



HAL
open science

Modèle et expériences pour la visite des musées en réalité augmentée sonore

Fatima-Zahra Kaghat Azough

► **To cite this version:**

Fatima-Zahra Kaghat Azough. Modèle et expériences pour la visite des musées en réalité augmentée sonore. Art et histoire de l'art. Conservatoire national des arts et métiers - CNAM, 2014. Français. NNT : 2014CNAM0926 . tel-01087127

HAL Id: tel-01087127

<https://theses.hal.science/tel-01087127>

Submitted on 25 Nov 2014

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

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

École Doctorale
Informatique, Télécommunications et Électronique de Paris
Laboratoire Cedric, Equipe ILJ

THÈSE DE DOCTORAT

présentée par : **Fatima Zahra KAGHAT**

soutenue le : **16 Mai 2014**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline/ Spécialité : **Informatique**

Modèle et expériences pour la visite de musées en réalité augmentée sonore

THÈSE dirigée par :

M CUBAUD Pierre

Professeur des universités, Cnam Paris

RAPPORTEURS :

M LEMORDANT Jacques

HDR, INRIA Rhône-Alpes

M ZREIK Khaldoun

Professeur des universités, Université Paris 8

EXAMINATEURS :

M BOISSARIE Xavier

CEO, Orbe

M GRESSIER Eric

Professeur des universités, CNAM Paris

M NATKIN Stéphane

Professeur des universités, CNAM Paris

Remerciements

Cette thèse doit beaucoup aux nombreuses personnes qui m'ont encouragé, soutenu et conforté au long de toutes ces années. Qu'elles trouvent dans ce travail l'expression de mes plus sincères remerciements.

Je souhaite tout d'abord remercier Pierre Cubaud pour avoir proposé, accepté et encadré cette thèse. J'aimerais le remercier pour la confiance qu'il m'a accordée du début à la fin du travail, pour sa disponibilité, sans oublier son soutien scientifique et psychologique qui m'a aidé à remonter le pont dans plusieurs reprises. Je n'oublie pas enfin son aide précieuse dans la relecture et la correction de ma thèse. La justesse de ses critiques a été très constructive et utile. Sincèrement, j'ai beaucoup appris à ses côtés et je suis chanceuse d'avoir eu l'occasion de travailler sous sa supervision.

Je tiens à remercier également tous les membres de mon jury Jacques Lemordant et Khaldoun Zreik d'avoir accepté de rapporter ma thèse et de l'intérêt qu'ils ont manifesté à mon travail. Ainsi que Stéphane Natkin, Eric Gressier et Xavier Boissarie pour avoir honoré ma soutenance par leur participations à mon jury en tant qu'examineurs.

Heureusement que mes parents, ma famille et mes amis ont tous étaient là. Ils ont tous cru en moi et maintenant, j'y suis ! Je dois beaucoup à mes parents Najia Berrada et Noureddine Kaghat qui m'ont donné un amour et une confiance indicibles. Ils m'ont éduqué sur l'amour du savoir et le goût de la lecture et de la recherche.

Je ne remercierai jamais assez mon époux Ahmed Azough, qui m'avait promis dès le début de notre union un soutien sans faille pour réussir une thèse de doctorat. Je lui dis qu'il a parfaitement tenu sa promesse. Je le remercie pour son soutien au quotidien, sa présence, sa patience, sa compréhension, sa confiance et sa capacité à me faciliter la vie durant ces années de thèse.

Je n'oublierai pas non plus le soutien de mes beaux-parents Hayat Hamidi et Ibrahim Azough qui m'ont toujours offert affection et compréhension, mais aussi mon frère Redouane, mes surs Asmae et Firdaous, mes beaux-frères et belles-surs et mes amis.

Je suis également reconnaissante à tous ceux qui m'ont aidé pendant l'évaluation de mes travaux : Zacharia Agedach, Cécile Leprado, Christophe Pietry, Stéphane Rodovski, ainsi que le personnel du Musée des Arts et Métiers. Sans oublier mes amis du laboratoire Cedric avec qui j'ai passé des agréables moments qui resteront inoubliables : Amélie Lambert, Sourour Layouni, Daniel Enselme, Sami Tak Tak, Viviane Gal, Aréti Damala, Guillaume Levieux et toute l'équipe ILJ.

J'en oublie certainement d'autres et je m'en excuse. Encore un grand merci à tous pour m'avoir conduit à ce jour mémorable.

Enfin, à mon petit Yacine qui m'a posé une sérieuse date limite, je lui dédie cette thèse.

Résumé

Cette thèse consiste à explorer l'usage de la réalité augmentée sonore pour la visite de musées. Notre objectif est de proposer un audioguide permettant d'immerger le visiteur dans une scène sonore constituée de sons ambiants et de commentaires associés aux objets exposés, et de minimiser ses efforts pour découvrir ces objets et interagir avec cet environnement sonore. La première contribution de cette thèse concerne la mise en place d'une preuve de concept de SARIM (Sound Augmented Reality Interface for visiting Museum). Cette preuve de concept a été développée en utilisant des capteurs de position et d'orientation filaire et non filaire. La deuxième contribution concerne la modélisation de la visite augmentée par la dimension sonore. Après une étude des modèles existants, l'objectif est de concevoir un modèle comprenant une représentation du visiteur, du paysage sonore et de la navigation et offrant une grande flexibilité pour créer l'environnement sonore. Ce modèle a comme finalité de faciliter la conception de différents types de scénarios de visite sur la base de la notion de zone d'audibilité. La troisième contribution de cette thèse est le travail d'évaluation mené dans un environnement réel qu'est le musée des arts et métiers de Paris, et qui a permis de confirmer l'utilisabilité, ainsi que l'apport didactique et ludique que procure la réalité augmentée sonore en général et le système SARIM en particulier pour prolonger et enrichir la visite de musées.

Mots clés : Réalité augmentée, visite musée, captation de mouvement, personnalisation, modélisation, spatialisation sonore, audioguide

Abstract

The goal of this thesis is to explore the use of sound to increase the museum visits. We aim to provide an audioguide to immerse the visitor in a soundstage consisting of ambient sounds and comments associated with the exhibits, and minimize its efforts to discover these objects and interact with the sound environment. The first contribution of this thesis is the implementation of a proof of concept of SARIM (Sound Augmented Reality Interface for visiting Museum). This proof of concept was developed using position sensors and guidance wired and wirelessly. The second contribution concerns the modeling of the augmented visit by the sound dimension. After a review of existing models, the objective is to design a model that includes a representation of the visitor, the soundscape and navigation, offering the flexibility to create the sound environment. This model has the purpose of facilitating the design of different types of scenarios based on the concept of audibility area. The Third contribution of this thesis is the evaluation conducted in a real environment what the Museum of Arts and Crafts in Paris, which confirmed the usability, as well as providing educational and ludic apport of the audio augmented reality in general and the SARIM in particular to extend and enrich the museum visits.

Keywords : Augmented reality, museum visit, motion tracking, context aware, adaptive system, user modeling, sound spatialization, audioguide

Table des matières

Résumé	ii
Table des matières	iii
1 Introduction	1
1.1 Contexte	1
1.2 Chronologie des travaux	2
1.3 Organisation de la thèse	4
1.4 Publications	6
2 État de l’art	7
2.1 Introduction	7
2.2 La réalité augmentée	8
2.3 Le son, un outil d’augmentation de la réalité	11
2.4 La spatialisation sonore pour la réalité augmentée	17
2.5 Cas d’étude : la visite et la médiation dans les lieux culturels	22
2.6 La visite augmentée	31
2.7 Conclusion	36
3 Dispositif du visiteur	41
3.1 Motivation	41
3.2 Expérimentations initiales	45
3.3 Captation de mouvement	49
3.4 Dispositif expérimental final	53
3.5 Conclusion	57
4 Modélisation de la visite augmentée	59
4.1 Introduction	59
4.2 Scénarios de visite	60
4.3 Personnalisation du contenu	63
4.4 Description du modèle	66
4.5 Scène sonore	68

4.6	Visiteur	76
4.7	Navigation	83
4.8	Exemple développé	92
4.9	Extension du modèle pour les visiteurs malvoyants	99
4.10	Conclusion	101
5	Evaluation de la visite augmentée	105
5.1	Introduction	105
5.2	Évaluation des dispositifs de réalité augmentée	106
5.3	Expérience 1 : Étude comparative de deux dispositifs dans la salle des communications	113
5.4	Expérience 2 : Étude de deux configurations de scènes sonores dans la salle des automates	129
5.5	Conclusion	141
6	Conclusion et travaux futurs	143
6.1	Bilan	143
6.2	Perspectives	145
A	Annexe	153
A.1	Calcul de l'orientation avec Sparkfun IMU-6DOF-V4	153
A.2	Consignes	157
A.3	Formulaire de recrutement	158
A.4	Questionnaires	159
	Bibliographie	185

1

Introduction

1.1 Contexte

La réalité augmentée permet de mélanger les objets réels et virtuels en utilisant les dispositifs d'interaction 3D de la réalité virtuelle dans un espace réel. La réalité augmentée dans sa dimension visuelle a été depuis longtemps la plus explorée pour améliorer la perception. Bien qu'elle ait permis, en l'appliquant à différents domaines, d'enrichir l'environnement visuel de l'utilisateur, elle a cependant le défaut de contraindre ce dernier à regarder constamment un écran mobile qui va superposer des éléments fictifs sur l'environnement observé, et ainsi restreindre le champs de vue de l'utilisateur.

Ce travail a pour objectif d'explorer la réalité augmentée, non pas dans sa dimension visuelle, mais auditive, en ajoutant numériquement des sons virtuels à l'espace physique. Par ce procédé, nous faisons profiter l'utilisateur des informations sonores qui accompagnent les objets présentés tout en assurant une interaction fluide entre l'utilisateur et les informations associées aux objets. Une interaction qui a comme objectif de diminuer au maximum la quantité d'attention demandée à l'utilisateur pour interagir avec son environnement.

Dans cette thèse, nous décrivons SARIM, un système de réalité augmentée sonore appliqué à la visite du musée des arts et métiers de Paris (MAM)¹. Le MAM a la particularité d'abriter de nombreuses machines mécaniques, électriques, musicales et autre, qui en cas de marche produisent des sons qui contribuent pleinement à créer leurs identité. Malheureusement, de nombreuses contraintes empêchent la mise en marche de ces

1. <http://www.arts-et-metiers.net/>

machines au sein du musée, ce qui prive les visiteurs d'informations pertinentes sur leur fonctionnement. L'incorporation des sons et des bruits de fonctionnement des machines participera à transformer le musée en une usine en marche, d'éveiller la curiosité des visiteurs et ainsi de les faire voyager dans le temps.

L'objectif de ce travail est de proposer, aux visiteurs des musées, une visite à la fois agréable et instructive, en augmentant les œuvres du musée par des sons virtuels. Bien qu'il n'est pas destiné à se restreindre au seul domaine applicatif des musées, SARIM s'inscrit dans la continuité des systèmes de réalité augmentée sonore appliquées à la visite de musée, tels que *ec(h)o* [Wakkary07, Hatala04], *Listen* [Zimmermann08, Eckel01a, Eckel01b, Delerue02, GoBmann02], et *Sound Delta* [Mariette09]. SARIM a été conçu et développé avec le souci de minimiser au maximum le degré d'attention requise de la part de l'utilisateur pour interagir avec son environnement, tout en lui procurant des moyens intuitives d'interaction avec ce dernier. En effet, à la différence des audioguides usuels pour la visite des musée, SARIM ne propose pas de dispositifs manuels pour sélectionner et contrôler le paysage sonore. SARIM assure toutefois une interaction fluide et continue en se basant sur la captation de la position et de l'orientation du visiteur. En plus de transmettre au visiteur le contenu audio compatible à son emplacement et son orientation au sein de l'espace sonore, SARIM permet également de détecter essentiellement les gestes des utilisateurs, qui une fois interprétés permettent de contrôler la scène sonore. L'autre objectif du travail est de concevoir un système, sensible aux intérêts des visiteurs, capable de les capter au cours de la visite, et pouvant adapter la scène sonore en temps réel pour s'accorder au mieux avec les préférences des visiteurs. Sur le plan architectural, le modèle du système SARIM est conçu d'une manière modulaire qui décompose le système en différentes parties disjointes. Chaque partie, est elle même structurée en plusieurs couches. Ceci permet une gestion indépendante des différentes parties, et facilite la paramétrisation du système.

La problématique de ce travail est donc dans la croisée de plusieurs domaines : la réalité augmentée, le son et la spatialisation sonore, l'IHM, la géolocalisation, ainsi que la médiation culturelle (figure3.2).

1.2 Chronologie des travaux

Notre travail a été mené dans plusieurs directions et a été organisé d'une manière itérative (figure 1.2). En effet, chaque axe est développé selon plusieurs dimensions analytique, théorique, pratique et expérimentale.

Le premier axe du travail était la problématique de la captation du mouvement. Nous avons commencé dès le début de l'année 2009 par étudier un module de captation de mouvement filaire, puis nous avons développé un modèle de captation de mouvement basé sur ce module. Ensuite nous avons opté, dès le deuxième semestre 2009, pour les micro-capteurs sans fils de type MEMS (chapitre 3). Nous avons conçu

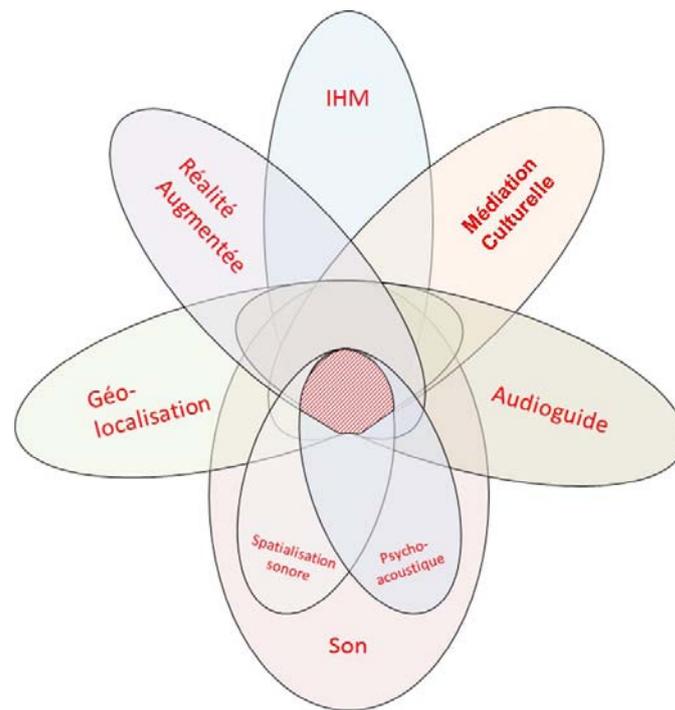


FIGURE 1.1 – Disciplines concernées par ce travail

en premier lieu un modèle de captation de mouvement utilisant les mesures magnétiques d'un capteur inertiel, puis nous avons conçu et développé des algorithmes de détection de mouvement permettant de combiner les mesures magnétiques et gyroscopiques. Les différents développements des solutions de captation de mouvement ont tous été suivis de phases d'expérimentations.

L'axe de l'édition de la scène sonore a pris une part importante dans le travail de thèse, pour ainsi permettre au concepteur de scène sonore de la visualiser et de la contrôler via une interface de rendu visuel 2D ou 3D. Afin d'avoir un retour visuel, nous avons commencé à développer, à partir de 2010, plusieurs versions d'interfaces de visualisation de la scène sonore, comprenant une interface d'édition et une interface de démonstration. Nous avons également développé une interface de suivi de la navigation du visiteur lors de sa visite au musée. Cette interface a été utilisée dans la phase d'expérimentation qui a été réalisée à partir du deuxième semestre 2012. Un travail d'étude et de conception a été ensuite entrepris avec le but de concevoir une interface homme-machine pour l'édition des scènes sonores adaptées à la visite augmentée de musées basé sur notre modèle de visite.

En parallèle, l'axe de la modélisation de la visite augmentée était étudié. Cet axe, développé courant 2011, a permis, après une étude des modèles existants, de concevoir un modèle comprenant une représentation du visiteur, du paysage sonore et de la na-

vigation du visiteur dans le paysage sonore, suivant plusieurs niveaux d'abstraction. Le développement et l'évaluation d'une partie de ce modèle ont été ensuite réalisés pendant l'année 2012.

1.3 Organisation de la thèse

Ce mémoire est composé de cinq autres chapitres. Outre ce chapitre d'introduction, le chapitre suivant décrit un état de l'art (chapitre 2) qui porte sur la notion de réalité augmentée en se concentrant sur l'augmentation sonore. Nous décrivons ensuite les principaux systèmes de réalité augmentée sonore conçus pour le domaine de la visite du musée. Cette partie évoque aussi la problématique de la visite du musée, et les systèmes de guidage et de médiation.

Le chapitre 3 présente l'idée de base qui a été le fil conducteur de cette thèse : celle de rajouter numériquement la dimension sonore à la visite de musées. Nous présentons l'architecture matérielle et logicielle d'un premier prototype, du dispositif du visiteur et des expérimentations préliminaires que nous avons menées pour examiner la faisabilité du projet.

Le chapitre 4 concerne la modélisation de la visite augmentée, qui repose sur la modélisation des différentes entités qui la composent. Dans ce chapitre, nous commençons par illustrer notre objectif à travers des scénarios types qui ont motivé la modélisation. Puis nous enchaînons avec la description du modèle lui-même. Nous présentons la structure du paysage sonore adoptée, composée de trois couches physique, virtuelle et sémantique. Nous modélisons ensuite le visiteur à travers ses informations d'emplacement et ses gestes mais également à travers ses préférences et ses intérêts (section 4.6), et finalement nous décrivons le modèle de navigation du visiteur dans le paysage sonore (section 4.7).

Après avoir proposé une conception du dispositif, des scénarios interactifs, et un modèle de visite, le chapitre 5 décrit la partie expérimentale de notre travail. Elle a pour objectif la mise en application du modèle, des scénarios et du dispositif de l'utilisateur en situation réelle, celle de la visite du MAM. Cette étape d'expérimentation a été nécessaire pour valider nos hypothèses théoriques, et pour pouvoir mesurer l'acceptabilité de notre système (SARIM), jusqu'alors peu connu par les visiteurs du musée.

Ce document se termine par une conclusion et un aperçu sur les travaux futurs (chapitre 6). Nous pouvons citer parmi ces travaux futurs l'outil auteur pour lequel nous avons commencé la conception. Il permettra de concevoir les scènes sonores et de tester virtuellement la mise en place de SARIM.

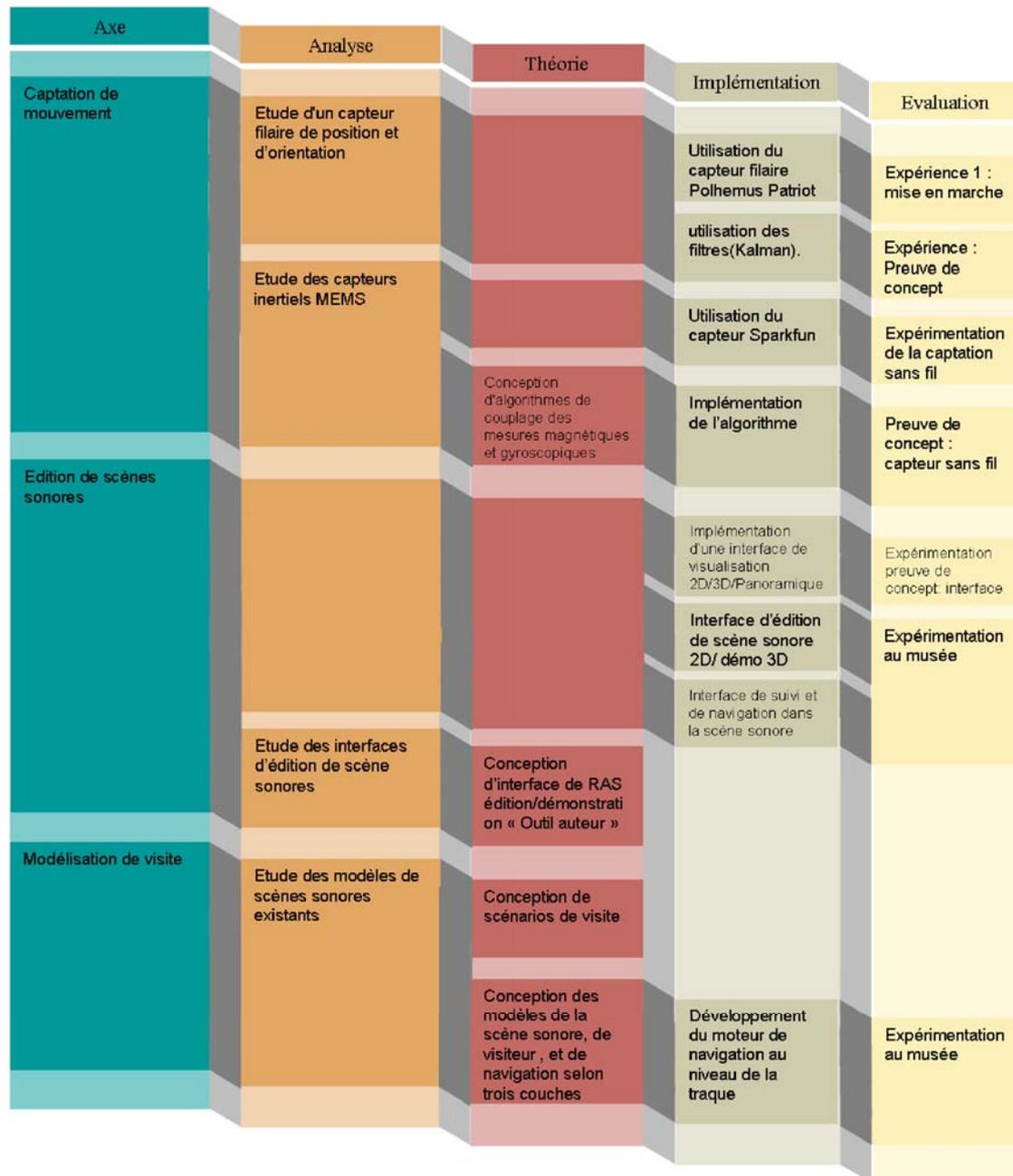


FIGURE 1.2 – Chronologie des travaux de recherche

1.4 Publications

Plusieurs parties de cette thèse ont été publiées dans quelques articles scientifiques dont voici un résumé. D'autres parties ne sont néanmoins pas encore publiées.

