*. Proceedings of the 4th Alvey Vision Conference, 1988.
- [Harris 90] Chris Harris & Carl Stennet. *RAPiD A video-rate object tracker*. British Machine Vision Conference, 1990.
- [Hein 98] George E. Hein. *Learning in the Museum*. Routledge, 1998.
- [Hollerer 04] Tobias H. Hollerer & K. Feiner Steven. *Mobile Augmented Reality . Telegeoinformatics : Location-Based Computing and Services*, 2004.
- [Holzmann 12] Clemens Holzmann & Matthias Hochgatterer. *Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision*. ICDCS Workshops, 2012.
- [Hooper Greenhill 94] Eilean Hooper Greenhill. *Museums and Their Visitors*. Routledge. New York, 1994.
- [Hooper-Greenhill 03] Eilean Hooper-Greenhill, Jocelyn Dodd, Theano Mousouri, Ceri Jones, Chris Pickford, Catherine Herman, Marlene Morrison, John Vincent & Richard Toon. *Measuring the Outcomes and Impact of Learning in Museums, archives and Libraries*. Research Centre for Museums and Galleries (RCMG), University of Leicester, 2003.
- [Hough 62] P. V. C. Hough. *A Method and Means for Recognizing Complex Patterns*. US Patent, 1962.
- [Hu 62] Ming-Kuei Hu. *Visual pattern recognition by moments invariants*. IEEE Transactions on Information Theory, vol. 8, 1962.
- [Huber 81] Peter J. Huber. *Robust statistics*. Wiley New York, 1981.
- [Hugues 11] Olivier Hugues. *Réalité augmentée pour l' aide à la navigation maritime*. These de doctorat, 2011.
- [Huttenlocher 93] Daniel P. Huttenlocher, Gregory A. Klanderman & William J. Rucklidge. *Comparing Images Using the*

-
- Hausdorff Distance*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, 1993.
- [Indyk 98] Piotr Indyk & Rajeev Motwani. *Approximate Nearest Neighbors : Towards Removing the Curse of Dimensionality*. Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing, 1998.
- [Inkila 05] Keijo Inkila. *Homogenous least squares problem*. The photogrammetric Journal of Finland, vol. 19, 2005.
- [Interlando 11] J.Carmelo Interlando & Cynthia Padilla. *Decoding the (41,20,10) quadratic residue code beyond its error-correcting capability*. Applied Mathematical Sciences (Ruse), vol. 5, 2011.
- [Jodelet 06] Denise Jodelet. *Place de l'expérience vécue dans le processus de formation des représentations sociales*. Les Presses universitaires de Rennes, 2006.
- [Juan 09] Luo Juan & Oubong Gwon. *A Comparison of SIFT, PCA-SIFT and SURF*. International Journal of Image Processing (IJIP), vol. 3, 2009.
- [Kalantari 06] Mahzad Kalantari, Franck Jung, Guillaume Moreau & Jeanpierre Guédon. *Détection entièrement automatique de points de fuite dans des scènes architecturales urbaines*. CORESA, 2006.
- [Kanbara 02] Masayuki Kanbara & Naokazu Yokoya. *Geometric and Photometric Registration for Real-time Augmented Reality*. In IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR, 2002).
- [Kanungo 02] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman & Angela Y. Wu. *An Efficient k-Means Clustering Algorithm : Analysis and Implementation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, 2002.
- [Kawashima-Bertrand 99] Atsuko Kawashima-Bertrand. *Expérience esthétique au musée : impact du parcours et des cartels*. PhD Thesis, Saint Etienne, 1999.
- [Kelly 08] Spencer D Kelly, Sarah M Manning & Sabrina Rodak. *Gesture gives a hand to language and learning :*