Année	Article
2009	[KID09] F. Kaghat , C. le Prado , A. Damala , P. Cubaud - Experimenting with Sound Immersion in an Arts and Crafts Museum, ICEC 09, 8th Int. Conf. on Entertainment Computing, Paris-France, September 2009, pp.173-178, Series LNCS 5709.
2010	[KC10] F. Kaghat , P. Cubaud - Fluid interaction in audio-guided museum visit : Authoring tool and visitor device, VAST'10, Virtual Reality, Archaeology and Cultural Heritage, The 8th EUROGRAPHICS Workshop on Graphics and Cultural Heritage, Full Papers, Paris, France, September 21-24, 2010, September 2010, September 2010.
2012	[KSC12] F. Kaghat , F. Sailhan , P. Cubaud - Application de la réalité augmentée sonore pour les visites de musées par les malvoyants, Handicap 2012, 7ème congrès sur les aides techniques pour les personnes handicapées., July 2012, pp.50, Paris, France.

2

État de l'art

2.1 Introduction

La réalité augmentée occupe de plus en plus de place dans la visite de lieux culturels. Elle permet de mélanger les objets réels et virtuels en utilisant les dispositifs d'interaction 3D de la réalité virtuelle dans un espace réel. Elle offre une expérimentation du monument et une notion de temporalité jusqu'alors imperceptible, sans dénaturer le lieu. Elle permet ainsi de visiter un même lieu de patrimoine à la fois tel qu'il est dans son environnement actuel et, par l'intermédiaire de la réalité augmentée, tel qu'il a été par le passé ou qu'il peut être dans le futur [Durand11]. Cependant, la dimension visuelle a été depuis longtemps la plus explorée en réalité augmentée pour améliorer la perception plus que la dimension sonore.

Le son fait partie intégrante de notre perception du monde. Il est en effet très rare de se trouver dans un environnement dénué de toute stimulation sonore. Lorsqu'un événement sonore se produit, par exemple en montagne ou dans le désert, une sensation étrange presque inquiétante est ressentie [Bouvier09]. En règle générale, le son rassure, on se sent moins seul. Ceci constitue une des raisons qui militent pour que la stimulation sonore soit présente dans un environnement virtuel. Deuxièmement, ne pas percevoir de stimulation sonore peut provoquer un effet de manque chez l'utilisateur qui peut être préjudiciable à l'expérience et donc à l'émergence du sentiment de présence. Troisièmement, le son fournit des informations qui, croisées avec les autres perceptions (visuelles notamment) permettent une bonne compréhension de l'environnement.

L'objectif de ce chapitre est d'étudier quelques systèmes de réalité augmentée sonore existants dans le domaine de la visite de musées. Ce chapitre est divisé en trois sections principales. La première section aborde la notion de réalité augmentée (Section 1.2), et met l'accent sur la dimension sonore dans l'augmentation de la réalité, une dimension qui est souvent négligée (Section 1.3), en terminant par la réalité augmentée sonore utilisant les techniques de la spatialisation sonore pour orienter et immerger les utilisateurs (Section 1.4). La deuxième section se concentre sur la problématique de la visite, le guidage et les systèmes d'aide à la visite ou audioguides, tout en évoquant la problématique de la géolocalisation dans la visite du musée et la personnalisation des contenus. Nous soulignons également dans cette section le manque constaté dans le développement des outils auteurs pour la visite des musées (Section 1.5). La dernière section s'intéresse elle à la croisée des deux sections précédentes ; celle d'étudier les expériences existantes qui visent à augmenter la visite de musées par la dimension visuelle et sonore (Section 1.6).

2.2 La réalité augmentée

La naissance de la réalité augmentée se situerait au milieu des années 1960, avant même que les termes de réalité virtuelle (RV) ou augmentée (RA) aient été envisagés. Elle est survenue au moment où Sutherland [Sutherland68] inventa le premier casque qui combinait la vue de la réalité et une image de synthèse. L'affichage était effectué par des tubes cathodiques miniatures et le suivi des mouvements de la tête de l'utilisateur était assuré par un bras mécanique lié au casque. Un intérêt soudain pour la réalité augmentée s'est révélé. Ainsi, la réalité augmentée est rendue possible par un système capable de faire coexister spatialement et temporellement un monde virtuel avec l'environnement réel. Cette coexistence a pour objectif l'enrichissement de la perception qu'a l'utilisateur de son environnement réel par des augmentations visuelles, sonores ou haptiques [Mallem08].

Quelques définitions

Trouver une définition précise et consensuelle de la RA n'est pas une tâche facile. Cela dépend de la provenance des chercheurs qui la définissent et de l'angle d'étude de leurs travaux. En effet la notion de RA couvre plusieurs définitions traduisant différents courants de recherche menés en parallèle dès les années 90.

Wendy Mackay [Mackay93] définit la RA comme une manière de réintégrer *"l'information électronique dans le monde physique"*. Il s'agit de *"permettre aux gens de tirer parti de leurs compétences dans l'interaction avec le monde de tous les jours, tout en profitant de la puissance des réseaux informatiques"*. Cette définition s'appuie résolument sur la conservation et l'amplification du réel, par exemple, la feuille de papier physique, objet familier, est augmentée de capacités de traitement de l'information [Mackay93, Mackay96]

Pour Azuma [Azuma01] un système de RA complète le monde réel avec des objets virtuels qui semblent coexister dans le même espace que le monde réel. Un système de RA a les propriétés suivants :

- Il combine des objets réels et virtuels dans un environnement réel,
- Il fonctionne de manière interactive en temps réel,
- Il fait coïncider des objets réels avec des objets virtuels.

La réalité augmentée est alors un moyen d'augmenter les sens de l'utilisateur, de transformer des événements imperceptibles en phénomènes visibles audibles ou touchables, c'est vu comme une symbiose entre l'homme et la machine comme jamais auparavant. De son côté Milgram [Milgram99, Milgram94b, Milgram94a] définit la RA comme l'ensemble des cas où un environnement réel est augmenté au moyen d'objets virtuels.

La réalité virtuelle et la réalité augmentée

Depuis le début des années 1990, l'évolution des technologies de la réalité virtuelle permet d'immerger complètement l'utilisateur dans un environnement virtuel. Toutefois, cette immersion ne permet pas à l'utilisateur de percevoir et d'interagir avec son environnement réel [Malle08]. La réalité augmentée est donc parfois définie comme prenant le contre-pied de cette approche immersive, et au lieu d'enfermer les personnes dans un monde artificiel, se propose d'utiliser l'ordinateur pour simplement augmenter le monde réel des utilisateurs [Wellner93]. Ainsi, la réalité augmentée permet justement la perception et l'interaction de l'utilisateur avec les deux environnements. Par conséquent, la réalité augmentée complète la réalité et n'a pas vocation à la remplacer complètement [Azuma97] comme dans le cas de la réalité virtuelle.

Afin de clarifier les concepts, Paul Milgram a proposé une unification de ces concepts en considérant un "continuum réel-virtuel" qui va du monde réel jusqu'aux environnements entièrement virtuels tels que proposés par la RV [Milgram94a, Milgram94b, Milgram99]. Ainsi le concept d'environnement de réalité augmentée peut être placé quelque part le long de ce continuum. Les auteurs définissent ainsi le terme plus générique de "réalité mixte", qui sous-entend différents degrés de fusion des mondes réels et virtuels. La réalité mixte se décline en deux sous cas : la Réalité Augmentée (RA) et la Virtualité Augmentée (VA) selon la proposition d'objets réels ou virtuels (Figure 2.1). Dubois [Dubois01a] a une approche différente car il distingue la RA et la VA selon deux concepts distincts (Figure 2.2), en considérant l'objet de la tâche qui est soit réel (interaction homme-monde réel) ou virtuel (interaction homme machine) :

Le continuum de la Réalité Augmentée : l'interaction se passe dans le monde réel et elle est augmentée par l'ordinateur.

Le continuum de la Virtualité Augmentée : l'interaction avec l'ordinateur est augmentée par des objets et des actions dans le monde réel.



FIGURE 2.1 – Le continuum réel-virtuel [Milgram94a]

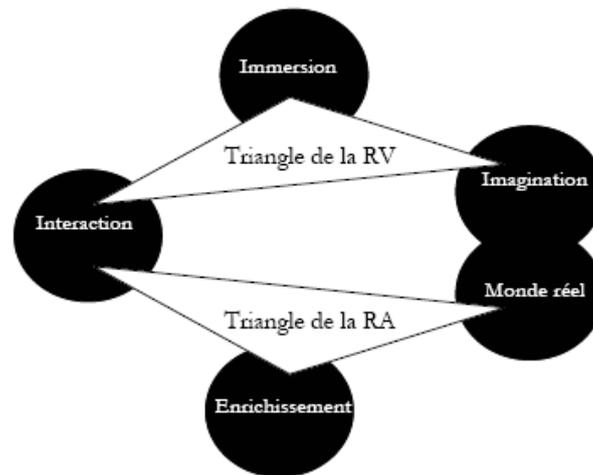


FIGURE 2.2 – Triangle de la Réalité Augmentée et la Réalité virtuelle [Dubois01b]

Le concept de l'augmentation

Une augmentation consiste en la synthèse par ordinateur d'un élément virtuel destiné à l'un des sens de l'utilisateur. Ce virtuel peut être visuel, haptique, sonore, voire olfactif. Par exemple, la réalité augmentée permet la perception d'indices dans l'environnement que l'utilisateur ne pourrait pas percevoir avec ses propres sens, d'où le terme augmentation. Cette assistance permet alors à l'utilisateur de réaliser des tâches complexes qu'il ne pourrait mener à bien autrement [Malle08].

De son côté Wendy Mackay distingue trois types d'augmentation.

Augmentation de l'utilisateur : grâce au dispositif porté par l'utilisateur généralement sur la tête ou sur la main.

Augmentation des objets physiques : l'objet physique est modifié en incorporant des périphériques d'entrée ou de sorties dessus.

Augmentation de l'environnement des utilisateurs et des objets : ni l'utilisateur ni les objets ne sont directement affectés. A la place, des périphériques indépendants fournissent et collectent de l'information de l'environnement, en affichant l'information sur les objets et en capturant l'information sur l'interaction de l'utilisateur avec ceux-ci.

Thomas Baudel [Baudel95], quant à lui définit ce concept d'augmentation comme visant à augmenter les propriétés des objets de notre entourage de capacité de traitement d'informations. En plus de leurs fonctions matérielles, ils acquièrent une fonction informatique, par leur capacité de réagir non pas aux seuls phénomènes physiques auxquels ils sont soumis, mais aussi aux informations qu'ils captent sur l'état de leur entourage (personnes, environnements, autres objets augmentés...).

Gonot [Gonot08] définit l'augmentation comme étant la superposition d'un modèle virtuel 3D ou 2D à la perception (visuelle, auditive, etc.) que nous avons naturellement de la réalité et ceci en temps réel.

Nous avons vu que la définition de la notion de l'augmentation n'est pas fixe et dépend des auteurs, de leur communauté scientifique et de leurs intentions. En Graphique, la RA dérive par exemple de vrais pixels d'origine vidéo qui sont fusionnés à des pixels de synthèse [Azuma97]. En interaction homme-machine, la RA, en réaction à l'immersion virtuelle, s'appuie résolument sur la conservation et l'amplification du réel [Chalon04].

Le concept de l'augmentation de la réalité nous paraît encore plus complexe que l'on peut penser, parce qu'il peut s'agir aussi de diminution de la réalité. En effet, si les développeurs de réalité augmentée ont pour but de rajouter de l'information aux objets qui nous entourent. En revanche, les développeurs de la réalité diminuée visent à éliminer les parties du monde qui nous entoure. Même si la réalité diminuée semble être le contraire de la réalité augmentée, elle est souvent considérée comme étant un type de réalité augmentée [Rolim12].

Même si la modalité sensorielle privilégiée par la réalité augmentée est la vision, elle est applicable aux autres sens de l'utilisateur incluant l'audition, le toucher, et l'odorat [Azuma01]. Nous nous focalisons dans cette thèse sur l'audition pour augmenter la réalité en posant les questions suivantes : comment l'audition peut amplifier et augmenter notre perception ? Qu'entendons-nous par réalité augmentée sonore ? Quel est l'apport du son pour la réalité augmentée ? Et quelle technique de restitution sonore peut-on utiliser pour amplifier la réalité ?

2.3 Le son, un outil d'augmentation de la réalité

Un système de réalité augmentée, comme nous avons vu, permet de mixer le virtuel et le réel. Il met en correspondance ces derniers en cas de mobilité de l'utilisateur ou de l'environnement et permet à l'utilisateur d'interagir en temps réel, avec le système. Le concept de réalité augmentée sonore caractérise les techniques qui permettent d'étendre le son d'un composant réel par un environnement sonore virtuel et par des scénarios de communication. La réalité augmentée sonore permet donc d'augmenter la perception sonore, en associant des objets virtuels (notamment des sons) aux objets réels. La différence fondamentale entre un environnement audio réel et virtuel, est que

les sons virtuels sont originaires d'un autre environnement où ils sont créés artificiellement. Ainsi, la réalité augmentée sonore combine ces aspects d'une manière que les scènes sonores réelles et virtuelles sont mélangées afin que les sons virtuels soient perçus comme une extension des sons réels.

Il convient de signaler que les premiers travaux sur l'augmentation audio n'utilisaient pas le recalage au sens visuel mais plutôt du son localisé dans l'espace. Il s'agissait de son localisé avec l'utilisateur que celui-ci pouvait déclencher par sa présence à proximité d'un lieu en utilisant des systèmes de radio fréquences dont les descendants sont aujourd'hui le RFID [Lyons00]. Néanmoins, des travaux plus récents tendent à apporter une augmentation audio spatialisée (et donc avec recalage) à l'utilisateur.

Dans un premier temps, nous présentons brièvement les phénomènes physiques et physiologiques qui interviennent lors de la perception humaine du son. Ensuite nous allons identifier les multiples apports du son pour la réalité augmentée, en citant quelques systèmes de RA sonores.

2.3.1 De l'audition à la perception

L'audition avec la vision sont deux sens humains très utilisés pour s'informer. L'audition est omnidirectionnelle et par là, elle capte tous les sons de l'environnement qu'ils soient utiles ou non. En contrepartie, elle nous permet de réagir extrêmement efficacement à tout événement sonore, y compris lorsque l'origine du son n'est pas directement visible, qu'elle soit derrière nous, ou cachée derrière un obstacle. L'aspect perceptif de l'audition est un facteur de premier plan dans les applications d'acoustique virtuelle, puisqu'elles s'adressent en premier lieu à des auditeurs humains. L'oreille est donc un récepteur exceptionnel qui n'a rien de commun avec les autres récepteurs artificiels comme les microphones. En combinaison avec notre système auditif, elle nous permet d'analyser d'une manière fine tant du point de vue fréquentiel que directionnel, les ondes sonores qui nous atteignent.

Le système auditif : anatomie et physiologie

Pour mieux comprendre le fonctionnement de notre système auditif, jetons un coup d'œil sur l'anatomie et la physiologie de notre système auditif, afin de comprendre comment une variation de pression est transformée en impulsion électrique compréhensible par le cerveau. Cette transformation se déroule dans les trois régions anatomiques distinctes qui composent l'oreille : l'oreille externe, l'oreille moyenne et l'oreille interne. L'organisation globale de l'oreille est visible sur la figure (Figure 2.3). Selon [Bourdot06], le processus est traité comme suit :

L'oreille externe : l'oreille externe se compose du pavillon et du conduit auditif. Elle fonctionne comme un résonateur qui permet d'amplifier les vibrations qui sont

comprises entre 1500 et 7000 Hz. La forme du pavillon en entonnoir permet une captation du son plus efficace puis concentre l'onde en direction du conduit ou canal auditif. Le pavillon a une configuration particulière faite de bosses et de creux qui est propre à chaque auditeur. Ces reliefs provoquent de multiples réflexions et diffractions qui modifient le spectre du son. Ces modifications du spectre sont exploitées par la suite pour localiser une source sonore. L'onde se propage alors dans le canal auditif jusqu'à rencontrer le tympan.

L'oreille moyenne : suite à la rencontre entre l'onde sonore et le tympan, celui-ci va osciller à son tour. Les oscillations du tympan propagent les variations de pression jusqu'à l'oreille interne par l'intermédiaire de trois osselets : le marteau, l'enclume et l'étrier. Jusqu'ici les ondes se propagent dans un milieu d'impédance faible, or dans l'oreille interne se situe l'organe de l'audition appelé cochlée. Celle-ci est un organe creux rempli d'un liquide dont l'impédance est forte. Le tympan transmet donc les vibrations aux osselets dont le rôle est de transmettre et d'amplifier les vibrations sonores afin de contrebalancer l'impédance forte de la cochlée.

L'oreille interne : l'oreille interne est composée de la cochlée et du vestibule. La cochlée appelée aussi colimaçon, est un tunnel osseux qui est rempli d'un liquide appelé endolymphe. Elle a la forme d'une spirale tapissée de cils. Ces cils sont mis en mouvement par les vibrations acoustiques transmises par l'oreille moyenne. Chaque cil oscille plus particulièrement pour une fréquence donnée puis transmet une impulsion électrique au cerveau par l'intermédiaire du nerf auditif. La spécialisation d'un cil pour répondre à une fréquence donnée permet par la suite au cerveau de distinguer la hauteur des sons.

La perception auditive

Il convient de noter qu'il y a une différence entre ce qu'on entend et ce que notre cerveau interprète. Pour traiter cette différence, la psycho-acoustique est née afin d'étudier les sensations auditives chez l'humain. Elle étudie comment les ondes sonores sont captées par le système auditif et la manière dont elles sont interprétées par le cerveau. Il existe des méthodes de quantification de la perception du son chez l'être humain, telle que l'intensité perçue appelée aussi *sonie*. Elle n'est pas uniquement liée à l'amplitude de la pression acoustique mesurée mais aussi à la fréquence. D'un autre côté, la sensibilité en fréquence varie beaucoup suivant les individus, chez l'enfant le seuil supérieur d'audition se situe autour de 20 kHz. Chez l'adulte, il tombe considérablement jusqu'au-dessus de 15 kHz. L'oreille est particulièrement sensible aux variations de fréquence dans une zone précise entre 1 kHz et 3 kHz, ce sont les fréquences rencontrées dans la voix humaine. Nous rappelons qu'un auditeur moyen peut distinguer environ 2000 changements de hauteur fréquentielle. Comme deuxième méthode de quantification de la perception humaine du son, il y a la sensation de "hauteur" subjective d'un son, appelé aussi *tonie*, elle évolue en fonction de l'intensité de l'onde sonore [Bourdot06].

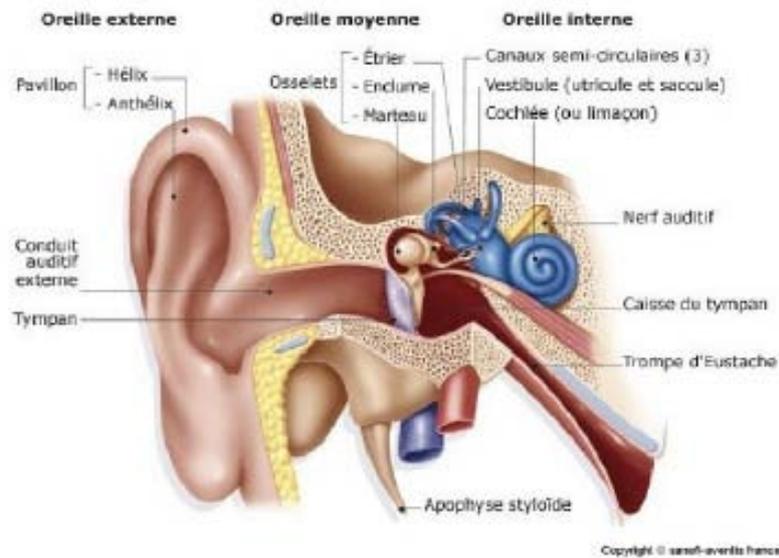


FIGURE 2.3 – Anatomie de l'oreille humaine [Dubois01b]

2.3.2 L'apport de la modalité sonore pour la réalité augmentée

Par rapport à la vision, l'audition possède les caractéristiques suivantes : elle est omnidirectionnelle, ne peut être suspendue (les oreilles n'ont pas de paupières) et il est même possible d'entendre dans l'obscurité. Ainsi, le son contribue pleinement à augmenter notre perception de l'environnement, et ceci est visible à travers le nombre de domaines qui s'appuient sur la dimension auditive pour augmenter la réalité, que ça soit pour la communication, la représentation des données, la compréhension, ou pour favoriser l'immersion, l'émotion ou alors la décoration ou l'effet de surprise. Des exemples d'utilisation du son dans ces différents domaines sont cités dans ce qui suit.

La communication

Parmi les fonctions du son, nous trouvons la possibilité de communiquer avec des agents (humains ou non) par la voix. Si ces agents sont des machines, la tâche devient alors plus ardue car il faut mettre en place un système de synthèse et/ou de reconnaissance vocale efficace. De plus, si le dialogue est ouvert, une intelligence artificielle performante est nécessaire. La voix de l'utilisateur peut aussi être utilisée comme outil d'interaction. C'est-à-dire que l'utilisateur peut communiquer de l'information au système par l'intermédiaire de sa voix pour indiquer un déplacement par exemple ou manipuler un objet. A priori ce schéma d'interaction ne permet pas d'assumer les affordances perçues. Toutefois, dans certaines conditions une interface à commande vo-

cale peut se révéler utile. La voix peut communiquer une information au système par l'intermédiaire d'un ordre, mais elle peut aussi en communiquer par les intonations présentes. Si le système est capable de détecter les émotions portées par la voix on peut alors imaginer adapter l'environnement virtuel ou les événements à celles-ci. Enfin, le son peut être utilisé comme canal pour communiquer, de façon non parlée, de l'information. C'est par exemple le cas pour signaler des situations d'urgence. La réalité augmentée sonore dans cette utilisation a touché plusieurs domaines, notamment le domaine militaire, la maintenance ou alors des applications de bibliothèques augmentées.

La représentation des données

Le son comme facteur d'augmentation de la réalité peut être considéré comme une dimension supplémentaire pour représenter des données. Cette fonctionnalité a été testée dans plusieurs domaines comme l'architecture et la bureautique. Dans le domaine de l'architecture : l'application Ariel [Mackay98] est destinée à assister des responsables de chantier qui manipulent des plans d'architecture sur une table à dessin. Les utilisateurs manipulent une tablette tactile pour capturer les annotations et les commandes de l'utilisateur et un vidéoprojecteur pour l'affichage sur la table à dessin. L'utilisateur peut faire des annotations textuelles, audio ou vidéo.

La compréhension et la mémorisation

Le son est très efficace pour indiquer ou décrire un lieu ou une situation, Deux cas de figure se proposent alors. Soit une perception visuelle est disponible. Il est alors possible de se contenter de sons approximatifs car l'essentiel des informations servant à la compréhension de la situation est fourni par l'image. Soit, du fait du caractère omnidirectionnel de l'audition, l'événement à l'origine du son n'est pas visible, il faut alors proposer des sons parfaitement identifiables par l'utilisateur [Chueng02]. Le fait d'entendre un son en plus de l'image favorise aussi la mémorisation de l'objet en question. La partie sonore fournit des informations supplémentaires qui, combinées ou non avec les autres informations sensorielles reçues par l'utilisateur permettent une meilleure compréhension de l'environnement [Burr06].

L'immersion

Lors des visites culturelles, et sans simulation sonore, l'utilisateur risque se ressentir un manque [Gilkey95] qui est préjudiciable à l'immersion [Bouvier09]. Par ailleurs, l'immersion dans un environnement sonore évite (avec un bon matériel de rendu) à l'utilisateur d'entendre les sons inutiles provenant de l'environnement physique. Ceci permet d'augmenter plus efficacement l'environnement réel avec des éléments virtuels plus appropriés.

La décoration

D'après Crawford [Crawford03], il est assez fréquent que les concepteurs de jeux fassent le choix de privilégier la forme (typiquement, le design graphique et éventuellement sonore) au détriment du fond (i.e. le gameplay). En effet, il semblerait, que les technologies de réalité virtuelle, telles que les synthèses d'image et de son 3D, soient employées dans leur usage le plus commun à des fins essentiellement cosmétiques. Cela n'est pas nécessairement dû à une méconnaissance des capacités expressives et fonctionnelles de la dimension spatiale du son, mais plutôt aux qualités supposées des jeux eux-mêmes [Gonot08].

La surprise

Le son possède une forte capacité de suggestion qui laisse une plage de liberté à l'imagination de l'utilisateur. Par exemple, il est courant que les attentes des auditeurs associées à une situation particulière soient différentes des sons que l'on aurait pu enregistrer dans une situation identique réelle. Ceci pointe encore une fois le rôle des attentes que nous avons déjà mentionné. On peut alors s'intéresser au travail des bruiteurs qui usent souvent de moyens détournés pour assurer une identification rapide des sources sonores.

L'aide au déplacement

Le projet européen Autonomie [Lemordant11], vise à faciliter le déplacement de piétons déficients visuels à l'intérieur et à l'extérieur des bâtiments en milieu urbain. Pour guider les malvoyants, le projet se base sur l'établissement des cartes à l'intérieur ou à l'extérieur des bâtiments, et sur une classification des points d'intérêts (POI). En outre, le guidage des malvoyants est réalisé en utilisant les techniques de localisation de la personne PDR (Pedestrian DeadReconning) basée sur la centrale inertielle disponible sur les smart phones (gyroscope, accéléromètre et magnétomètre). Les auteurs du projet prévoient aussi une synchronisation du système avec des informations de localisation supplémentaires comme le GPS, des bornes Bluetooth ou une infrastructure de RTLS (Real time Localisation System).

Le projet développé par Russell Davies, nommé *SAP*¹ (Situating Audio Platform) [Davies09], repose sur la même finalité qu'est l'aide au déplacement en se basant sur les techniques de réalité augmentée sonore. Il vise à utiliser le son pour guider et donner des indices aux piétons au lieu d'avoir les informations marquées à l'écran. Les auteurs prévoient également une utilisation particulière pour les non-voyants. Le projet a pour but de reprendre toutes les application actuelles qui utilisent la boussole et le

1. <http://www.la-realite-augmentee.fr/videos/la-realite-augmentee-sonore>

GPS (Global Positioning System) du téléphone et d'y ajouter de la voix pour aiguiller l'utilisateur.

Jusqu'à présent, nous nous sommes contentés d'étudier les applications de la modalité sonore sans prendre en compte l'aspect spatial de la diffusion. Dans la section suivante nous étudions la spatialisation sonore en répondant aux questions suivantes : qu'est-ce que c'est que la spatialisation sonore ? À quoi sert-elle ? Comment peut-elle contribuer à l'augmentation de la réalité ? Et quelle sont ses domaines d'application ?

2.4 La spatialisation sonore pour la réalité augmentée

La spatialisation du son occupe aujourd'hui une position similaire à celle que l'art d'orchestration occupait au dix-neuvième siècle. Déployer l'espace ou spatialiser revient à chorégraphier le son : positionner les sources sonores et animer le mouvement [Roads07].

Selon les mêmes auteurs, la spatialisation du son possède deux aspects : le virtuel et le physique. Dans la réalité virtuelle du studio, les compositeurs spatialisent le son en imposant des retards, des filtres, des panoramisations et de la réverbération, ce qui conduit à l'illusion du son émergeant d'environnements imaginaires. Dans le monde physique des salles de concert, les sons peuvent être projetés par un système sonore à partir de plusieurs positions : autour, au-dessus, en dessous, ou à l'intérieur du public [Roads07]. Dans cette thèse, c'est plutôt le premier aspect qui nous intéresse en l'occurrence le virtuel pour créer l'illusion que des sons proviennent de diverses directions de l'espace et organiser ainsi des scènes sonores en trois dimensions.

Afin d'augmenter la réalité par du son spatialisé, le concepteur sonore ou le musicien se trouve face à trois choix de conception : soit il opte pour une restitution fidèle de la scène sonore d'origine d'une manière à ce qu'elle soit identique à la scène réelle, ou il peut ajouter, en temps réel ou au moment du mixage, des effets artificiels de spatialisation supplémentaires, par exemple pour donner plus de consistance aux instruments et voix : réverbération, écho, chorus, etc. Il peut finalement procéder à une restitution basée sur des sons qui ne sont pas reliés à la scène sonore réelle. La spatialisation résultera ainsi entièrement des effets artificiels visant à synthétiser pour l'auditeur un espace sonore virtuel.

2.4.1 Perception spatiale du son

Par perception spatiale d'un son, nous entendons la capacité de déterminer la provenance d'un son. Cela signifie la perception de sa direction et de la distance qui nous sépare de la source sonore. Percevoir la direction c'est connaître son azimut (angle dans le plan horizontal) et son élévation (angle dans le plan vertical). Percevoir la distance c'est connaître la distance qui nous sépare de la source sonore.

2.4.1.1 Perception de la direction

Contrairement à notre vue doté d'un champs limité, notre audition nous permet une localisation des sources sonores dans l'ensemble de l'espace. Cette localisation est rendue possible par des mécanismes complexes de diffraction créée par notre tête et notre torse ainsi que des résonances dans nos oreilles [Blauert83, Middlebrooks91]. Ces indices dépendent de la direction d'incidence de l'onde sonore et nous permettent de localiser les sons d'une manière très précise, en particulier dans le plan. En effet, dans les conditions optimales, l'angle minimum audible entre deux directions voisines est d'environ 1 degré. De nombreuses études ont ainsi montré que les deux principaux indices utilisés pour la localisation auditive sont la différence de temps d'arrivée à nos deux oreilles (interaural time difference, ITD), liée à leur écartement, ainsi que la différence de niveau inter-aurale (interaural level difference ILD), liée à l'ombre acoustique de la tête [Blauert83, Middlebrooks91, Faller04, Zahorik05].

En pratique, les variations d'ILD et d'ITD ne sont pas liées à la direction de provenance du son de manière simple. Elles dépendent également de la fréquence du son en raison de l'interaction complexe des ondes sonores avec notre corps. Les études [Loomis98, Middlebrooks91, Brungart99] ont montré que le pavillon de l'oreille joue un rôle prépondérant dans notre perception spatiale des sons [Blauert83]. Ces effets se traduisent donc par un filtrage du son émis par les sources suivant la direction d'incidence, les filtres correspondants étant communément référencés dans la littérature sous le nom de "Head Related Transfer Functions", ou HRTFs. Dans le cadre d'applications de réalité virtuelle ou augmentée, les techniques de restitution sonore spatiale visent donc, d'une manière plus ou moins précise, à reproduire aux oreilles de l'auditeur ces indices de localisation fréquentiels pour des sources sonores virtuelles.

2.4.1.2 Perception de la distance

La perception de la distance est principalement guidée par deux facteurs : la décroissance de l'amplitude de la pression sonore en fonction de la distance à la source et l'atténuation des hautes fréquences due à l'absorption atmosphérique [Coleman68, Loomis98]. D'autres facteurs peuvent venir s'y ajouter, dans le cas d'environnement réverbérant. Un son lointain a donc une intensité plus faible et voit ses hautes fréquences atténuées. Néanmoins, ces critères, bien que nécessaires, peuvent être insuffisants pour donner une impression directe de la distance à la source, qui reste très variable suivant les auditeurs [Middlebrooks91, Zahorik05]. En particulier, dans la plupart des cas, la distance à la source tend à être sous-estimée, ce qui pourrait se traduire en pratique par une plus grande marge de manœuvre pour éviter les obstacles potentiels dans notre environnement [Wuestefeld92, Zahorik05].

Il convient de dire que la localisation d'une source sonore est influencée par d'autres sens que l'audition [Murat08]. Par exemple, si la source sonore et le stimulus visuel

correspondant occupent des positions spatiales décalées, alors la source sonore semble provenir de la position du stimulus visuel. Les stimuli visuels capturent les stimuli sonores. On parle alors d'effet ventriloque [Alais04]. Les mouvements de la tête affinent la localisation, l'homme tend l'oreille pour obtenir en quelque sorte une deuxième image sonore de l'environnement. Les sensations vestibulaires et proprioceptives sont alors combinées avec les différentes images sonores pour déduire une localisation plus précise.

2.4.2 L'apport du son spatialisé à la RA

Le son spatialisé ou communément appelé "le son 3D" peut contribuer à favoriser l'immersion dans un environnement 3D. Écouter le son localisé dans l'espace ressemble au mappage de texture qui est utilisée pour améliorer la qualité des images ombrées. Le son spatialisé offre également une forte capacité à s'orienter dans l'espace. Un calcul de réverbération fournit des informations sur la nature et le volume de l'endroit où se trouve l'utilisateur [Pontus07]. Une cathédrale n'a pas la même réverbération qu'une bibliothèque ou qu'une scène en extérieur. Le son 3D peut également aider l'utilisateur à estimer la position d'objets sonores dans l'espace tant pour la direction que la distance. D'un autre côté, le son 3D offre un canal important de rétroaction (feedback) qui soit peut être utilement redondant pour le cas de la déficience visuelle, et peut aussi fournir une rétroaction pour les actions et les situations qui sont hors du champ de vue de l'auditeur. Il s'agit d'une utilisation du son en 3D pour les domaines axés sur la situation ou le contexte "situational awareness ou context awareness". De multiples sources sonores virtuelles peuvent être synthétisées et peuvent se produire n'importe où dans l'espace à 360 degrés autour d'un auditeur. Avec un retour audio, le centre de l'attention entre les sources sonores virtuelles peut être commuté à volonté. Une vision, en revanche, exige le mouvement des yeux ou la tête. Ceci est important en particulier pour l'intelligibilité de la parole au sein de multiples canaux de communication. Enfin, au-delà d'améliorer l'intelligibilité d'une mise en scène sonore complexe, cette spatialisation virtuelle apporte une composante esthétique intégrée à l'œuvre elle-même.

2.4.3 Exemples de systèmes augmentés par le son spatialisé

L'histoire de la technologie de la spatialisation sonore est illustrée en détail dans la thèse de Nils Peters [Peters10]. De nos jours, nous la trouvons dans beaucoup d'autres domaines. Dans la téléconférence par exemple, le son spatialisé peut servir comme outil de détection et de suivi des interlocuteurs [Fodróczy08]. Les systèmes virtuels de téléconférence basés sur la spatialisation sonore ont été étudiés par de nombreux auteurs [Virolainen01, Koizumi92]. Généralement, en téléconférence le positionnement de chaque locuteur peut être fait librement, typiquement il est basé sur un "environnement virtuel" prédéfinis [Koizumi92] ou une carte d'une salle de réunion virtuelle [Kang96]. Le système de vidéoconférence de CAR/PE en est un exemple. Il comporte

les flux vidéos des participants disposés autour d'une table virtuelle [Regenbrecht04], et permet à trois participants de différents endroits de communiquer au travers d'un réseau dans un environnement simulant une réunion traditionnelle comme si les participants étaient assis autour d'une table réelle. Le son spatialisé est véhiculé à travers des écouteurs stéréophoniques ou par les systèmes de reproduction phonique à huit canaux.

En jeux vidéo, la spatialisation sonore permet d'enrichir l'information visuelle en aidant les joueurs à localiser des endroits, des événements et des objets non-visuels, mais également à estimer la distance des objets visuels. Sodnik [Sodnik05] a exploré la capacité des utilisateurs à localiser un son virtuel spatialisé dans un environnement de RA avec différentes configurations spatiales de la scène virtuelle. Afin de réaliser l'expérimentation, une scène de RA a été créée et observée par un dispositif monté sur la tête (HMD pour Head Mounted Device). L'utilisateur porte des écouteurs stéréo afin d'entendre les sons spatialisés. Le logiciel fourni utilise la bibliothèque de la vision par ordinateur *ARToolKit*. Thomas Gaudy [Gaudy08], quant à lui propose un jeu sonore de labyrinthe accessible aux personnes aveugles, avec un système d'apprentissage non langagier qui permet un démarrage rapide de la partie de jeu.

Childs et al [Childs03] exploitent la spatialisation sonore pour représenter des données météorologiques, Taylor et al [Taylor07] utilisent le son 3D pour faciliter la compréhension de la dynamique des fluides. D'autres recherches tentent de remplacer les graphiques visuels par des graphiques sonores [Nees08]. Une façon classique de procéder est de faire correspondre l'axe des ordonnées à des variations de fréquence et l'axe des abscisses à un décalage temporel.

Spatialisation sonore et géolocalisation

Outre les systèmes de spatialisation sonore où la position de l'utilisateur est manuellement définie, d'autres travaux de recherche ont combiné les techniques de spatialisation sonore avec celles de la captation du mouvement de l'utilisateur, permettant de mettre à jour automatiquement le contenu sonore avec la position et l'orientation de l'utilisateur à l'intérieur de l'environnement sonore.

Dans le projet *Bristol CyberJacket* [Carro06], un ensemble de sons virtuels sont projetés dans un espace physique. L'utilisateur expérimente l'environnement sonore produit par le biais d'un ensemble d'écouteurs attachés à un ordinateur portable. Une boussole est fixée sur les écouteurs de l'utilisateur afin de fournir les informations sur l'orientation de sa tête. L'ordinateur portable se rend également compte de l'endroit de l'utilisateur. Il emploie le GPS pour les environnements extérieurs et un système de positionnement ultrasonique pour les installations d'intérieur. Avec ces deux sources d'information le dispositif portable peut déterminer où un son devrait être placé par rapport à l'utilisateur.

RA audio pour les malvoyants

Dans l'objectif d'aide au déplacement pour les personnes malvoyantes, plusieurs dispositifs sonores allient géolocalisation des personnes et personnalisation des contenus audio envoyés [Roentgen08]. Un des premiers dispositifs sonores d'aide à la navigation, qui utilise la spatialisation sonore est le système d'orientation personnelle (PGS). Il se base sur un GPS, une boussole et du son spatialisé pour guider les utilisateurs aveugles [Golledge91, Golledge98]. Par ailleurs, d'autres systèmes similaires ont suivi, tels que Drishti [Ran04, Vazquez09] et MOBIC [Petrie96], bien qu'ils n'utilisent pas le son spatialisé. Le projet SWAN [Wilson07] se poursuit dans la veine de PGS, permettant d'expérimenter un rendu sonore qui change avec le déplacement de l'utilisateur dans un monde réel. SWAN requiert d'utiliser des périphériques externes tels qu'un capteur monté sur la tête, un système inertiel, ou une boussole numérique pour la détection de l'orientation.

Motivés par le besoin des gens d'utiliser leurs appareils électroniques, tout en se promenant, Holland et al. [Holland02] proposent un prototype d'une interface utilisateur audio spatialisée pour les mobiles à GPS. Le système a été conçu pour répondre aux besoins des personnes malvoyantes, mais aussi pour permettre aux autres utilisateurs d'effectuer les tâches avec un minimum d'efforts tandis que leurs yeux, leurs mains et leur attention sont engagés ailleurs.

Par ailleurs, et dans le but d'augmenter l'autonomie des déficients visuels dans la navigation, le projet ANR NAVIG [Katz10] a été développé. Le projet vise à concevoir un système permettant à des personnes visuellement déficientes de se déplacer plus facilement en intérieur comme en extérieur, ainsi que de pouvoir localiser des objets d'intérêt dans leur environnement. En plus de l'aide à la mobilité et à l'orientation, NAVIG souhaite permettre aux déficients visuels de localiser et de saisir des objets environnants, sans nécessité de marquer ces objets préalablement. Il repose sur la fusion de données de géolocalisation (GNSS) et de vision artificielle en entrée, et sur la synthèse de son spatialisé en sortie.

RA audio pour des jeux

La spatialisation sonore a également été expérimentée dans le domaine des jeux géolocalisés proposés dans les téléphones mobiles intelligents. Cela permet d'offrir plus de réalisme, d'immersion, et d'engagement émotionnel. Paterson et al. [Paterson10] présentent un prototype de jeu appelé Ghost Hunt Viking (VGH) permettant de localiser et d'interagir avec les vikings virtuels autour des vieux sites de vikings de Dublin. Le jeu s'exécute sur un mobile HTC équipé d'un casque stéréo et d'un GPS. Il offre diverses interfaces visuelles et audio pour localiser les activités paranormales et recueillir des preuves qui se manifestent visuellement ou auditivement. Malgré l'inexactitude du GPS, les perturbations imprévisibles, et les limites de processeur mobile, des ex-

périences ont montré l'importance de la spatialisation sonore dans le monde virtuel comme outil à la fois instructif et riche en émotion.

D'autres projets se sont également intéressés à une application ludique de la spatialisation sonore et la géolocalisation. Le projet *Ambient Horn* [Randell04] qui a exploré le potentiel de la réalité augmentée sonore dans les environnements extérieurs, et plus particulièrement lors d'une visite dans une région boisée. Les utilisateurs qui étaient des enfants, se sont déplacés vers un emplacement dans lequel une balise locale RF a été cachée. Un bruit a été déclenché et joué dans des enceintes sans fil à proximité, tandis que d'autres modules mis en œuvre ont permis aux enfants de recueillir et échanger les informations reçues.

Un autre système intéressant est le système audio-guide de la visite de zoo présenté par Stahl [Stahl07]. C'est un repère auditif partagé, qui transmet des connaissances spatiales et d'aide à la navigation. En utilisant la spatialisation sonore des enregistrements correspondants aux sons et chants des animaux, l'auditeur est en mesure d'identifier la distance et la direction de l'enclos des animaux. Le système présente également des clips audio avec des informations détaillées sur chaque animal.