- Perspectives from cognitive neuroscience, developmental psychology and education.* Language and Linguistics Compass, vol. 2, 2008.
- [Kim 03] Hansung Kim, Seung-Jun Yang & Kwanghoon Sohn. *3D Reconstruction of Stereo Images for Interaction between Real and Virtual Worlds.* ISMAR, 2003.
- [Kim 06] ZuWhan Kim. *Geometry of Vanishing Points and its Application to External Calibration and Realtime Pose Estimation.* Research report (University of California, Berkeley. Institute of Transportation Studies, 2006.
- [Klein 03] Georg Klein & Tom Drummond. *Robust Visual Tracking for Non-Instrumented Augmented Reality.* Proc. Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR), 2003.
- [Klein 07] Georg Klein & David Murray. *Parallel Tracking and Mapping for Small AR Workspaces.* Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR), 2007.
- [Kosecka 02] J. Kosecka & W. Zhang. *Video Compass.* Proceedings of European Conference on Computer Vision, 2002.
- [Kotler 99] N Kotler. *Delivering experience : marketing the museum's full range of assets.* Museum News, 1999.
- [Kwenzi Mikala 93] Tangu Jérôme Kwenzi Mikala. *La gestualité et les interactions dans la narration d'une épopée : l'exemple de Mumbwanga.* Pholia, vol. 8, 1993.
- [Kytö 11] Mikko Kytö, Mikko Nuutinen & Pirkko Oittinen. *Method for measuring stereo camera depth accuracy based on stereoscopic vision.* Proc. SPIE 7864, Three-Dimensional Imaging, Interaction, and Measurement, 2011.
- [Lai 09] Po-Lun Lai & Alper Yilmaz. *A new approach for vanishing line estimation.* ASPRS Annual Conference, Baltimore, Maryland, vol. 2, 2009.
- [Leonard 91] J.J. Leonard & H.F. Durrant-Whyte. *Simultaneous map building and localization for an autonomous mobile robot.* IEEE/RSJ Int. Workshop on Intelligent Robots and Systems, vol. 3, 1991.
- [Leone 10] Massimo Leone. *Le Repentir – Une énonciation fragmentaire.* Actes sémiotiques [En ligne], 2010.

- [Li 02] Yi Li & Linda G. Shapiro. *Consistent Line Clusters for Building Recognition in CBIR*. Proceedings of the 16th International Conference on Pattern Recognition, vol. 3, 2002.
- [Li 09] Zheng Li & Hansong Li. *Building line feature acquisition based on morphologic filter for infrared images recognition*. MIPPR : Multispectral Image Acquisition and Processing, 2009.
- [Lin 08] Ming-Wei Lin, Yun-Maw Cheng, Wai Yu & Frode Eika Sandnes. *Investigation into the feasibility of using tactics to provide navigation cues in pedestrian situations*. Proceedings of the 20th Australasian Conference on Computer-Human Interaction : Designing for Habitus and Habitat, 2008.
- [Loffler-Laurian 93] Anne Marie Loffler-Laurian. *Les Connaissances scientifiques et le discours muséal*. La lettre de l'OCIM, 1993.
- [Louwerse 05] Max M Louwerse. *Focusing attention with deictic gestures and linguistic expressions*. Proceedings of the 27th Annual Meeting of the Cognitive Science Society, 2005.
- [Lowe 99] David G. Lowe. *Object Recognition from Local Scale-Invariant Features*. Proceedings of the International Conference on Computer Vision, 1999.
- [Lugez 04] Adeline Lugez. *La médiation dans les musées en France*. Rapport de recherche bibliographique, ENSSIB, 2004.
- [Luo 14] Hao Luo, Hideki Shimada & Kenya Sato. *Interactive Co-presence Environment with Mixed Reality Using Remote and Local Avatar*. International Journal on Computer Science and Engineering, vol. 6, 2014.
- [Lutton 94] Evelyne Lutton, Henri Maitre & Jaime Lopez-Krahe. *Contribution to the Determination of Vanishing Points Using Hough Transform*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 16, 1994.
- [Mannion 12] Shelley Mannion. *Beyond Cool : Making Mobile Augmented Reality Work for Museum Education*. Museums and the Web (MW), 2012.
- [Mante 13] Christian Mante. *Accueil numérique dans les offices de tourisme*. 2013.

- [Marton 93] F Marton, G Dall’Alba & E Beaty. *Conceptions of learning*. International Journal of Educational Psychology, vol. 19, 1993.
- [Mason 12] Marco Mason. *The Dimensions of the Mobile Visitor Experience : Thinking beyond the Technology Design*. The International Journal of the Inclusive Museum, vol. 5, 2012.
- [Matas 02] J Matas, O Chum, M Urban & T Pajdla. *Robust Wide Baseline Stereo from Maximally Stable Extremal Regions*. Proceedings of the British Machine Vision Conference, 2002.
- [McLean 97] Fiona McLean. *Marketing the Museum*. Routledge, 1997.
- [McManus 92] Paulette McManus. *Attention! Les visiteurs lisent vos textes dans les musées*. La lettre de l’OCIM, 1992.
- [McNeill 92] David McNeill. *Hand and Mind : What Gestures Reveal about Thought*. University of Chicago Press, 1992.
- [Merleau-Ponty 10] Claire Merleau-Ponty. *La transmission culturelle, nouveaux modes de médiation*. Bibliothèque Nationale de France, 2010.
- [Milgram 94] Paul Milgram & Fumio Kishino. *A Taxonomy of Mixed Reality Visual Displays*. IEICE Trans. Information Systems, vol. E77-D, 1994.
- [Miyashita 08] Tsutomu Miyashita, Peter Georg Meier, Tomoya Tachikawa, Stephanie Orlic, Tobias Eble, Volker Scholz, Andreas Gapel, Oliver Gerl, Stanimir Arnaudov & Sebastian Lieberknecht. *An Augmented Reality museum guide*. IEEE International Symposium on Mixed and Augmented Reality 2008 (ISMAR), 2008.
- [Mooser 07] Jonathan Mooser, Lu Wang, Suyu You & Ulrich Neumann. *An Augmented Reality Interface for Mobile Information Retrieval*. IEEE International Conference on Multimedia and Expo (ICME), 2007.
- [Mottaz Baran 01] Arlette Mottaz Baran. *Publics et musées : représentations emblématiques et rituel social*. Info AMS/VMS, Bâle, 2001.
- [Muja 09] Marius Muja & David G. Lowe. *Fast approximate nearest neighbors with automatic algorithm configuration*.