RA audio pour des applications artistiques

Certaines applications artistiques basées sur la réalité augmentée audio ont également été développées. Certaines de ces recherches ont été menées pour la mise en œuvre de la spatialisation du son sur les dispositifs mobiles équipés de systèmes de positionnement. Strachan et al. [Strachan05], ont décrit la mise en œuvre du prototype *gpsTunes*. C'est un système de navigation qui combine les fonctionnalités d'un lecteur MP3 et d'un capteur GPS, dont la mission est de guider l'utilisateur vers sa destination tout en écoutant la musique. Le dispositif change le volume et l'orientation de la musique avec le déplacement des utilisateurs. Le volume de la musique augmente lorsque l'utilisateur s'approche de sa destination. Les expériences montrent qu'une simple adaptation du volume du son, avec des sources sonores fixes, permet aux gens de naviguer avec succès le long d'une trajectoire spécifiée. La nouvelle version du système [Strachan09] vise à améliorer le suivi des orientations des utilisateurs en utilisant un système inertiel combinant des accéléromètres, des magnétomètres et des gyroscopes.

2.5 Cas d'étude : la visite et la médiation dans les lieux culturels

La visite du musée permet de mettre les personnes en contact direct avec le passé et le présent et permet également d'éveiller la sensibilité au domaine de l'art et de l'histoire. Pour les visiteurs, un bon musée ou plutôt une bonne visite de musée est celle qui permet d'informer et de former, en offrant au visiteur les moyens de bien comprendre lui-même l'œuvre exposée, ses mécanismes (comment il marche) et de situer son rôle

dans une dynamique historique, économique, sociologique et humaine. En plus de son côté pédagogique indéniable, la visite de musée suscite un intérêt émotionnel, elle peut porter une charge affective, manifestée par un sentiment de nostalgie que peut provoquer certains artefacts ou certaines parties du musée. De nos jours, la visite de musée, est devenue également une forme de distraction, grâce aux jeux qui sont proposés de plus en plus pour les visiteurs de musées. En plus, la visite de musée est devenue pour les concepteurs de visite, une occasion pour saisir l'intérêt du visiteur, pour lui proposer du contenu personnalisé. C'est pour cela que la fonction de guidage est de plus en plus présente dans presque tous les musées, et la majorité des musées tente de la développer pour s'adapter aux exigences du public.

2.5.1 Guidage et médiation

Pour favoriser l'aspect didactique et pédagogique des visites, bon nombre de musées proposent aujourd'hui des "prestations" visant à faciliter l'accès des visiteurs aux contenus exposés. On peut citer comme exemples la présence d'un animateur, d'un guide conférencier, ou d'un démonstrateur ; la mise à disposition de cartels, de guides papiers, de catalogues papiers ou numériques, d'ateliers, de réseaux sociaux, de sites web, de guides audio, de guide multimédia, worksheets, de bornes multimédia sur place, d'écran interactif, de feuille d'aide à la visite et la liste est longue. Fauche [Fauche08] distingue deux types d'assistances directe et indirecte.

L'aide directe à la visite : incontournable dans la plupart des missions du musée, elle implique la présence physique du guide appelé aussi médiateur. On parle aussi de médiation orale, ou encore de médiation humaine. Pour le médiateur, ce qui s'y joue se situe à la fois sur le plan des savoirs, et sur le plan de sa relation au public. Quant à la relation qui se noue entre le savoir et le public : il s'agit d'abord pour le médiateur d'accueillir l'autre tel qu'il se donne, à la fois en tant qu'individu et en tant que membre d'un groupe qui a sa dynamique propre.

L'aide indirecte à la visite : consiste à élaborer divers documents, soit écrits (fiches d'aide à la visite, dossier pédagogique), soit destinés à un autre support (audio ou vidéo). En effet, à côté des visites traditionnelles guidées dispensées par des conférenciers, d'autres programmes d'animation sont parfois mis en œuvre et assument la dimension humaine de la relation. Il s'agit d'une palette de supports de médiation qui visent l'autonomie du visiteur. L'audioguide vient ainsi compléter et pallier d'autres types d'aide à la visite désormais courantes au musée ou en voie de l'être dans les musées d'art rénovés [Deshays02] :

- Guides ou petits dépliants papier remis au visiteur à l'entrée.
- Dispositifs signalétiques plus ou moins aboutis, avec titres et sous-titres (parfois assortis de courts textes introductifs).
- Documentation sélective : des fiches, ou feuillets ou "petit journal", disponibles en salle, aux audiovisuels et aux bornes d'information multimédia.

Si l'audioguide peut être proposé comme une aide possible parmi un éventail d'aides, dans certains musées ou exposition il est devenu le principal outil d'assistance, voir le seul [Deshays02]. Pour certains professionnels des musées, l'audioguide est un substitut à la visite guidée.

2.5.2 L'audioguide, une alternative à la visite guidée

Au départ, l'introduction de l'audioguide a été liée à la question de l'accessibilité des publics étrangers. Il s'agissait de leur proposer une visite accompagnée dans leur langue. Si la fonction principale de l'audioguide était la traduction. Il est devenu aujourd'hui un outil de médiation à part entière, à destination du grand public. D'après [Vilatte07], l'audioguide peut aussi répondre au problème particulier des personnes handicapées visuelles ou malentendantes.

Sophie Deshayes [Deshays02] constate que l'audioguide est perçu par le visiteur comme une alternative à la visite guidée, cette dernière apparaissant comme une visite contraignante et dis-qualifiante dans la mesure où elle situe le visiteur dans une logique de dépendance. L'audioguide, quant à lui, permet une visite individuelle qui épargne la visite en groupe (inscription préalable, heure fixe, problème de visibilité, rythme imposé, etc.). Il permet également d'éviter l'inconfort de la visite guidée : problème de stratégie de positionnement à toujours développer vis-à-vis des œuvres en fonction du groupe, et le problème du rythme de la visite imposé par le guide. La liberté du choix des contenus et le respect du rythme de visite propre à chacun sont des opportunités fortement valorisées par les utilisateurs d'audioguide. Ils se plaisent à se sentir accompagnés, guidés, bien renseignés mais non pas contraints par une visite formatée imposant des arrêts obligés et un circuit prédéterminé.

Toutefois d'après l'étude sur les visiteurs menée par Bernard et Hélène Lefebvre [Lefebvre91], deux aspects fondamentaux semblent convaincre les visiteurs des musées des bénéfices de l'audioguide : l'aspect mobilité qui permet aux visiteurs d'écouter le contenu tout en regardant l'œuvre exposée, et l'aspect individualisation du contenu qui donne aux visiteurs le sentiment de solitude de calme et même de détente et qui permet surtout aux visiteurs d'apprendre à leur rythme.

2.5.3 L'audioguide : fonctions et attentes

Loin de sa fonction de traduction, la fonction principale de l'audioguide à laquelle les visiteurs s'attendent c'est "d'apprendre des choses en contexte" [Deshays02, Deshays04]. L'usage de l'audioguide s'affine pour devenir d'une simple aide à la visite, à la contemplation intelligente des œuvres et des collections. Même si, le positionnement de l'audioguide, tel qu'il est perçu par le public, est éminemment pédagogique, les études conduites sur l'usage des audioguides [Lefebvre91], montrent que les visiteurs s'attendent quand même à s'amuser en utilisant l'audioguide. En effet, les utili-

sateurs d'un audioguide attendent à ce qu'il transmette les savoirs d'une manière ludique, et aussi à transmettre quelque chose d'insoupçonné et peu attendu, qu'il révèle la face cachée des choses et qu'il permette au visiteur d'accéder au sens de l'artefact exposé et donner du sens à sa visite. Pour cela, nous proposons de retenir cinq types de contenus :

Contenu qui oriente : les visiteurs ne doivent pas se perdre en repérage. Les éléments d'orientation doivent être définis en situation de visite et ancrés dans l'espace à partir d'éléments de repères visuels susceptibles de structurer le parcours et de guider précisément le regard, par des points cardinaux ou des éléments visuels ou sonores.

Contenu qui informe : qui peut être le contexte historique, le sens d'une représentation (une scène par ex, des symboles), l'usage d'un objet, la démarche d'un artiste, l'intention conceptuelle s'il y a lieu, ou alors les hypothèses scientifiques et/ou les sources de questionnement possibles.

Contenu qui surprend : le contenu doit toujours avoir une vocation de surprendre le visiteur, en contenant des informations imprévisibles, ce qui épargne le visiteur d'écouter du contenu redondant et prévisible, et qui l'encourage à écouter plus pour en connaître d'avantage.

Contenu qui évoque : des extraits littéraires, autres références hors musée, citations, musiques d'ambiance etc. Les extraits musicaux sont appréciés des visiteurs, ils ménagent des pauses à l'écoute et qui peuvent être propices à la rêverie, certains passages musicaux peuvent à l'occasion, réhabiliter la libre déambulation promenade au musée ou au sein de certaines salles.

Contenu qui amuse : le contenu doit être de types différents, musique, sons ambiants, voix, dialogues, etc., l'audioguide ne doit pas négliger son aspect ludique tant pour le public jeune que pour les moins jeunes. Le visiteur ne doit pas sentir qu'il est face à un contenu didactique qui le plonge à l'école, le contenu doit amuser le visiteur et rendre sa visite plus agréable.

2.5.4 Types d'audioguides

Les systèmes d'audioguide proposés dans les musées sont multiples. Ils varient par rapport au média utilisé, à la navigation, aux technologies de localisation des visiteurs, au contenu délivré, au rôle de l'audioguide et au public ciblé. Nous proposons une classification des audioguides portables proposés dans les visites de musées. Cette classification est basée sur le moyen de géolocalisation utilisé dans l'audioguide (manuelle ou automatique), les technologies utilisées pour localiser les visiteurs, ainsi que l'existence ou pas de personnalisation pour chaque visiteur. D'autres classifications d'audioguides ont été proposés dans la littérature [Damala09] selon d'autres aspects non liés à notre travail.

2.5.4.1 La géolocalisation

La géolocalisation

Nous entendons par géolocalisation la manière avec laquelle le visiteur se déplace dans le musée, communique sa position et interagit avec le contenu proposé. Le visiteur du musée a besoin de savoir comment le contenu qu'il reçoit est lié avec l'objet physique exposé dans le musée. Proctor [Proctor05] a défini le lien entre le contenu délivré et l'objet exposé par trois méthodes, soit il est dans l'environnement physique du musée, soit il est sur ou dans le dispositif d'interprétation (l'audioguide), soit il est présenté dans les deux environnements.

Plusieurs techniques de géolocalisation sont utilisées dans la visite de musées, allant des techniques de géolocalisation explicite où le visiteur muni d'un clavier doit saisir un numéro correspondant au numéro affiché devant l'objet exposé, jusqu'aux techniques de géolocalisation qui visent à localiser le visiteur pour lui envoyer automatiquement, le contenu correspondant à son emplacement et éventuellement aussi à son orientation.

Géolocalisation explicite

De nombreux musées ont opté pour la géolocalisation explicite appelée aussi navigation manuelle. Cette approche repose sur un système d'étiquetage, où les œuvres du musée qui sont sujettes à un commentaire audio se voient munir une étiquette placée à côté. L'étiquette porte un nombre, et pour écouter le contenu audio associé, le visiteur doit saisir le nombre affiché sur l'étiquette à l'aide d'un baladeur fourni au début de sa visite. Cette méthode est utilisée dans l'audioguide proposé par le musée des arts et métiers de Paris². Une autre méthode de navigation est utilisée dans le guide multimédia que propose le musée de Louvre de Paris³. Elle consiste à associer aux étiquettes, des photos pour aider le visiteur à identifier l'objet exposé. Une troisième méthode utilisée dans ce contexte, consiste à fournir une carte ou un plan digital aux visiteurs pour se repérer. Cette méthode est largement utilisée dans les systèmes d'aide à la visite notamment, dans Tate Modern Gallery⁴ qui propose cette méthode dans ses parcours destinés aux adultes et aussi aux enfants.

Géolocalisation implicite

La géolocalisation implicite ou la navigation automatique est très utile autant pour les visiteurs que pour le personnel du musée. En effet pour les visiteurs, cela permet aux visiteurs de se localiser automatiquement dans le musée, et ainsi de recevoir les contenus audio correspondant aux œuvres qui les intéressent. Du point de vue du personnel du musée, cela permet de créer des parcours et des expériences de visite individualisés, sans empêcher de créer des applications pour le groupe. Elle permet aussi de localiser chaque visiteur en temps réel et par conséquent de gérer le flux des visiteurs.

2. <http://www.arts-et-metiers.net/>

3. <http://www.louvre.fr>

4. <http://www.tate.org.uk/visit/tate-modern>

Plusieurs technologies et techniques sont expérimentées pour localiser les visiteurs de musée. Le GPS reste largement le plus utilisé pour la navigation en extérieur. Quant à la navigation en intérieure, nous citons d'une manière non exhaustive quelques technologies de localisation utilisées dans le cadre de la visite des lieux culturels.

Infrarouge Utilisée par la cité de la science et de l'industrie à Paris, le musée national du cinéma à Turin en Italie⁵ et le musée des instruments de musique à Bruxelles⁶. Nécessitant un chemin clair entre le marqueur (émetteur) et le récepteur infrarouge contenu dans le dispositif mobile (par exemple un PDA). Cette technologie est utilisée surtout pour le déclenchement des contenus (stockés localement ou délivrés par un serveur), et pour le transfert de données.

Bluetooth Développé comme un protocole à basse puissance et à petite bande passante (1 Mb/s). Le Bluetooth est basé sur les signaux radio et présente plusieurs avantages. Premièrement il n'exige pas de ligne de vue claires entre les outils communiquant. Deuxièmement jusqu'à huit appareils peuvent être connectés simultanément dans un rayon de 10 m. Dans les musées, il est utilisé comme une technologie de déclenchement pour les petites portées. Cependant, Il est incapable de transférer des contenus multimédias riches. C'est pourquoi il a été testé dans le contexte du musée une seule fois (à notre connaissance) par le musée Madame Tussauds⁷ à Londres pour déclencher le PDA.

RFID C'est un terme générique désignant l'ensemble des technologies qui utilisent les ondes radio pour l'identification automatique des objets. La puce RFID comprend : une micro puce et une antenne minuscule qui transmet les données de la puce au lecteur. La puce peut être attachée aux objets qui sont faciles à transporter notamment un bracelet, une carte ou une montre. Cet item envoie alors son identificateur au réseau par le biais du transceiver posé près de l'objet exposé. Deux types d'RFID existent, les passifs qui nécessitent de rapprocher l'émetteur du récepteur pour qu'il y ait communication, et les actifs dotés d'une portée plus importante.

Si on utilise des RFID passifs, on est obligé de s'approcher beaucoup de l'objet, d'où l'intérêt des RFID actifs. Quelques musées utilisent la technologie RFID pour la sécurité et la maintenance de leurs objets. Les RFID sont utilisées aussi dans quelques musées notamment au Danemark pour localiser les enfants perdus dans le parc. La puce RFID est utilisée aussi pour garder trace de la visite. Elle a été utilisée aussi dans le musée des arts et métiers en France⁸ dans le cadre du projet PLUG [Damala10].

5. http://www.museocinema.it/quick_info.php?l=fr

6. <http://www.mim.be/fr>

7. <http://www.madametussauds.com/London/>

8. <http://www.arts-et-metiers.net/>

WIFI Cette technologie permet à un grand nombre de visiteurs de se connecter à intranet et internet pour une large portée. Sa grande bande passante lui permet de transférer des données et leur mise à jour depuis un serveur central. Le WIFI nécessite un serveur central où les logiciels de positionnement sont installés, et où le calcul de localisation est fait, ainsi que plusieurs points d'accès qui relient le signal radio entre les visiteurs et le serveur. La détection de position du visiteur portant un terminal Wifi passe par une triangulation en se basant sur les mesures d'au moins trois points d'accès. Cependant la technologie souffre d'un manque de précision dans la mesure de localisation, et d'une latence pour délivrer le contenu. Ces deux critères peuvent être améliorés en augmentant le nombre de points d'accès. Le WIFI est déployé dans le musée Tate Modern à Londres et au centre de la science à Singapour.

NFC (Pour Near Field Communication), est une technologie de communication sans fil de courte portée qui commence à équiper la majorité des téléphones mobiles intelligents. Avec NFC, la communication se produit entre deux appareils séparés de quelques centimètres. La technologie NFC est basée sur la technologie d'identification par radiofréquence (RFID). Elle peut fonctionner sous plusieurs modes, mode passif ou émulation de carte dans lequel le terminal mobile se comporte comme une carte à puce, mode active dans lequel le terminal mobile devient lecteur de carte, ou en mode pair-à-pair où les terminaux peuvent échanger les informations.

Cette technologie est utilisée dans le musée national de Corée⁹, le musée Hecht¹⁰, le musée national de Cracovie¹¹, etc. L'idée de base est d'orienter les utilisateurs de téléphones mobiles NFC qui ont aussi une application de navigation intérieure embarquée. Un utilisateur peut ainsi indiquer ou déterminer sa position actuelle à l'intérieur d'un bâtiment en touchant son appareil mobile avec les balises réparties à l'intérieur de bâtiment.

UWB La technologie UWB (pour ultra wideband) est une technique de modulation radio qui est basée sur la transmission d'impulsions de très courte durée souvent inférieure à la nanoseconde. Un consortium européen souhaite que cette technologie soit utilisée pour chercher des personnes disparues ou des objets non identifiés. L'UWB peut être utilisé pour le positionnement de mobiles : les récepteurs, qui sont répartis dans l'environnement, peuvent évaluer la distance des balises UWB, situées sur les mobiles, par mesure du temps de propagation des signaux émis par ces balises. Ceci permet au système de calculer la position des mobiles par triangulation.

9. <http://www.museum.go.kr/main/index/index002.jsp>

10. <http://mushecht.haifa.ac.il/Defaulteng.aspx>

11. <http://www.muzeum.krakow.pl/>

2.5.4.2 La personnalisation

La notion de personnalisation est utilisée pour faire référence à un système qui se met à jour et modifie ses caractéristiques selon les besoins de l'utilisateur [Filippini03]. L'origine de la personnalisation du contenu remonte aux premiers jeux de rôles interactifs, où les joueurs choisissent un personnage dont la puissance et les caractéristiques se développent au fur et à mesure de l'avancement du jeu [Gee04]. La personnalisation du contenu peut être statique ou dynamique (appelée aussi interactive). La personnalisation statique signifie que le contenu va être adapté à une catégorie de visiteur, avant ou après la visite. La personnalisation dynamique signifie que le contenu envoyé au visiteur est adapté et est mis à jour durant la visite.

Personnalisation statique

Damala a répertorié neuf critères de personnalisation statique qui sont généralement utilisés dans les pratiques des audioguides de musées [Damala09]. Nous en citons quelques uns :

- L'âge** : le contenu de l'audioguide peut être adapté aux différentes tranches d'âge de visiteurs. Nombre de musées offrent ce genre de service en proposant des parcours pour les enfants et des parcours pour les adultes. Le guide de Sycomore¹² présenté dans le grand palais à Paris, en est un exemple.
- La langue** : le contenu de l'audioguide doit s'adapter à la langue du visiteur. C'est pour cette raison que le guide multimédia du musée de Louvre est disponible en sept langues parlées et en langue des signes françaises.
- La thématique** : le contenu de l'audioguide doit proposer plusieurs parcours thématiques.
- Les capacités spéciales des visiteurs** : l'audioguide doit être adapté aux besoins et capacités spéciales des visiteurs, comme dans le cas des visiteurs à déficience visuelles ou auditives.
- Le groupe de visiteurs/communauté** : pour favoriser l'aspect social de la visite d'un musée, certains audioguides sont destinés à être utilisés par plusieurs visiteurs à la fois.
- La plateforme de livraison de contenu** : certains audioguides proposent aux visiteurs d'utiliser leur propres terminaux (MP3, téléphone,...) pour accéder au contenu de l'audioguide.

L'adaptation du contenu de l'audioguide aux besoins de visiteurs en amont de la visite aide à personnaliser le contenu. Par exemple le projet *HyperAudio* [Petrelli99], qui utilise un modèle de visiteur statique à l'aide d'un questionnaire rempli par le visiteur au début de la visite. Cela reste insuffisant, dans la mesure où le visiteur peut changer ses préférences et découvrir d'autres centres d'intérêts pendant sa visite.

12. <http://www.sycomore.eu/index.php/fr/>

Personnalisation dynamique

La personnalisation du contenu selon les intérêts de l'utilisateur est un sujet d'études important qui est largement étudié dans la littérature, notamment dans la recherche sur internet [Mladenic96]. Nous trouvons également dans la littérature les systèmes qui génèrent des suggestions personnalisées pour les touristes. Nous en citons quelques exemples :

Le projet CyberGuide [Abowd97] produit une série de prototypes d'un guide mobile, qui prend en compte le contexte des visiteurs d'un laboratoire de recherche. Les systèmes prennent en compte seulement les informations concernant l'emplacement et l'orientation actuelle de l'utilisateur en compte, et les applications qui nécessitent une connaissance rudimentaire de l'utilisateur (par exemple les langues qu'il parle) sont envisagées aussi.

Le système ILEX [Oberlander97, O'donnell01] produit les "Étiquettes intelligentes" pour les œuvres d'art qui peuvent être sous forme de pages web, dans une galerie électronique, ou de la parole synthétisée dans une galerie physique. Le système prend en compte les objets que l'utilisateur a visité ou qui étaient mentionnés à l'avant, et se réfère à eux lors de la génération des étiquettes.

GUIDE [Cheverst00, Cheverst02] est un guide touristique électronique portatif qui crée des visites de la ville qui sont adaptés à chaque visiteur. Le système prend en compte les données personnelles des visiteurs (par exemple, nom, âge, sexe) ainsi que leurs intérêts (par exemple maritime, histoire, diététique), leur emplacement actuel, les monuments qu'ils ont déjà visités, et les parcours de visites qu'ils ont emprunté. Le système retrace le parcours de chaque visiteur et met à jour son profil d'intérêts en conséquence.

Le projet "Deep Map" [Fink02] du Laboratoire européen des médias à Heidelberg en Allemagne vise à développer des guides touristiques mobiles et personnels. Afin de permettre à Deep Map de fournir un comportement personnalisé, un modèle de caractéristiques individuelles des utilisateurs (à savoir leurs intérêts et préférences individuels) et des groupes d'utilisateurs est acquis et maintenu par le système de modélisation de l'utilisateur. Les auteurs présentent un modèle d'utilisateur et d'adaptation. L'adaptation est basée sur les intérêts explicite et les données d'usage, sans prendre en compte le focus (l'orientation) de l'utilisateur. Les auteurs ne proposent pas de modèle pour la scène sonore.

La personnalisation dynamique du contenu de la scène sonore dans le cadre de la visite des musées sera mieux décrite dans la section 4.3.

2.6 La visite augmentée

La réalité augmentée occupe de plus en plus de place dans la visite des lieux culturels. Elle permet par exemple de visiter un même lieu du patrimoine à la fois tel qu'il est visible dans son environnement actuel et, par l'intermédiaire de la réalité augmentée, tel qu'il a été par le passé [Durand11]. La RA peut être utilisée sur des objets réels pour les définir et les compléter. Elle peut aussi être utilisée comme moyen de substitution ou de remplacement des objets réels. Elle permet aux visiteurs de manipuler les objets virtuellement sans avoir besoin de voir l'objet réel. Par conséquent, ils permettent dans certaines applications de manipuler aussi les objets des musées qui ne sont pas exposés [Damala09]. La visite des lieux culturels peut être augmentée par la modalité visuelle, haptique, sonore, olfactive ou même affective. Nous allons citer, dans ce qui suit, des exemples de visites augmentées par les modalités visuelles et sonores.

2.6.1 La visite augmentée par la modalité visuelle

La RA pour les visites de musée est encore à ce jour (2013), dans une phase expérimentale. Toutefois, plusieurs prototypes ont été développés, dont les résultats sont prometteurs.

Nous pouvons distinguer deux types de systèmes de réalité augmentée : les systèmes de RA fixe et les systèmes de RA mobiles. Le qualificatif fixe signifie que le système de RA est indépendant de l'emplacement de l'utilisateur et n'assure pas son suivi. Ce dernier est plutôt assurée dans les systèmes de RA mobiles. La vitrine virtuelle [Bimber03] est un exemple de systèmes de la RA fixes, conçu pour permettre aux visiteurs du musée de visualiser les images stéréoscopiques des expositions augmentées. Le scénario initial visait à montrer aux visiteurs les mesures prises par les paléontologues dans la reconstitution des muscles, des tissus mous et les os du crâne d'un dinosaure. Le concept de la RA fixe a été testé aussi par le centre Ename en Belgique [Owen05]. Le système superpose une représentation 3D sur les objets réels exposés dans le musée. L'inconvénient de cette approche est son caractère invasif et non personnalisable. Ces systèmes nécessitent le déplacement récurrent des visiteurs entre le virtuel et le réel.

Parmi les systèmes de RA mobiles, nous pouvons citer le projet *Archeoguide* [Gleue01]. L'objectif de ce système est d'assister les visiteurs pendant leurs visites. Le visiteur porte un HMD (Head Mounted Display), des écouteurs, un capteur de position qui détermine son emplacement. Le projet *Lifepius* [Papagiannakis02] peut être considéré comme le successeur de *Archeoguide*. Il permet aux visiteurs de visualiser les aspects de la vie quotidienne des Romains. Le Musée américain d'histoire naturelle à New York, quant à lui, utilise une application gratuite appelée *Explorer*¹³. Cette application permet de naviguer dans le musée, agissant comme un guide personnel

13. <http://www.amnh.org/apps/explorer.php>

pour une visite personnalisée, tout en permettant aux visiteurs du musée de choisir les pièces qu'ils veulent voir. Au Musée d'Art Moderne de New York, les artistes médiatiques utilisent la réalité augmentée pour créer un spectacle pour le festival Conflux en Octobre 2010. Ils ont utilisé le navigateur de réalité augmentée *Layar* dans un smartphone, pour montrer aux visiteurs les galeries du musée via la caméra du téléphone. En Australie, le Powerhouse Museum a une application de réalité augmentée mobile qui tire des images du musée provenant de Flickr et les montre à travers le navigateur *Layar*¹⁴. *Museum Lab* s'efforce de trouver des solutions à la problématique de la visite de musées avec une combinaison de panneaux imprimés et des animations projetées. Il propose une visite de musée tout à fait novatrice, en superposant des images virtuelles sur celles qui sont perçues par les visiteurs dans la zone de présentation [Mina10].

Une autre application intéressante est celle de l'institut néerlandais d'architecture à Rotterdam, qui propose une visite de la ville par l'intermédiaire d'une application de réalité augmentée et de la technologie *Layar* qui décrit l'histoire sur les divers monuments de la ville. Elle utilise des images, des informations audio et des modèles 3D de bâtiments futurs qui montrent à quoi ils peuvent ressembler à l'avenir. Les utilisateurs ont trouvé que, bien que le guide humain était attentif, l'application de réalité augmentée pouvait fournir plus d'informations, offrant plus de flexibilité en fonction des préférences individuelles des visiteurs [Boulin10].

Le projet ARTSENSE (Augmented Reality Supported adaptive and personalized Experience in a museum based on processing real-time Sensor Events) vise quant à lui à appliquer une augmentation visuelle, acoustique et affective à la visite de musée [Stojanovic12, Damala13, Damala12]. Les auteurs introduisent de nouvelles technologies portables pour détecter en continu et de manière non intrusive le contexte de l'utilisateur afin de déterminer l'intérêt actuel de l'utilisateur. Afin de visualiser le contenu augmenté, ils utilisent une paire de lunettes qui contiennent un capteur du mouvement du visiteur. Les événements acoustiques sont captés en temps réel en utilisant des microphones omnidirectionnels et une unité d'acquisition de données portée par le visiteur. Les auteurs analysent aussi les réactions physiologiques des visiteurs en utilisant des biocapteurs pour mesurer les battements du cœur, la conductance de la peau et l'activité cérébrale.

2.6.2 La visite augmentée par la modalité auditive

Jacomy a été l'un des premiers à étudier profondément la question de la visite augmentée par la modalité sonore dans le contexte muséographique [Jacomy02]. Il trouve que les spécialistes du patrimoine ont principalement été les conservateurs des musées chargés par la société de collecter, de conserver et de restituer les traces matérielles de notre passé. On a donc avant tout conservé ce qui se voit et se touche. Jacomy estime que la majorité des musées n'ont aucune préoccupation particulière en matière de la

14. <http://www.powerhousemuseum.com/layar/>

dimension sonore et que peu de musées ont une véritable approche sonore des présentations. Cependant notre mémoire est également olfactive, gustative et auditive. En prenant l'exemple des musées techniques, en l'occurrence le Musée des Arts et Métiers (MAM) de Paris, ce musée comporte de grosses machines, par exemple la locomotive à vapeur. Il dit " *il est essentiel que les générations puissent regarder et toucher les machines, mais si les enfants qui les approchent ne peuvent sentir les odeurs de graisse chaude et du fumée, ou entendre les sifflements et les grondements de la vapeur, ils n'auront qu'une perception partielle d'un objet qui a pourtant profondément marqué la vie de leurs ancêtres. Faute de pouvoir appréhender ce patrimoine "total", nous pouvons déjà nous pencher sur l'une de ses composantes, importante et pourtant souvent négligée : notre environnement sonore*".

Dans cette section nous allons citer quelques tentatives de sonorisation des musées. Nous allons commencer par les systèmes qui offrent le même contenu sonore à tous les visiteurs du musée, et puis les systèmes audio basés sur l'utilisation de casque audio pour une individualisation et une personnalisation du contenu pour chaque visiteur.

2.6.2.1 Vers un musée sonore

Comme dans le cas de l'augmentation visuelle, nous distinguons deux types de systèmes augmentés par la modalité auditive : les RA sonores fixes et mobiles.

Parmi les systèmes de RA sonores fixes, [Jacomy02] cite quelques exemples d'installations sonores qui ont été effectuées dans certains musées. Un exemple de ces systèmes se trouve dans le musée de la Coutellerie de Thiers, où une salle du sous-sol dégageait un bruit industriel intense. " *Dans cette salle les murs étaient recouverts de photos d'ateliers grandeur nature, quelques machines-outils ont été exposées et une bande sonore reconstituait l'ambiance de l'atelier, avec un niveau acoustique proche du niveau réel*. Il cite aussi le cas du musée français du chemin de fer de Mulhouse. Une grosse locomotive à vapeur est présentée avec une bande sonore qui reproduit, le bruit de la locomotive au démarrage. Cette locomotive est montée sur un banc de roulement qui permet de faire tourner les roues. Ces systèmes de RAS fixes souffrent des mêmes inconvénients qui sont l'aspect intrusif et non personnalisable. Pour y remédier, les systèmes de RA sonores mobiles ont été introduits. Ce sont des systèmes qui visent à localiser les visiteurs du musée, et à personnaliser le contenu audio délivré à chaque visiteur en utilisant des casques audio.

Bederson [Bederson95] a été parmi les premiers à développer un prototype de guide électronique pour les musées se basant sur l'interaction implicite. Celui-ci permet aux visiteurs de recevoir automatiquement les contenus audio appropriés en fonction de leurs positions individuelles. Ce prototype se compose d'un certain nombre de dispositifs que le visiteur emporte avec lui : une source audio numérique à accès aléatoire (Sony MiniDisc™), un microprocesseur (Motorola M68332) et un récepteur infrarouge qui indique à l'ordinateur la position du visiteur. Un petit émetteur infrarouge est placé dans le plafond au-dessus de chaque objet à décrire. Il transmet un nombre unique que

le système emploie pour identifier l'endroit du visiteur. Ensuite, le système joue ou arrête les sources audio sur le terminal du visiteur en fonction de sa position.

Dans le projet européen *Hips* [Benelli99], le visiteur est également localisé. La position et l'orientation du visiteur sont calculées respectivement grâce à la technologie infrarouge et une boussole électronique. Les données de localisation sont envoyées au serveur central qui se charge de délivrer les informations appropriées aux visiteurs.

Dans le projet *CyberJacket* [Carro06], un ensemble de sons virtuels sont projetés dans un espace physique. L'utilisateur expérimente l'environnement sonore produit par le biais des écouteurs attachés à un ordinateur portable. Une boussole est fixée sur les écouteurs de l'utilisateur afin de fournir les informations sur l'orientation de sa tête. L'ordinateur portable se rend également compte de l'endroit de l'utilisateur. Le système utilise un système GPS pour les environnements extérieurs et un système de positionnement ultrasonique pour les environnements intérieurs. Avec ces deux sources d'information, le dispositif portable peut déterminer où un son devrait être placé par rapport à l'utilisateur.

L'un des prototypes que nous avons pu tester est *Perceprio*. C'est un audioguide *Percipio* (du latin " percevoir par les sens ") [Apeloig08] qui permet de déterminer la position du visiteur et aussi ce qu'il regarde, sans intervention manuelle. *Percipio* (figure 2.4), se présente sous la forme d'un appareil audio relié à un boîtier pendu autour du cou. A son arrivée, à l'accueil, le visiteur reçoit *Percipio* dans les mêmes conditions qu'un audio guide classique. Il peut alors se promener librement dans le site, qu'il soit intérieur ou extérieur. Ce système est conçu pour calculer le vecteur visuel de l'utilisateur afin de lui envoyer les commentaires audio liés à l'exposition qu'il regarde. Il est relié à une plateforme multimédia (PDA) et peut fournir un contenu personnalisé en fonction de l'intérêt du visiteur. Le système utilise la technologie infrarouge pour déterminer la position du porteur dans des environnements intérieurs, et le GPS pour les environnements extérieurs. Pour le calcul de l'orientation de la tête (le lacet), *Percipio* utilise un compas magnétique.

Le Musée de Londres a lancé *StreetMuseum* [Farman12], une application iPhone gratuite qui utilise ses collections de photos d'archives associées à la géolocalisation et à la réalité augmentée. Cette application, consacrée à la capitale anglaise, permet de retrouver les anciennes photos de lieux actuels, avec en plus un descriptif audio de la photo prise à l'époque.

2.6.2.2 Spatialisation sonore pour la visite des musées

En plus de la localisation du visiteur dans le musée, notamment le suivi de sa position et son orientation, quelques chercheurs ont pensé à augmenter la visite du musée par du son spatialisé pour favoriser l'aspect immersif de la visite.

Le projet *ec(h)o* [Hatala05] est conçu comme un audioguide interactif pour le musée.

FIGURE 2.4 – Le casque PERCIPIO¹⁵

Le système capte la position du visiteur qui se déplace dans le musée en utilisant la technologie RFID et les caméras, et par conséquent, le contenu audio relatif à ce que regarde le visiteur du musée lui est envoyé via le casque. En plus de la captation de mouvement, le visiteur interagit avec le contenu audio à l'aide d'un cube en bois de forme asymétrique qu'il manipule durant sa visite. En dépit de sa fiabilité, l'efficacité de l'approche est discutable car l'utilisateur spécifie manuellement son orientation à l'aide du cube. En outre, le système utilise des fonctionnalités limitées de la spatialisation sonore, notamment l'effet stéréo. Par contre, l'effet binaural n'est pas offert.

Le projet *Listen* [Eckel01b, Delerue02, GoBmann02, Leprado07, Zimmermann08], quant à lui a comme but d'explorer l'immersion dans les environnements augmentés par l'audio, en superposant un environnement sonore virtuel à l'environnement réel. Ce système a été installé au musée Kunstmuseum de Bonn dans le cadre d'une exposition comprenant les œuvres du peintre August Macke. Le suivi du visiteur est effectué en utilisant des caméras infrarouges afin de reconnaître l'objectif monté sur le casque et sur les objets exposés. Une unité centrale recueille les données de chaque visiteur correspondant à sa position absolue et à son orientation. Puis, le contenu auditif est sélectionné, spatialisé en temps réel et envoyé à l'utilisateur à travers un casque d'une manière binaurale. Néanmoins, Le dispositif de suivie ne permet de suivre que huit visiteurs simultanément.

Le projet *Sound Delta* propose, de son côté, une architecture distribuée en combinant deux architectures locale et centralisée. Il s'agit d'une installation artistique pour l'interaction en temps réel et de l'exploration simultanée de paysage sonore [Mariette09, Mariette10]. L'utilisateur porte un ordinateur utilisé comme récepteur au-

dio et il est invité à une expérience d'écoute en se déplaçant dans un espace physique. L'utilisateur a l'impression que les sources sont en mouvement grâce au changement des paramètres des sources sonores musicales en fonction de la position et de l'orientation de l'utilisateur. Le système mobile est constitué d'un ordinateur de poche ordinateur (Sony Vaio VGN-UX1XN, Windows XP SP3), La localisation est assurée par le GPS et l'orientation avec le capteur XSens MTi miniature. L'utilisateur porte un casque ouvert (Sennheiser HD485). Le système distant se compose d'un point d'accès Ethernet sans fil et un serveur (Apple PowerMac).

Plus récemment, le projet *Corona*¹⁶ a proposé une visite culturelle en réalité augmentée audio de l'hôtel de ville d'Aachen, en Allemagne. L'espace audio est composé de plusieurs sources sonores virtuelles [Heller11, Heller09], correspondant à des monuments historiques de la ville. Le processus du rendu sonore récupère la position et l'orientation de l'auditeur ainsi que la position des sources sonores virtuelles pour émuler une propagation naturelle du son. Les visiteurs découvrent cet espace virtuel audio en continu uniquement en se déplaçant dans la salle du couronnement. *Corona* tourne sur iPhone et son système iOS, à travers lequel l'utilisateur sélectionne l'œuvre qu'elle l'intéresse. Le tracking de la position et de l'orientation de l'auditeur est réalisé à l'aide d'Ubisense. Il s'agit d'un système de tracking en temps réel qui trace la position de l'auditeur avec des tags RF intégrés dans les écouteurs. Pour suivre l'orientation de sa tête, une boussole numérique est également montée sur les écouteurs. Les informations recueillies sont ensuite remis à un moteur de rendu audio. Apple iOS SDK est livré avec une mise en œuvre d'OpenAL, qui assure l'effet stéréo, les auteurs ont ensuite étendu ce rendu par l'ajout d'un filtre passe-bas pour assurer l'effet devant/derrière, et ajouter des réverbérations.

En comparant les projets cité selon le moyen d'interaction avec le paysage sonore. Nous trouvons que dans les projets *ec(h)o* et *Corona* le visiteur sélectionne les œuvres dont il souhaite entendre le contenu grâce à une interface manuelle (cube en bois dans *ec(h)o* et iPhone dans *Corona*). Quant aux projets *Listen* et *Sound Delta*. Ils ont opté pour une interaction avec le paysage sonore basée uniquement sur la déambulation des visiteurs dans le musée. Si cette solution a l'avantage de libérer les mains des visiteurs et de les immerger dans le paysage sonore. Elle a toutefois, l'inconvénient de ne pas donner au visiteur le moyen de contrôler le contenu audio qu'il désire entendre (arrêter ou jouer le contenu, augmenter ou diminuer le volume, etc).

2.7 Conclusion

Dans ce chapitre, nous avons commencé par définir le concept de la réalité augmentée, qui possède un périmètre de définition très large et mouvant. Nous avons situé la réalité augmentée par rapport à la réalité virtuelle, et nous avons défini le concept de

16. <http://hci.rwth-aachen.de/corona>

l'augmentation de la réalité qui peut se faire sur le plan visuel, haptique, sonore, voire olfactif. Nous nous sommes plus particulièrement focalisés sur l'augmentation sonore. Après avoir présenté les concepts de l'audition et de la perception humaine du son. Nous avons présenté l'apport de la dimension sonore dans les systèmes de RA. Dans un deuxième temps nous avons limité nos champs d'étude aux systèmes de RA sonore dans le cas particulier de la visite des lieux culturels. Dans la visite de lieux culturels force est de constater la prolifération d'outils de guidage. Parmi ces outils nous trouvons les guides multimédias qui se chargent d'assurer cette médiation autrefois assurée exclusivement par un médiateur humain. Il convient de noter que plusieurs taxonomies d'audioguides se trouvent dans la littérature. Damala [Damala09] estime que l'on peut avoir huit critères de classification des audioguides. Ils peuvent être classifiés selon le média utilisé, la géolocalisation, le contenu stocké, la personnalisation, le contexte social, la continuité de l'usage et le caractère ludo-éducatif. Nous nous sommes focalisés dans cette étude de l'état de l'art sur les deux critères suivants : la géolocalisation et la personnalisation.

Dans la dernière section nous avons fait le lien entre la réalité augmentée et la visite du musée, nous avons vu, à travers des exemples, comment une visite de musée peut être augmentée, que ça soit par la modalité visuelle ou sonore. Les tableaux 2.2, et 2.4 représentent un récapitulatif de quelques systèmes de RA étudiés dans ce chapitre.

- La première colonne donne le nom de l'application ou de son auteur principal.
- La deuxième colonne indique le domaine couvert par cette application.
- La troisième colonne précise la nature de l'augmentation du point de vue de la perception par l'utilisateur. Nous distinguons principalement les augmentations visuelles et sonores.
- La quatrième colonne indique les dispositifs utilisateur. Ce qui signifie, le moyen d'interagir avec l'environnement augmenté et le type d'ordinateur pour traiter les données.
- La cinquième colonne indique si l'application permet ou pas le suivi de la position de l'utilisateur, et si oui par quelle technologie.
- La sixième colonne permet de préciser si l'application permet ou pas le suivi de l'orientation de l'utilisateur et par quel type de capteurs.
- La septième colonne précise si la technique de la spatialisation sonore est utilisée ou pas.
- Enfin la huitième colonne précise si l'application supporte la mobilité. Nous distinguons : F (l'application est fixe), M (l'application est mobile).