- In VISAPP International Conference on Computer Vision Theory and Applications, 2009.
- [Nagel 04] Willibald Nagel. *Allgemeine Einleitung zur Physiologie der Sinne*. Handbuch der Physiologie des Menschen, vol. 3, 1904.
- [Nieto 10] Marcos Nieto & Luis Salgado. *Real-time robust estimation of vanishing points through nonlinear optimization*. Proc. SPIE, 2010.
- [Nieto 11] Marcos Nieto & Luis Salgado. *Simultaneous estimation of vanishing points and their converging lines using the EM algorithm*. Pattern Recognition Letters, 2011.
- [Ogre 00] Ogre. <http://www.ogre3d.org/>. 2000.
- [openscenegraph 98] openscenegraph. <http://www.openscenegraph.org/>. 1998.
- [Packer 02] Jan Packer & Roy Ballantyne. *Motivational Factors and the Visitor Experience : A Comparison of Three Sites*. Curator : The Museum Journal, vol. 45, 2002.
- [Packer 06] Jan Packer. *Learning for Fun : The Unique Contribution of Educational Leisure Experiences*. Curator : The Museum Journal, vol. 49, 2006.
- [Pekarik 99] Andrew J Pekarik, Zahava D Doering & David A Karns. *Exploring Satisfying Experiences in Museums*. Curator : The Museum Journal, vol. 42, 1999.
- [Peter N. 97] Belhumeur Peter N., Hespanha P. & Kriegman David J. *Eigenfaces vs. Fisherfaces : Recognition using class specific linear projection*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, 1997.
- [Peyrin 12] Aurélie Peyrin. *Focus : les paradoxes de la médiation culturelle dans les musées*. Informations sociales, 2012.
- [Pflugfelder 05] R. Pflugfelder & H. Bischof. *Online Auto-Calibration in Man-Made Worlds*. Digital Image Computing : Techniques and Applications, 2005. DICTA ' 05. Proceedings, 2005.
- [Piaget 69] Jean Piaget. *Psychologie et pédagogie*. Denoel, 1969.
- [Pielot 10] Martin Pielot, Benjamin Poppinga & Susanne Boll. *PocketNavigator : vibro-tactile waypoint navigation for everyday mobile devices*. Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, 2010.

- [Platonov 06] Juri Platonov, Hauke Heibel, Peter Meier & Bert Grollmann. *A Mobile Markerless AR System for Maintenance and Repair*. Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2006.
- [PocketNavigator 10] PocketNavigator. <http://www.androidguys.com/2010/04/19/pocketnavigator-tactile-pedestrian-navigation/>. 2010.
- [Quan 99] Long Quan & Zhongdan Lan. *Linear n-point camera pose determination*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, 1999.
- [Reitmayr 04] Gerhard Reitmayr & Dieter Schmalstieg. *Collaborative Augmented Reality for Outdoor Navigation and Information Browsing*. Proceedings of the Second Symposium on Location Based Services and TeleCartography, 2004.
- [Rekimoto 95] Jun Rekimoto & Katashi Nagao. *The World Through the Computer : Computer Augmented Interaction with Real World Environments*. Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology, 1995.
- [Resnick 91] Lauren B Resnick, John M Levine & Stephanie B Teasley. *Perspectives on socially shared cognition*. Contemporary Sociology, vol. 21, 1991.
- [Ritchhart 07] Ron Ritchhart. *Cultivating a Culture of Thinking in Museums*. Journal of Museum Education, vol. 32, 2007.
- [Roberts 97] Lisa C. Roberts. *From Knowledge to Narrative : Educators and the Changing Museum*. Washington, DC : Smithsonian Institution Press, 1997.
- [Roth 01] Wolff Michael Roth. *Gestures : their role in teaching and learning*. Review of Educational Research, 2001.
- [Rother 00] Carsten Rother. *A new approach for vanishing point detection in architectural environments*. In Proc. 11th British Machine Vision Conference, 2000.
- [Salisbury 04] Kenneth Salisbury, François Conti & Federico Barbagli. *Haptic Rendering : Introductory Concepts*. IEEE Computer Graphics and Applications, vol. 24, 2004.
- [Seifert 07] Kurt Seifert & Oscar Camacho. *Implementing Positioning Algorithms Using Accelerometers*. Free Scale Semiconductor Application Note, AN3397, 2007.

- [Servant 10] Fabien Servant, Pascal Houlier & Eric Marchand. *Improving monocular plane-based SLAM with inertial measures*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010.
- [Shao 03] Hao Shao, Tomas Svoboda & Luc Van Gool. *Zubud-zurich Buildings Database for Image based Recognition*. Technical Report 260, Swiss Federal Institute of Technology, 2003.
- [Sheth 91] Jagdish N Sheth, Bruce I Newman & Barbara L Gross. *Why we buy what we buy : A theory of consumption values*. Journal of Business Research, vol. 22, 1991.
- [Short 93] J Short, E Williams & B Christie. *Visual communication and social interaction*. Readings in groupware and computer-supported cooperative work, 1993.
- [Siltanen 12] Sanni Siltanen. *Theory and applications of marker-based augmented reality*. VTT Science 3, 2012.
- [Stewenius 06] Henrik Stewenius, Christopher Engels & David Nister. *Recent developments on direct relative orientation*. ISPRS Journal of Photogrammetry and Remote Sensing, 2006.
- [Sutherland 65] Ivan E. Sutherland. *The Ultimate Display*. Proceedings of the IFIP Congress, 1965.
- [Sutherland 68] Ivan E. Sutherland. *A head-mounted three dimensional display*. Proceedings of the December 9-11, 1968, fall joint computer conference, part I, vol. New York, NY, USA, 1968.
- [Swain 91] Michael J. Swain & Dana H. Ballard. *Color indexing*. International Journal of Computer Vision, vol. 7, 1991.
- [Tardif 92] Jacques Tardif. *Pour un enseignement stratégique, l'apport de la psychologie*. Editions Logiques, Montréal, 1992.
- [Tardif 09] Jean-Philippe Tardif. *Non-iterative approach for fast and accurate vanishing point detection*. ICCV, 2009.
- [Torr 97] P. H. S. Torr & D. W. Murray. *The development and comparison of robust methods for estimating the fundamental matrix*. International Journal of Computer Vision, vol. 24, 1997.

- [Torres-Solis 10] Jorge Torres-Solis, Tiago H. Falk & Tom Chau. *A Review of Indoor Localization Technologies : towards Navigational Assistance for Topographical Disorientation*. Ambient intelligence, 2010.
- [Triggs 00] Bill Triggs, Philip Mclauchlan, Richard Hartley & Andrew Fitzgibbon. *Bundle adjustmenta modern synthesis*. 2000.
- [Trinh 07] Hoang-Hon Trinh, Dae-Nyeon Kim, Hyun-Uk Chae & Kang-Huyn Jo. *Geometrical feature-based wall region and SIFT for building recognition*. International Symposium on Electrical-Electronics Engineering(ISEE), 2007.
- [Tsukada 04] Koji Tsukada & Michiaki Yasumura. *Activebelt : Belt-type wearable tactile display for directional navigation*. in Lecture Notes in Computer Science. Springer-Verlag GmbH, 2004.
- [Turk 91] Matthew Turk & Alex Pentland. *Eigenfaces for Recognition*. J. Cognitive Neuroscience, vol. 3, 1991.
- [Van Winkle 12] Christine M. Van Winkle & Katya Lagaya. *Learning during tourism : the experience of learning from the tourist's perspective*. Studies in Continuing Education, vol. 34, 2012.
- [Wagner 06] Daniel Wagner, Dieter Schmalstieg & Mark Billinghurst. *Handheld AR for Collaborative Edutainment*. Proceedings of the 16th International Conference on Advances in Artificial Reality and Tele-Existence, 2006.
- [Yang 10] Gi-Hun Yang, Moon-sub Jin, Yeonsub Jin & Sungchul Kang. *T-mobile : Vibrotactile display pad with spatial and directional information for hand-held device*. The IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 10, 2010.
- [Zhang 96] Zhengyou Zhang. *Determining the Epipolar Geometry and its Uncertainty : A Review*. International Journal of Computer Vision, vol. 27, 1996.
- [Zhao 03] Shanyang Zhao. *Toward a Taxonomy of Copresence*. Presence : Teleoper. Virtual Environ., vol. 12, 2003.