Nom de l'application ou de l'auteur	Domaine d'application	Nature de l'augmentation	Dispositif	Suivi position	Suivi orientation	Spatialisation	Mobilité
Ariel [Mackay98]	Architecture	Visuelle, sonore	Table, localisateur	caméras	Non	Non	F
Autonomie [Lemordant11]	Aide à la navigation	Sonore	Smart-phone	GPS	centrale inertielle	Non	M
CARE/PE [Regenbrecht04]	Téléconférence	Sonore	Ecouteurs stéréo, ordinateur	Non	Non	Stéréo	M
Sodnik [Sodnik06]	Jeu vidéo	Sonore	Ecouteur stéréo, HMD, PC windows	caméras	Non	Binaural	F
Labyrinthe [Gaudy08]	Jeu vidéo	Sonore	Casque stéréo, PC standard	Non	Non	Binaural	F
PGS [Golledge98]	Aide à la navigation	Sonore	Casque, tél mobile Nokia	GPS	Boussole	Non	M
Drishti [Vazquez09]	Aide à la navigation	Sonore	Casque	GPS	Boussole	Non	M
MOBIC [Petrie96]	Aide à la navigation	Sonore	Casque	GPS	Boussole	Non	M
SWAN [Wilson07]	Aide à la navigation	Sonore	Casque, micro PC(Sony Vaio)	GPS	Boussole	Non	M
GpsTunes [Strachan09]	Navigation	Sonore, visuelle	UMPC Sumsung	GPS	Système inertielle	stéréo, panning	M
VGH [Paterson10]	Jeu vidéo	Sonore, visuelle	casque	RFID	Non	Non	M
Ambient Horn [Randell04]	Promenade ludique	Sonore	Casque, prototype Ambient Horn	Non	Non	Non	F
CyberJacket [Carro06]	Navigation	Sonore, visuelle	carte Gumstix	GPS	Boussole	stéréo	M
Sound Delta [Mariette09]	Installation sonore	Sonore	casque stéréo, micro ordinateur(Sony vaio), serveur	GPS	Boussole	Stéréo	M

TABLE 2.2 – Tableau de synthèse de la revue des applications de RA (Partie 1)

Nom de l'application ou de l'auteur	Domaine d'application	Nature de l'augmentation	Dispositif	Suivi position	Suivi orientation	Spatialisation	Mobilité
Audio GPS [Holland02]	Navigation, aide à la visite	Sonore	casque, ordinateur standard	GPS	Boussole	Stéréo	M
Stahl [Stahl07]	Visite de zoo	Sonore	casque	GPS	Boussole	Stéréo	M
Terracota Warios [Thian12]	Visite de musée	Visuelle	iPhone	GPS	Boussole	Non	M
Life plus [Papagiannakis02]	Visite de musée	Sonore, visuelle	HMD	Non	Non	Non	M
Archeo-guide [Gleue01]	Visite de musée	Visuelle	visio-casque	Oui	Non	Non	M
Emane [Owen05]	Visite de musée	Visuelle	-	Non	Non	Non	F
Street Museum [Farman12]	Visite de la ville	Visuelle, sonore	iPhone	GPS	Boussole	Oui	M
Explorer [AMNH]	Visite de la ville	Visuelle	Ipod touch	GPS	Non	Non	M
Museum Lab [Mina10]	Visite de la ville	Visuelle	Baladeur	Non	Non	Non	M
Bederson [Bederson95]	Visite de la ville	Sonore	PC portable	IR	Non	Non	M
Percipio [Apeloig08]	Visite du musée	Visuelle, sonore	PDA	GPS, IR	Boussole	Non	M
Hips [Oppermann99]	Visite du musée	Sonore	casque	IR	Boussole	Non	M
Ec(h)o [Hatala05]	Visite du musée	Sonore	casque stéréo, PC fixe	RFID	Cube, cameras	Stéréo	M
Listen [Zimmermann08]	Visite du musée	Sonore	casque stéréo, ordinateur fixe (serveur)	Caméras IR	Boussole	Binaural	M
Navig [Katz10]	Aide à la visite	Sonore	casque stéréo, PC portable	GPS	Boussole	Binaural	M
Corona [Heller11]	Visite du musée	Sonore	casque stéréo, iphone	Ubi-sense	Boussole	Stéréo	M
Artsense [Stojanovic12]	Visite du musée	Sonore, visuelle, affective	Caméras, HMD, eye tracking	-	-	Binaural	M

TABLE 2.4 – Tableau de synthèse de la revue des applications de RA (Partie 2)

3

Dispositif du visiteur

3.1 Motivation

3.1.1 Contexte

Le Musée des Arts et Métiers (MAM) interpelle par la richesse fonctionnelle et technique des œuvres qu'il abrite. En effet, le MAM de Paris, expose environ 4000 objets de science et de technologie de différents domaines : communication, construction, transport, mécanique et autres (Figure 3.2). Un ensemble d'objets exposés dans le musée est présenté dans la figure 3.1. Des bijoux de technologie uniques aussi bien par leurs formes étonnantes, à la fois les grands (le fardier de Cugnot, l'aéroplane de Clément Ader, etc.) et les moins grands (les automates par exemple), que par les sons qu'ils produisent quand ils se mettent en état de marche. Ces sons peuvent être complexes ou simples, intenses ou légers, tellement caractérisants qu'ils contribuent pleinement à créer l'identité des objets qui les émettent.

Bruno Jacomy a structuré les objets qui produisent des sons dans le MAM [Jacomy02]. Il a développé une typologie d'objets sonores se trouvant dans le MAM en les répartissant en quatre catégories :

Les objets conçus pour produire des sons agréables comme la musique. Comme exemple nous pouvons citer la salle des automates qui contient plusieurs objets qui produisent de la musique comme "La joueuse de tympanon".

Les objets qui produisent des sons fonctionnels (horloge, alarme, télévision, phonographe).



FIGURE 3.1 – Quelques objets des collections du musée des arts et métiers.

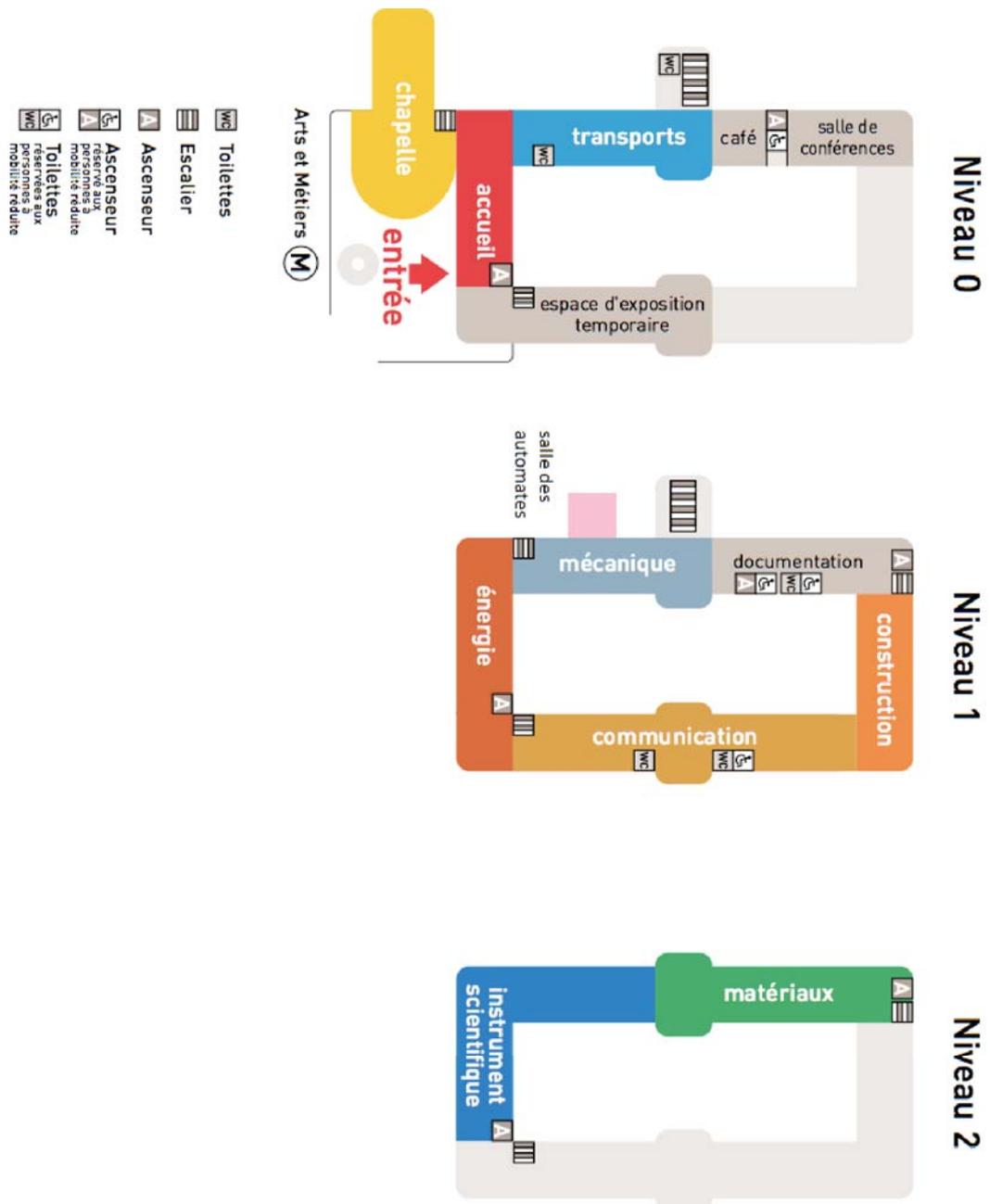


FIGURE 3.2 – Plan des salles du musée des arts et métiers.

Les objets dont le fonctionnement produit des sons (moteurs, le métier à tisser, outils, instruments scientifiques).

Éléments sonores liés à l'environnement du musée et à ses éléments architecturaux (appareil multimédia, sol, murs d'écho).

De nombreuses raisons pratiques empêchent de mettre en marche ces machines, ce qui priverait les visiteurs du musée d'informations importantes concernant le fonctionnement de ces machines. Activer les sons associés aux objets exposés par haut-parleurs placé à coté de chaque machine serait également inapproprié. En effet, ceci priverait le visiteur d'une expérience individuelle, et transformerait la visite du musée d'une expérience personnelle vers une excursion en groupe.

L'idée est donc d'associer à travers de la RAS un ou plusieurs contenus sonores virtuels aux objets réels exposés au musée, puis de renvoyer à chacun des visiteurs un contenu audio personnalisé approprié à sa position et à son orientation. Le concepteur du scénario de visite peut choisir des sons ambiants qui donnent une idée sur l'objet visité, ou des sons de paroles présentant l'objet exposé, ou même des sons non familiers pour susciter la curiosité des visiteurs et favoriser l'effet de surprise. Ce chapitre vise à présenter la structure générale de SARIM, le système que nous avons conçu et développé pour réaliser notre scénario de spatialisation sonore. Le chapitre commence par décrire les premières conditions expérimentales qui ont motivé le choix du dispositif, puis décrit les dispositifs de captation de mouvement utilisés tout au long de la thèse, avant de finir par présenter le dispositif final retenu.

3.1.2 Choix de conception

SARIM est conçu pour permettre au visiteur du musée de naviguer dans un environnement physique augmenté par une scène sonore virtuelle. Cette dernière comporte plusieurs objets sonores liés chacun à une œuvre exposée dans le musée. Pour se faire, le visiteur doit porter un casque audio sur lequel est fixé un capteur d'orientation et de position, ce capteur transmet les informations sur la position et l'orientation du visiteur au système de la spatialisation sonore. Ce dernier consulte les sons contenus dans sa base de données, met à jour les paramètres de la spatialisation sonore, puis transmet le son spatialisé au moteur audio qui le renvoie à son tour d'une manière binaurale au casque du visiteur.

Bien que le système SARIM s'inscrit dans la ligné de quelques systèmes similaires de RAS appliqués à la visite de musées, tels que *ec(h)o* [Hatala05], *Listen* [Zimmermann08], *Sound Delta* [Mariette09, Mariette10] et *Corona* [Heller11, Heller09], cités dans le chapitre 2. L'approche suivie est différente par rapport aux approches existantes. En effet SARIM se distingue de ces systèmes par quelques caractéristiques. Premièrement, en le comparant à *ec(h)o* [Hatala05] ou à *Corona* [Heller11, Heller09] par exemple, SARIM est un dispositif à utilisation main libre. Il est capable de capter l'orientation de la tête du visiteur et de calculer son vecteur visuel, sans aucune action de ce

dernier à l'exception de son déplacement et de ses gestes. Deuxièmement, et en le comparant cette fois à *Listen* [Zimmermann08] ou à *Sound delta* [Mariette09, Mariette10], SARIM offre, en plus de l'adaptation de l'espace sonore aux intérêts du visiteur, un moyen d'interaction intuitif basé sur les gestes. Celui-ci permet à l'utilisateur de mieux contrôler l'espace sonore. En plus de ça, SARIM est un système entièrement distribué, ce qui signifie qu'il dispose d'un dispositif autonome pour chaque utilisateur, et que la gestion du système est réalisée localement. Par conséquent, SARIM peut être utilisé simultanément par un grand nombre de visiteurs.

3.2 Expérimentations initiales

3.2.1 Scénario : combinaison de l'immersion et la description

Les premières expérimentations informelles de notre prototype ont été réalisées en utilisant le capteur de position et d'orientation *Polhemus Patriot* (voir section 3.3.1). Au sein du laboratoire, dans une salle de 4x4 mètres, trois photos correspondant à des œuvres du musée ont été disposées sur les murs, tandis que la même scène est reproduite en temps réel sur une interface visuelle 3D. Ces expérimentations ont été menées auprès d'une dizaine de sujets, qui ont expérimenté notre prototype. Par la suite, d'autres expérimentations informelles ont suivi, et ont été réalisées cette fois dans la salle des automates du MAM en utilisant le capteur d'orientation *Sparkfun IMU 6 DOF* (figure 3.6). L'auditeur porte un casque d'écoute stéréo, sur lequel le capteur de mouvement est fixé. Son mouvement est analysé en permanence.

Le scénario de visite que nous avons proposé initialement est un scénario qui combine l'immersion sonore dans le musée à travers des sons ambiants et de la parole décrivant les objets. Lorsque le visiteur entre dans l'environnement sonore, sa position et l'orientation de sa tête sont automatiquement envoyées au système de spatialisation sonore. Celui-ci récupère le contenu audio approprié en fonction de la position de l'utilisateur et de son orientation, puis envoie à l'utilisateur le contenu audio sous format binaural via le casque. Les actions suivantes ont été expérimentées :

Action 1 : Quand l'utilisateur se dirige vers un objet exposé, il commence à entendre le son ambiant correspondant à cet objet. Le son est envoyé au visiteur d'une manière spatialisée. Le volume et l'orientation du son, dépendent de la distance qui sépare le visiteur de l'objet et de l'orientation de la tête du visiteur par rapport à cet objet

Action 2 : Quand l'utilisateur se positionne entre deux ou plusieurs objets exposés, il entend simultanément les sons ambiants de toutes les sources sonores correspondant à ces objets, tandis que chaque son est envoyé au visiteur en fonction de sa position et l'orientation de sa tête. Puisque des sons différents proviennent de dif-

férentes directions et positions, le visiteur doit être capable de localiser les objets qui l'entourent et d'estimer la distance qui le sépare de ces objets.

Action 3 : Si l'utilisateur s'approche d'un objet exposé et le fixe durant un certain nombre de secondes (spécifié par le concepteur en utilisant l'outil d'écriture de scènes sonores), il commence à entendre uniquement la description de cet objet.

Tous les changements dans le contenu audio que le visiteur entend sont effectués d'une manière fluide et progressive, de sorte que le visiteur ne sente pas de coupures dans le paysage sonore. Lorsqu'on passe d'une action à l'autre et que le contenu sonore doit être changé, le volume du dernier son est diminué progressivement au fur et à mesure que le volume du nouveau son augmente jusqu'à ce que le nouveau son remplace l'ancien. Cette phase de transition dure quelques seconde.

3.2.2 Rendu sonore

SARIM est conçu d'une manière modulaire qui exige l'utilisation d'un moteur de rendu sonore binaural. Pour des raisons pratiques, le rendu "binaural" réalisé dans notre travail s'appuie sur l'API *NativeFmodex*¹ (la version native de *FMOD*). *FMOD*² est une API (Application Programming Interface) de programmation et un outil pour la création et la lecture du son interactif. Pour créer "l'effet binaural", nous avons utilisé l'HRTF statique de *FMOD*, à travers le mode *FMOD_INIT_SOFTWARE_HRTF* qui permet l'initialisation du système sonore *FMOD*. Ce mode ajoute un effet filtre passe-bas qui fait que le son paraît plus terne lorsque la source sonore se situe derrière l'auditeur.

3.2.3 Rendu visuel

Dès le début des développements, nous avons senti le besoin d'avoir un retour visuel : une interface qui montre les conséquences des gestes du visiteur et le retour sonore qu'ils produisent. Pour créer un retour visuel, nous avons commencé par concevoir et réaliser une simulation virtuelle d'une visite du musée guidée par le son. Cela a été réalisé en créant un mini musée virtuel dans le laboratoire. Chaque objet exposé est associé à différents types de contenus audio : une description vocale de l'objet, un son ambiant ou une représentation musicale correspondante à cet objet. Deux modes d'exécution sont possibles dans ce système : Un mode "virtuel" où la position et l'orientation du visiteur sont saisies manuellement à l'aide de la souris et du clavier. L'utilisateur peut ainsi simuler un déplacement dans l'environnement virtuel, approcher les objets et entendre le contenu audio qui leur sont liées. Le deuxième mode, est le mode "réel". Il est utilisé pour surveiller et monitorer la visite. La position du visiteur et son orientation sont récupérées à partir du capteur de mouvement. Cet interface n'est pas destinée

1. <http://jerome.jouvie.free.fr/nativefmod/index.php>

2. <http://www.fmod.org/>

à être utilisée par le visiteur qui aura lui les mains libres, mais plutôt par le concepteur de scénarios de visite.

De même, deux interfaces graphiques ont été attachées au système : Une interface 3D et une interface 2D (figure 3.3). L'interface 3D, de vue subjective, reproduit en temps réel le champ de vision du visiteur au même temps que son déplacement. Dans le cas où le mode virtuel est sélectionné, le visiteur peut naviguer en utilisant un dispositif de pointage connecté au système (souris, clavier). Dans l'interface 2D, de vue de dessus, le système affiche la carte des objets sonores et la position et l'orientation du visiteur au sein de l'environnement du musée. Les objets sont représentés sous forme d'icônes visuelles tandis que le visiteur est dépeint par une icône et une flèche. L'interface permet alors au responsable de la visite d'activer ou de désactiver l'émission des son particulier en cliquant simplement sur l'icône associée. Le type du contenu est représenté par une liste pour permettre au responsable de la visite de modifier le type de contenu sonore.

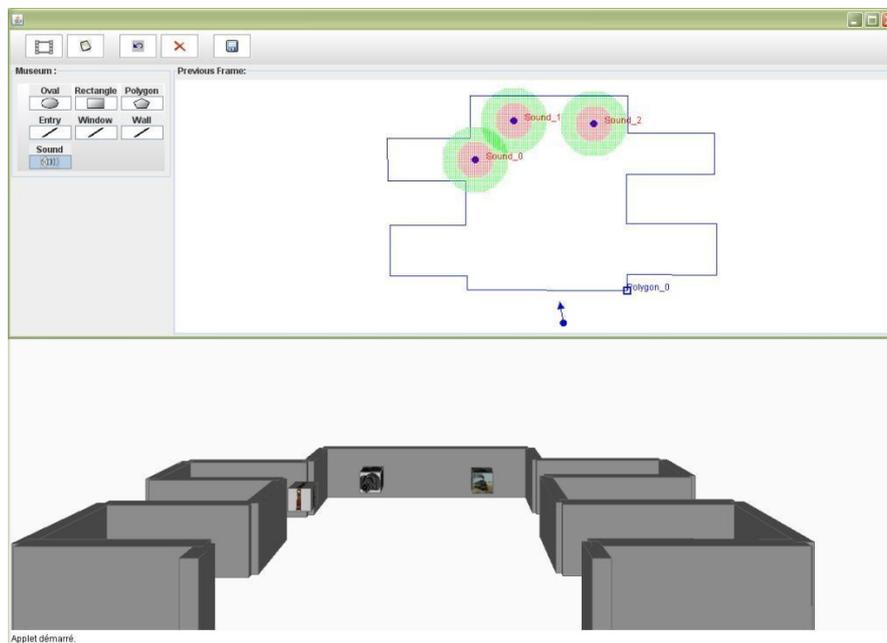


FIGURE 3.3 – Interfaces graphiques pour les expérimentations initiales

Une autre interface de vue panoramique a été également développée est utilisé dans le cadre des expérimentation informelles réalisées au MAM (figure 3.4).

3.2.4 Perception simultanée des contenus sonores

Bien que l'architecture du système prend en charge un nombre illimité de contenus audio, la question d'une perception optimale des sons est fondamentale pour notre sys-



FIGURE 3.4 – Vue panoramique des trois objets sonorisés

tème. Pour cette raison, nous avons effectué un test sur le nombre maximum de sons différents qui peuvent être distingués, s'ils sont joués en même temps et dans la même salle. Les participants devaient se mettre sur un point fixe et être acoustiquement entouré par un nombre variable d'objets audio. Ils avaient le droit de tourner uniquement leurs têtes. La plupart d'entre eux ont été en mesure de percevoir simultanément jusqu'à à six sons à la fois. Au-delà de ce nombre, la localisation des objets audio et leur distinction semblaient être difficile. Pour cette raison, lors de la conception du scénario, le concepteur doit définir et ajuster le nombre maximum d'objets audio qui vont être perçus simultanément.

3.2.5 Création de contenus sonores

La création du contenu sonore pour cette première phase a été effectuée par un compositeur de son³, qui s'est chargé de préparer le contenu. La préparation repose sur l'enregistrement des sons des objets de l'expérience, ou la création de sons similaires en mixant plusieurs sons, ou en créant des ambiances sonores qui donnent une idée sur l'objet. Dans cette expérience, nous avons choisi une petite salle du musée qui comporte trois objets (figure 3.4) : une grande machine électrostatique du 18ème siècle, un impressionnant moteur de bateau du 19ème siècle. Les deux objets sont placés l'un devant l'autre à l'est et à l'ouest de l'entrée de la salle. Le troisième objet est l'entrée dans une seconde grande salle dédiée aux médias. Elle est situé au nord de l'entrée de la salle. Trois sons ambiants devraient être entendus dans la salle : deux sons provenant des deux machines et un son composé en provenance de la salle des médias. Aucune de ces deux machines ne peut être mise en marche, car elles sont très vieilles,

3. Cécile Leprado

de ce fait, nous ne pouvions pas enregistrer aucun son dans cette salle. L'ingénieur du son s'est chargé par la suite de composer des sons artificiels correspondants à chaque objet de la salle. Le son du moteur du navire a été créé en sur-échantillonnant, ralentissant, et effectuant un filtrage passe-bas à un son d'un petit bateau contemporain. Le son de la machine électrostatique a été, à son tour, créé en mélangeant un son illustrant le frottement mécanique des disques de verre et de la manivelle, avec le son de certains éclairs électriques. Le son provenant de la salle des médias a été créé par le mélange des sons de certaines technologies des médias connus tels que des imprimantes et des télégraphes. La création du contenu sonore reste tout de même une discipline à part entière qui n'a pas été étudiée comme problématique de recherche dans cette thèse.

3.3 Captation de mouvement

De manière générale, la captation du mouvement est employée dans les environnements virtuels pour enregistrer et suivre dans l'espace la position et l'orientation des objets physiques. A la différence de quelques autres systèmes similaires de RAS, qui ont consacré un large espace dans leurs travaux à la problématique de captation de mouvement du visiteur (voir chapitre 2), nous n'avons pas étudié la captation en tant que problématique de recherche dans notre travail. Nous avons plutôt testé plusieurs capteurs, en notre possession, tout en se concentrant sur la captation de l'orientation plutôt que sur celle de la position.

3.3.1 Polhemus Patriot

Le premier capteur utilisé est le capteur *PATRIOT*⁴ de la société *Polhemus*⁵, qui permet le suivi de la position et de l'orientation. Bien qu'il s'agisse d'un capteur filaire, inadapté à une visite de musée, *PATRIOT* a été utilisé en premier lieu afin de valider, dans une situation réelle de captation de mouvement, le système de spatialisation sonore développé. *PATRIOT* comprend, un système d'unité électronique (SEU), une source magnétique, un capteur et un boîtier de contrôle.

Lorsque *PATRIOT* est en fonction, la source émet successivement trois champs magnétiques mutuellement perpendiculaires qui induisent un courant dans les trois bobines du capteur. A partir de l'intensité de ces courants, qui dépend de la position relative du récepteur par rapport à la source, le système calcule la position et l'orientation du capteur. Le champ magnétique étant de faible intensité, il n'est pas possible d'utiliser le système pour le suivi dans des grands espaces. Le *PATRIOT*, qui est caractérisé par une portée utile (Useful Range) réduite à 60 inches, offre ainsi un espace de travail de la forme d'une sphère de rayon égale à environ 152 mètres. *PATRIOT* est fournit

4. http://www.polhemus.com/?page=motion_patriot

5. <http://www.Polhemus.com>

avec un logiciel *Polhemus PiMgrGUI* qui permet la gestion et le transfert des mesures de position et d'orientation du capteur. Ce logiciel est adapté uniquement aux systèmes d'exploitation Windows 2000 et XP.

Pendant nos expérimentations informelles (figure 3.5), le temps de latence entre le changement de position et de l'orientation de l'utilisateur et la mise à jour du contenu audio est estimé à 17 millisecondes. Cette valeur est largement inférieure à la perception auditive humaine estimée à 50 millisecondes. En outre, la vitesse du mouvement de l'auditeur n'affecte pas la performance acoustique du système.



FIGURE 3.5 – Expérience informelle menée au laboratoire

Après avoir passé la première étape de validation de notre prototype avec le *PATRIOT*, l'étape suivante est de choisir un capteur approprié qui va assurer l'auto-localisation. Ce qui signifie, que le visiteur doit porter sur lui tout le système de localisation, est ce dernier communiquera directement les mesures de localisation au système de spatialisation sonore. Ceci va permettre à plusieurs visiteurs d'utiliser le dispositif simultanément. Nous nous sommes ainsi orientés vers les capteurs inertiels.

3.3.2 Centrale inertielle Sparkfun IMU 6 DOF

Les capteurs inertiels tirent leur intérêt de l'utilisation conjointe d'accéléromètres et de gyroscopes. L'orientation de l'objet sur lequel le capteur est posé, est calculée en intégrant conjointement les valeurs des gyroscopes, ces dernières étant proportionnelles à la vitesse angulaire autour de chaque axe. Ces capteurs sont rapides, ne nécessitent pas d'éléments extérieurs, et ne souffrent d'aucun problème d'occlusion par un autre objet, ni d'interférences magnétiques, et occupent un espace très réduit. Cependant, ils ne sont pas précis lorsque les changements de positions sont lents, et restent sensibles à des phénomènes de dérives, puisque les erreurs s'accumulent en permanence.

Plusieurs modèles de centrales inertielles existent sur le marché. Nous pouvons citer les produits des sociétés Xsens [B.V.09], Intersense [Inc.10], Sparkfun [Sparkfun07], etc. Nous avons choisie la *IMU-6DOF-V4* de Sparkfun (figure 3.6) vu qu'elle est peu encombrante, légère, ouverte, documentée et moins onéreuse. En revanche, elle est vendue sans logiciel de traitement de mesures.

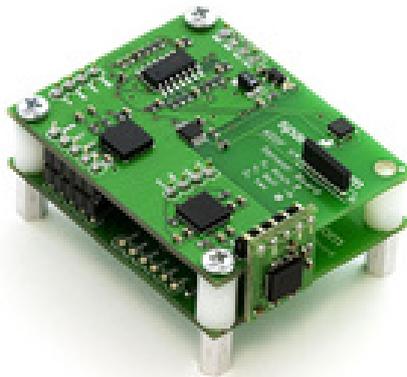


FIGURE 3.6 – Sparkfun IMU-6DOF-V4 [Sparkfun07]

Sparkfun *IMU-6DOF-V4* combine différents capteurs inertiels. Elle comporte :

- Un accéléromètre à triple axe Freescale MMA7260Q,
- Deux Gyroscopes InvenSense IDG500,
- Deux capteurs magnétiques Honeywell HMC1052L et HMC1051Z.

La centrale inertielle délivre des données brutes en 3 dimensions : vitesse angulaire, accélération et champ magnétique. En plus de la connexion filaire, elle supporte aussi la connexion Bluetooth.

Nous avons utilisé Sparkfun *IMU-6DOF-V4* dans le cadre de nos expérimentations informelles réalisées au MAM (figure 3.7). Cependant, la performance de la centrale inertielle est limitée en terme de précision et de stabilité du biais sur les mesures, ce qui est le principal facteur de dérive des techniques de navigation à guidage inertielle. Plusieurs essais ont été effectués afin de réaliser le calibrage du capteur et la conversion des mesures brutes vers des mesures fiables d'angles d'orientation (Voir Annexe A.1).



FIGURE 3.7 – Expérience informelle en utilisant Sparkfun IMU-6DOF-V4.

3.3.3 InertiaCube

Le dernier capteur d'orientation que nous avons utilisé est le *InertiaCube BT*⁶ de la société Intersense (figure 3.10). Il fournit en temps réel les données de l'orientation via une interface Bluetooth standard. L'*InertiaCube BT* est un capteur à trois degrés de liberté, permettant de mesurer les trois angles d'Euler : le lacet, l'élévation et le roulis, avec une précision annoncée de 1° pour le lacet et 0,5° pour l'élévation et le roulis. La latence moyenne du système serait de 40ms, mais cette dernière peut également varier selon l'équipement sans fil utilisé, à l'OS et à la connexion Bluetooth. Le capteur a une autonomie de 8 à 10 heures et a une connexion Bluetooth avec une portée allant jusqu'à 33m. Le capteur est livré avec un logiciel ISDemo permettant la gestion des mesures du capteur. Il permet de diffuser les données de captation sur le réseau en utilisant le protocole UDP. Il permet également de sauvegarder les données dans un fichier tampon

6. <http://www.intersense.com/pages/18/60/>

en local au fur est à mesure de leur réception. C'est cette dernière solution qui sera utilisée et maintenue pour nos prochaines expérimentations formelles.

3.4 Dispositif expérimental final

Après avoir réalisé de multiples expériences informelles, et testé quelques capteurs de mouvements, nous avons réalisé un dispositif du visiteur (logiciel et matériel) sur la base de nos travaux théoriques décrits au chapitre *Model* (chapitre 4). Ce même dispositif a été évalué durant des expérimentations formelles décrites au chapitre *Évaluation* (chapitre 5). L'architecture générale du dispositif est décrite par le schéma présenté par la figure 3.8.

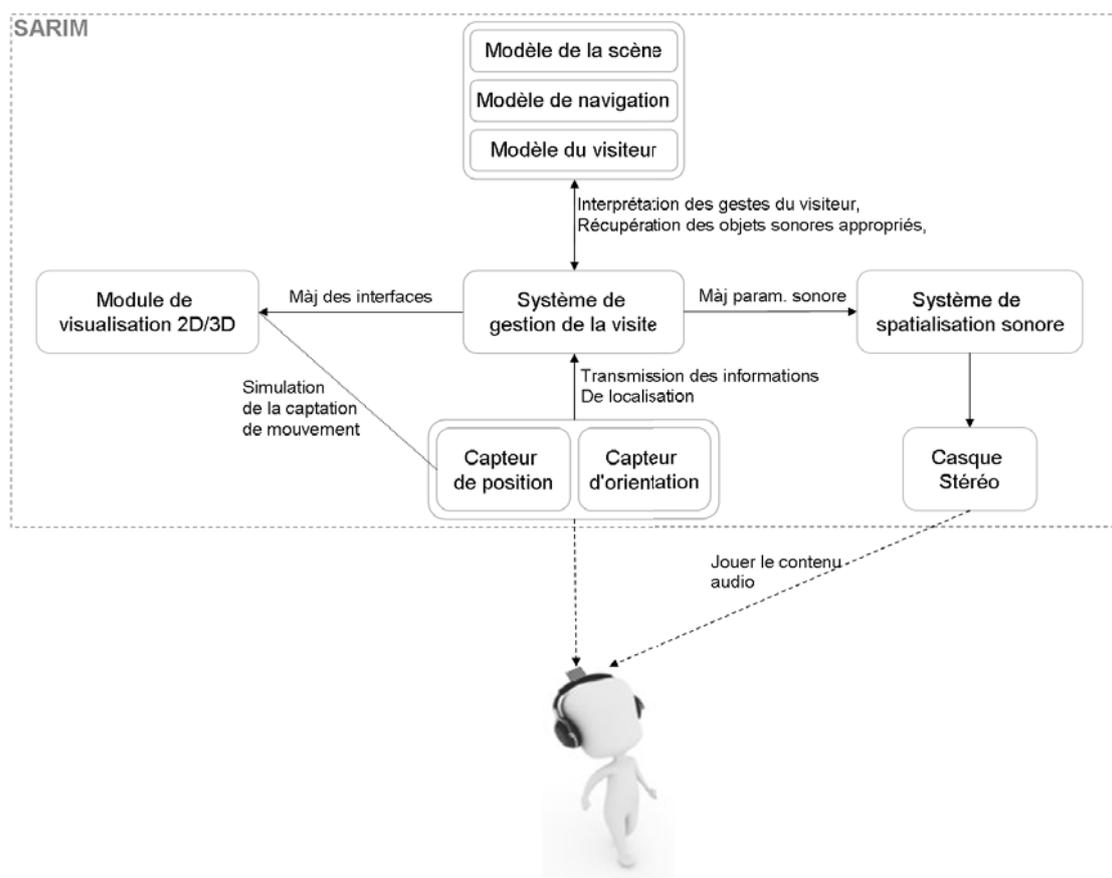


FIGURE 3.8 – Architecture générale du dispositif du visiteur de SARIM.

3.4.1 Localisation simulée (magicien d'Oz)

Puisque nous ne faisons pas de captation de position pour traquer la position des participants, nous avons décidé de simuler l'action de la captation de position par la technique du magicien d'Oz. Le magicien d'Oz est une technique expérimentale. Elle est utilisée pour simuler les fonctionnalités absentes d'un système existant, ou dans un système réel inexistant ou partiellement développé. Elle peut être utilisée également quand la technique est difficile à mettre en place. Dans la technique du magicien d'Oz, les participants sont mis en présence d'un système qu'ils croient réel, où une action ou plusieurs actions sont effectuées par un "magicien" caché. Le *comparse* doit, tout en restant à distance raisonnable du participant, à la fois observer l'interaction opérée par celui-ci et jouer son rôle, sans pour autant interférer avec la tâche du participant. L'utilisateur pense tester une interface réelle, alors que c'est une personne qui renvoie à l'utilisateur les réponses de l'interface. Dans notre expérience, cette technique a été utilisée pour simuler la captation de la position du participant. La simulation consiste à ce que l'expérimentateur porte l'ordinateur sur lequel l'application de simulation est installée. Il simule la navigation du visiteur en faisant déplacer un avatar qui représente le visiteur, au même temps que celui-ci se déplace. Par conséquent, le visiteur entend le son qui se déplace avec lui sans se rendre compte de ce que fait l'expérimentateur. Pour le visiteur, sa position est traquée automatiquement par le capteur qu'il porte. A la fin de l'expérience, une entrevue est menée par le responsable avec chaque participant à l'expérimentation afin de déterminer si celui-ci s'est aperçu de la manipulation et à quel moment. Il faut signaler qu'aucun participant à cette expérience n'a perçu que l'interaction était simulée par le comparse.

3.4.2 Équipement matériel

Nous utilisons un casque⁷ stéréo bluetooth de la marque Philips, modèle SHB9100, de dimensions 14,9 x 17,8 x 7 cm et d'un poids de 0,15kg, avec une portée de 15m et une autonomie de plus de 10h. Le format spatialisé du contenu audio qui est envoyé au visiteur justifie le caractère stéréo du casque. Nous avons choisi un casque Bluetooth pour ne pas encombrer les participants avec les fils. Le port du casque est confortable grâce aux coussinets, et peut être porté durant toute la visite du musée. Ce modèle de casque est aussi doté d'un arceau assez large permettant d'accueillir notre capteur d'orientation *InertiaCube*.

Le traitement des données de géolocalisation est effectué par un logiciel installé sur un ordinateur. L'ordinateur doit être porté par le visiteur au cours de sa visite. L'ordinateur où est installé notre application doit être équipé de la technologie Bluetooth pour pouvoir communiquer avec le casque et le capteur d'orientation, d'une carte graphique

7. http://www.philips.fr/c/ecouteurs/supra-aural-noir-shb9100_00/prd/

et d'une carte son, d'un processeur puissant, d'une autonomie raisonnable compatible au besoins de l'application et doit être léger (de préférence un netbook).



FIGURE 3.9 – Dispositif SARIM



FIGURE 3.10 – Capteur InertiaCube BT [Inc.10]

3.4.3 Architecture logicielle

Un logiciel a été conçu pour effectuer les tests de SARIM. Ces tests ont été réalisés en se servant du capteur *InertiaCube* qui permet de capter l'orientation de son porteur mais pas sa position. Il fallait intégrer au logiciel de test un outil permettant de simuler la position suivant la méthode du magicien d'Oz. Le logiciel que nous avons développé est composé de quatre modules principaux :

- La captation de mouvement.
- La spatialisation sonore.
- L'affichage graphique.
- La collecte d'informations.
- La lecture audio.

Le module de captation de mouvement est responsable de la récupération des données de l'orientation de la tête transmises par le capteur *InertiaCube*. Notre module se charge ainsi de récupérer ces données produits par ISDemo à partir de ce fichier tampon, d'en extraire les trois angles d'Euler (le lacet, le tangage et le roulis) définissant l'orientation du capteur, puis d'appliquer un filtre de Kalman sur chacune de ces trois valeurs. Ce module s'exécute d'une manière parallèle aux autres modules à travers la technique du multithreading. Ceci permet une communication plus efficace et évite le

blocage des autres modules dû aux latences imprévisibles susceptibles de se produire pendant la récupération du signal.

Le module de spatialisation sonore s'occupe lui de la spatialisation du son transmis à l'auditeur suivant sa position et son orientation. Le module se base sur la version native de la bibliothèque FMOD conçue pour Java. Le module offre plusieurs modes de spatialisation sonore :

La mode de spatialisation simple qui modifie la nature du son transmis à l'utilisateur suivant sa position et son orientation.

La sonorisation par l'orientation dans ce scénario, l'auditeur écoute uniquement le son de l'objet sonore vers lequel il est orienté.

Le mode de spatialisation composée dans ce cas, la spatialisation prend en compte la dualité zone d'audibilité partagée/exclusive (voir la définition des zones d'audibilité à la section 4.5.2). Elle permet d'arrêter les sons des zones partagées quand l'auditeur est en train d'écouter le son d'une zone exclusive. Ce mode de spatialisation a été utilisé pour l'expérimentation 1 (section 5.3).

La combinaison de la spatialisation et de la sonorisation par l'orientation ici l'auditeur commence par écouter tous les contenus sonores suivant le scénario de la spatialisation simple, puis dès que l'auditeur se focalise vers un objet sonore au delà d'un certain seuil temporel (fixé à 3 secondes), il commence à entendre uniquement le son provenant de cet objet. Puis passé un deuxième seuil temporel, l'auditeur entend le contenu audio correspondant à la description de cet objet. Ce mode de spatialisation a été utilisé pour l'expérimentation 2 (section 5.4).

De la même manière que le module précédent, ce module fait usage de la technique de multithreading en faisant en sorte que chaque objet sonore soit géré par un processus indépendant qui n'entrave pas l'exécution des autres traitements sur les autres objets sonores.

Le module d'affichage graphique est responsable de la création du rendu visuel 2D de la scène sonore, des zones d'audibilité exclusives et partagées liées à chacun des objets, et de la position et de l'orientation du visiteur en temps réel. Le module récupère l'orientation du visiteur à partir du module de captation de mouvement. Il offre au meneur de l'exécution (dans notre cas l'expérimentateur) la possibilité de spécifier manuellement la position du visiteur et ainsi simuler son déplacement au sein de l'environnement sonore suivant la méthode du magicien d'Oz (Voir Image 5.1).

Le module offre deux méthodes pour spécifier la position du visiteur. La première méthode fait usage des flèches directionnelles du clavier pour permettre à l'utilisateur de simuler le déplacement continu du visiteur au sein de la scène sonore. La deuxième méthode permet à travers un clic avec la souris de positionner directement le visiteur sur un endroit particulier dans la scène sonore. Ce module est développé suivant le patron de développement logiciel Modèle-vue-contrôleur (MVC) bien adapté pour ré-

pondre aux besoins des applications interactives en séparant les problématiques liées aux trois composants.

Le quatrième module est le module de la collecte des informations de l'expérience. Ce module, développé également suivant la technique du multithreading, traque en temps réel l'accès du visiteur aux différentes zones d'audibilité. Il permet ainsi après une visite de produire des fichiers de type tableur retraçant le parcours du visiteur au sein de la scène sonore, et marquant les temps d'entrée et de sortie aux différentes zones visitées.

Pour le programme que nous avons développé, nous avons utilisé le langage de programmation *java* et la librairie *NativeFmod* pour le rendu sonore. le programme comporte 8966 lignes de code.

3.5 Conclusion

Ce chapitre décrit le dispositif du visiteur et les étapes par lesquelles nous sommes passés afin de réaliser notre objectif de base qui concerne l'augmentation de la visite du musée par la dimension sonore. Nous avons choisit la visite du musée des arts et métiers comme cas d'étude pour les raisons suivantes :

- Le manque constaté de l'information sonore dans la présentation des œuvres du musée.
- L'importance du son pour le type d'œuvres exposées.
- La localisation du musée par rapport à notre lieu de travail, et la facilité d'expérimenter dans ce musée (logistique).
- Le contenu sonore qui nous été délivré par le personnel du musée.

Nous avons présenté l'architecture matérielle et logicielle des premiers dispositifs utilisés pour les premières expérimentations informelles que nous avons menées, puis nous avons présenté le dispositif final adopté. Ce dispositif a été développé pour réaliser le système de la modélisation de la visite augmentée décrit au chapitre *Modèle* (chapitre 4), il a servit également pour réaliser les expérimentations et évaluations décrites au chapitre *Évaluation* (chapitre 5).

4

Modélisation de la visite augmentée

4.1 Introduction

Modéliser une visite augmentée par la dimension sonore repose sur la description de tous les composants qui entrent en jeu dans cette visite. Cela revient à modéliser la scène sonore, le visiteur et sa navigation dans la scène sonore. Dans la première partie du chapitre, nous illustrons notre objectif par des exemples de scénarios de visite qui ont motivé ce travail. Nous avons choisi des scénarios libres et d'autres prédéfinis à vocation ludo-éducatives. Comme scénario prédéfini, nous donnons l'exemple d'un jeu didactique se déroulant dans la salle des automates ainsi qu'une expérience dans le laboratoire de Lavoisier. Le parcours libre sera, illustré par une visite dans les salles de communication qui abritent plusieurs catégories d'ouvrages, une découverte du Fardier de Cugnot, ainsi qu'une visite dans l'escalier d'honneur (section 4.2). Ces scénarios ont pour but de guider la conception d'un modèle le plus général possible de scènes sonores et de scénarios d'interaction pour les visites augmentées. La deuxième partie du chapitre viendra traiter la problématique de la modélisation de la visite augmentée dans le but de personnaliser le contenu envoyé à chaque visiteur du musée. Cette section commence par présenter la problématique de la personnalisation du contenu en citant quelques projets phares. Ensuite nous passons à la présentation des modèles que nous proposons des différentes entités qui composent la visite : l'environnement de la scène sonore sous trois couches : physique, virtuelle et sémantique (section 4.5), le visiteur avec des informations sur son emplacement et ses gestes mais également sur ses préférences et ses intérêts (section 4.6), et finalement le modèle de navigation du visiteur dans le paysage sonore (section 4.7).

4.2 Scénarios de visite

Plusieurs types de scénarios de visite peuvent être proposés. Les scénarios dépendent du mode de la visite (visite libre ou prédéfinie) et de sa finalité (visite à vocation ludique ou éducative). La visite libre signifie que le visiteur possède un "degré de liberté" de navigation et qu'il peut choisir librement son parcours. La visite prédéfinie signifie que le visiteur est invité à suivre un parcours guidé et défini par une chronologie (un ordre). Toutefois, la visite prédéfinie peut avoir une part de liberté, comme la visite libre qui peut avoir des indices de guidage. Il en est de même pour les finalités. Une visite peut être à vocation éducative tout en offrant des éléments ou un environnement ludique, et réciproquement. Avec les paramètres mode de la visite et finalité, nous pouvons concevoir un nombre illimité de scénario de visite.

Nous illustrons des exemples de parcours libres et prédéfinis avec des vocations ludo-éducatives. Comme parcours prédéfini, nous donnons l'exemple d'un jeu didactique à la salle des automates ainsi qu'une expérience dans le laboratoire de Lavoisier. Le parcours libre, sera illustré par une visite dans les salles de communication qui abritent plusieurs catégories d'œuvres, et une découverte du Fardier de Cugnot, ainsi qu'une visite à l'escalier d'honneur .

Ces scénarios de visite peuvent être classés en quatre catégories : visite prédéfinie et éducative (didactique), visite prédéfinie et ludique, visite libre et éducative et visite libre et ludique. Pour la première catégorie de visites : visite prédéfinie didactique, l'objectif est de proposer un parcours où l'ordre de passage devant les œuvres est important à des fins éducatifs. Pour cela on peut privilégier des contenus pédagogiques notamment des enregistrements vocaux correspondant à des explications, des témoignages, ou autre. Nous avons retenu l'exemple du laboratoire de Lavoisier. Pour le parcours prédéfini ludique, nous proposons un jeu ludo-éducatif dans la salle des automates.

Jeu ludique parmi les automates

Le scénario proposé s'inscrit dans un parcours prédéfini à vocation ludo-éducative. Il s'agit d'un jeu proposé aux enfants qui visitent le musée, une occasion pour eux de connaître les automates en jouant. L'expérience permet à l'enfant d'entendre toutes les explications des automates incluses dans le parcours, et c'est également un moyen de s'assurer qu'il n'est pas passé à côté d'un automate phare du musée. En pratique, il est demandé à l'enfant d'identifier l'automate produisant une musique en s'en approchant. Pour faciliter cette identification, la musique est spatialisée : le son s'adapte à la position et à l'orientation de l'enfant. Ainsi, l'enfant est guidé par une musique localisée, ce qui facilite la localisation et l'identification de l'automate qui la produit. Une fois l'automate trouvé, l'enfant se voit décrire l'automate, puis, une autre musique se déclenche (un autre automate devant être localisé).



FIGURE 4.1 – Laboratoire de Lavoisier

Une expérience didactique dans le laboratoire de Lavoisier

Le laboratoire de Lavoisier (Figure 4.1) comporte des instruments scientifiques, notamment deux calorimètres, deux gazomètres, deux balances de précision. Ces instruments ont été utilisés par Antoine Lavoisier pour découvrir la composition de l'eau. Le scénario vise à jouer l'expérience de la découverte de la composition de l'eau, par le son, autrement dit, le visiteur reçoit des sons correspondant à chacune des étapes de l'expérience en entendant le son produit par l'instrument qui est utilisé dans cette étape.

Le scénario est une sorte de mise en scène sonore dédiée aux enfants, où des souris font de la chimie pour fabriquer de l'eau. Ce scénario vise à faire connaître aux enfants l'expérience de la découverte de la composition de l'eau menée par Antoine Lavoisier. Dans cette mise en scène, l'enfant entendra différents types de sons : bouillonnement, crépitement et gouttes d'eau ainsi que des commentaires audio. Les différentes étapes de l'expérience sont illustrées par un ordre de déclenchement de son qui provient de l'instrument utilisé dans cette étape. Le son est joué d'une manière spatialisée. Ainsi pour illustrer par exemple l'étape de pesage de l'hydrogène, le visiteur entend un son provenant du gazomètre se trouvant à sa droite, ce son est composé d'une séquence du son produit si on pose une masse sur le fléau d'une balance, puis une voix dit : *"on pèse de l'hydrogène"*. Ainsi de suite jusqu'à la fin de l'expérience. Le détail du scénario est donné dans l'annexe.



FIGURE 4.2 – Le fardier de Cugnot

Découvrir le Fardier de Cugnot

Le fardier de Cugnot (Figure 4.2) est une œuvre phare du musée, c'est une œuvre complexe qui produit différents sons associés à chacun de ses composants¹. Par exemple la chaudière produit un bruit, le piston et les roues aussi, l'idée est de faire profiter le visiteur des sons produits par ses composants. Ainsi le visiteur entendra le son produit par chaque partie du fardier comme s'il était en marche. En se rendant près du fardier, le visiteur entendra un son mixé. Le mixage est fait à partir des sons produits par les trois composants principaux du fardier : la chaudière, le piston, et les deux grosses roues. Si le visiteur s'approche d'un composant particulier, il n'entendra que le contenu sonore associé. S'il y reste pendant un certain temps, il entendra les commentaires audio associés décrivant le composant en question. Si le visiteur est entre deux ou plusieurs composants, il entendra le contenu audio du composant qu'il regarde.

Jouer dans l'escalier d'honneur

L'escalier d'honneur (Figure 4.3) du MAM comporte l'aéroplane de Clément Ader accroché au plafond, l'horloge de Constantin (première étage) ainsi que le moteur à vapeur de l'aéroplane (deuxième étage). En bas des escaliers (rez de chaussée), se trouve le fardier de Cugnot. Pour de nombreux visiteurs, cet escalier est un passage obligé entre les étages plutôt qu'un lieu d'exposition. Pour valoriser ce lieu, nous proposons le scénario suivant en associant des contenus sonores et des descriptions didactiques à

1. Une copie fonctionnelle du fardier de Cugnot a été réalisée par l'ENSAM et des entreprises françaises en 2011 : http://www.dailymotion.com/video/xilowl_econstruction_du_fardier_de_cugnot_tech.UXj7LL9DQ3h

ces œuvres. Au pied de l'escalier, le visiteur entend le bruit de l'aéroplane lui provenant par le haut. Lorsque le visiteur s'approche du premier étage (respectivement du deuxième étage), il entend le bruit de l'horloge (respectivement du moteur à vapeur de l'aéroplane) qui lui parvient de haut à gauche. S'il s'approche de l'œuvre, le volume du bruit augmente, et s'il s'y arrête, il entend la description vocale de l'œuvre. Au bruit des œuvres s'ajoute le pas d'une personne qui accompagne le visiteur selon qu'il monte ou descende les escaliers. Le visiteur a la possibilité d'être accompagné : il rajoute une personne en se plaçant à l'endroit où il l'imagine et en hochant sa tête du haut en bas deux fois. Sinon, il continue sa visite seul. Si le visiteur court, il entend un son d'essoufflement qui l'accompagne. Ainsi, le son produit dépend des mouvements du visiteur durant sa vitesse.



FIGURE 4.3 – Escalier d'honneur.

4.3 Personnalisation du contenu

Les systèmes qui assurent la personnalisation du contenu sont les systèmes qui répondent aux besoins de chaque utilisateur, prenant ainsi en compte par exemple ses intérêts, ses préférences et ses connaissances de base. Ces systèmes à personnalisation dynamique doivent être en mesure d'observer l'utilisateur et son comportement et faire des déductions et des prédictions au sujet de ses intérêts pour lui envoyer du contenu personnalisé approprié [Kobsa01].

La personnalisation dynamique en cours de la visite des musées nécessite des systèmes interactifs capables de suivre le visiteur pendant sa visite et d'adapter automatiquement le contenu audio qu'il reçoit selon son comportement.

Jusqu'à présent, les projets phares modélisant dynamiquement le profil et le comportement du visiteur sont les projets *Hips*, *Hippie* [Petrelli01, Petrelli99, Benelli99, Specht99, Oppermann99], *peach* [Stock07], *ec(h)o* [Wakkary07, Hatala04] et *Listen* [Zimmermann08]. Ces systèmes proposent des modèles dynamiques qui se mettent à jour automatiquement, selon trois critères : le style du visiteur, l'historique de la visite et les intérêts du visiteur.

Le style du visiteur : le type de visiteurs dans le contexte du musée a été largement étudié dans la littérature [Dean94, Serrell96, Sparacino02]. Véron et Levasserrur [Véron96] classifient les visiteurs selon leurs manières d'explorer une exposition en les associant à des animaux. Ils déduisent ainsi *le style de la fourmi* qui suit le chemin du conservateur, puis *le style du poisson* qui a une vue holistique (globale), puis *le style du papillon* qui s'intéresse à toute l'exposition sans suivre le chemin prédéfini par le conservateur, et finalement *le style de la sauterelle* qui s'intéresse uniquement à des expositions spécifiques. Le projet *Listen* [Zimmermann08] adopte cette approche et définit le style du mouvement du visiteur comme un comportement stéréotypé d'un visiteur du musée en définissant quatre styles de visiteur : le visiteur déambulant qui marche doucement avec un regard promeneur, le visiteur guidé par des objectifs qui a un mouvement direct avec la vue dirigée vers un artefact spécifique, le visiteur debout dont le regard reste fixe sur un objet, puis le visiteur debout flou qui demeure debout ou assis et qui est caractérisé par un regard promeneur. Comme dans *Listen*, le type du visiteur est calculé à partir de la vitesse et de l'uniformité de son mouvement. Le projet *ec(h)o* [Wakkary07, Hatala04] classe les visiteurs en trois types également. Il définit le visiteur avare, qui n'est pas précipité, et qui aime connaître le plus possible d'information, le visiteur sélectif qui explore uniquement les objets qu'ils l'intéressent, et le visiteur occupé qui préfère faire un tour dans le musée et avoir une idée générale de ce qui est exposé.

L'historique de la visite : l'historique de la visite se construit à partir des objets rencontrés par le visiteur et des sons affiliés à ces objets. Il représente le comportement du visiteur, manifesté par son mouvement et/ou sa sélection d'objets. Dans les projets *Hippie* et *ec(h)o*, le visiteur est muni d'une interface tangible (cube en bois avec un bouton de sélection dans *ec(h)o*, et un PDA dans *HyperAudio*, *Hips* et *Hippie*) qui permet au visiteur de sélectionner les objets qui l'intéresse. Concrètement, dans *ec(h)o*, le système stocke les données de l'historique sous forme d'ensemble de points discrets (espace-temps). Dans *Listen* l'historique de la visite dépend de l'œuvre que le visiteur regarde à chaque moment de la visite, et de l'entité sonore que ce dernier écoute. Dans chaque unité de temps, les informations selon ses deux entités sont stockées en base de données.

L'intérêt du visiteur : afin d'individualiser et de personnaliser le contenu audio pour chaque visiteur, les chercheurs se sont interrogés sur les intérêts des visiteurs, autrement dit comment détecter les objets du musée qui intéressent plus les vi-

siteurs. Les projets *ec(h)o* et *Listen* adoptent une approche dynamique de l'intérêt des visiteurs. En effet, dans le projet *ec(h)o*, l'intérêt du visiteur sont représenté comme des concepts pondérés de l'ontologie des concepts, chaque objet exposé est annoté avec une suite de concepts de la même ontologie. Quant au projet *Listen*, il adopte l'approche heuristique de Pazzani [Pazzani97] qui suppose que plus le visiteur passe du temps devant un objet, plus il l'intéresse. Toutefois, cette hypothèse peut ne pas être vérifiée car il peut y avoir plusieurs raisons pour lesquelles l'utilisateur s'arrête quelque part (présence d'obstacles par exemple). Ainsi, les déductions erronées peuvent influencer négativement sur la détermination de l'intérêt de l'utilisateur. Pour prévenir de cette limitation, le projet *ec(h)o* [Wakkary07, Hatala04], propose un modèle d'utilisateur permettant au visiteur la sélection manuelle des objets intéressants, en plus de la captation de la position et l'orientation des visiteur. En effet, le visiteur peut sélectionner les objets qui l'intéressent et dont il veut écouter le contenu grâce aux boutons se trouvant sur une interface tangible que le visiteur porte durant sa visite. Le fait de sélectionner manuellement des objets peut en effet aider à déterminer avec précision les intérêts des visiteurs, mais l'idée d'avoir les mains libres proposée par *Listen* se justifie aussi par la volonté de ne pas encombrer le visiteur et de ne pas préoccuper son attention et ses yeux avec des boutons et une interface et de le détourner finalement de son but essentielle qui est de regarder l'œuvre exposée.

Afin de trouver un compromis entre la méthode heuristique et la méthode manuelle pour la détermination de l'intérêt du visiteur, nous optons pour une troisième voie pour l'expression des intérêts qui permet de minimiser au maximum le degré d'attention requise de la part de l'utilisateur pour interagir avec son environnement.

Le modèle SARIM que nous présentons permet aux visiteurs, qui le souhaitent, d'exprimer leurs intérêts par le moyen de gestes simples et intuitives de tête. Bien que nous nous basant également sur la méthode heuristique, ces gestes qui permettent au visiteur de contrôler la scène sonore seront pris en compte dans la déduction et ainsi la personnalisation de son parcours.

Le contrôle de la scène sonore par les gestes constitue l'une des contributions de ce chapitre. En outre la structure du modèle général de la visite augmentée proposée est également novatrice dans la mesure où, une telle modélisation des acteurs majeurs de la visite augmentée n'a pas été proposée auparavant. C'est une structure qui sépare d'une manière modulaire et efficace les trois parties du modèle ; la scène sonore, le visiteur et la navigation. Chacune de ces parties, est elle même structurée en plusieurs couches. Ceci permet une gestion indépendante des différentes parties, et facilite la paramétrisation du système. En effet bien que notre système permet trois niveaux différents de complexité de navigation (la traque, le contrôle et l'adaptation), il est devient plus aisé de choisir le niveau de navigation le plus approprié à chaque visiteur.

La troisième contribution est l'introduction du concept de Zones d'Audibilité (ZA) qui permettent une meilleur structuration de la scène sonore. Avec leurs caractéris-

tiques géométriques et conceptuelles interactives et hiérarchiques, ces ZA permettent de créer des scénarios à la fois riches et surprenants.

4.4 Description du modèle

La conception d'une visite augmentée par la dimension sonore nécessite le développement d'un modèle décrit à travers ses différentes composantes. Trois entités essentielles composent ce modèle ; la *scène sonore*, le *visiteur*, et la *navigation* faisant interagir les deux premiers.

La *scène sonore* est l'entité décrivant l'environnement du musée. Elle est constituée de trois différents niveaux d'abstraction. Le premier niveau représente l'environnement physique du musée. Ce niveau est augmenté par un deuxième qui représente l'environnement sonore virtuel. Ce dernier est lui même augmenté par un troisième niveau représentant les caractéristiques sémantiques liées aux objets exposés et aux contenus audio qui leurs sont attribués.

Pour sa part, le *visiteur* est modélisé à travers différents descripteurs personnels et environnementaux qui le caractérisent. Ces caractéristiques sont réparties sur plusieurs niveaux permettant de capter l'emplacement du visiteur, son comportement et ses préférences et intérêts déduits à partir de l'analyse de son parcours au sein du musée.

Les différents gestes et déplacements du visiteur peuvent être interprétés en tant qu'interactions avec la scène sonore à travers l'entité du *navigation*. Ce troisième composant relie la *scène sonore* et le *visiteur* et permet de définir différents modes d'interactions entre les deux. La complexité des interactions varie de l'analyse simple de la position et de l'orientation du visiteur (niveau traque) pour effectuer la spatialisation sonore, jusqu'à l'adaptation instantanée de la scène sonore (niveau adaptation) suivant les préférences du visiteur, en passant par l'interprétation des gestes du visiteur pour contrôler la scène sonore (niveau contrôle).

L'architecture globale du système est décrite par le schéma présenté dans la figure 4.4.

L'architecture globale de notre système diffère de celle proposée dans le cadre des projets *Hips*, *ec(h)o* et *Listen* ou encore *Deep Map*, par la séparation des différents acteurs majeurs de la visite augmentée. En effet, si le projet *Deep Map* propose un modèle visiteur pour personnaliser le contenu, il ne propose pas de modèle de scène sonore. Les projets *Hips*, *ec(h)o* et *Listen*, quant à eux, proposent des modèles de scène sonore et un modèle visiteur qui mêlent les données concernant le visiteur et sa navigation dans le musée. Notre modèle de visite vise à définir séparément les trois acteurs principaux, à savoir la scène sonore, le visiteur et la navigation. Tout en proposant une spécification pour chaque acteur.

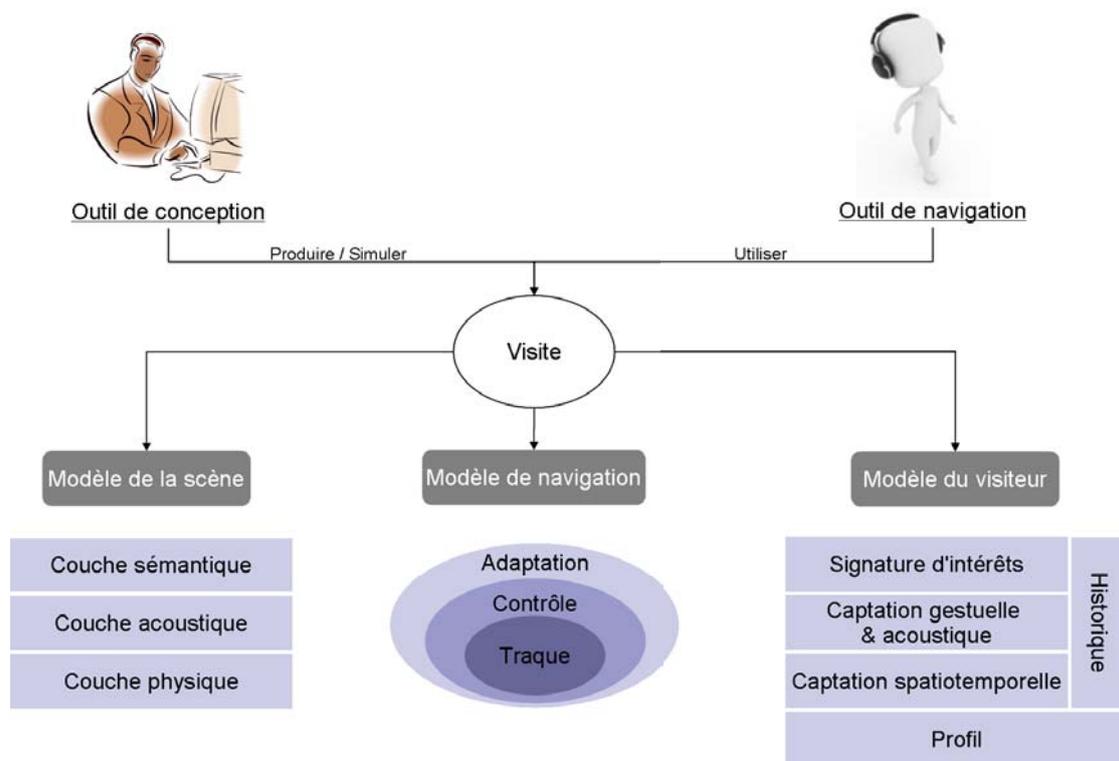


FIGURE 4.4 – Représentation de la structure de la visite

4.5 Scène sonore

Le paysage sonore (en anglais *soundscape*), est un néologisme subjectif, résultant de la contraction des mots *landscape* et *sound*. A l'origine, il fut défini par le canadien Robert Murray Schafer comme un environnement de sons. Ce terme s'applique aussi bien à un environnements physique (au sens réel) qu'à des constructions abstraites, telles que les compositions musicales ou les montages sur bande, en particulier lorsqu'ils sont considérés comme faisant partie du cadre de vie [Schafer10, Bosseur96].

Dans un musée proposant un système d'audioguidage, la distanciation entre l'espace réel et l'espace virtuel est accentuée du fait de commentaires audio décrivant des groupes d'œuvres, leurs histoires, ou par le son qu'ils produisent. Ainsi, l'entrechoquement d'éprouvettes et les bouillonnements emplissent le laboratoire de Lavoisier au musée du MAM et transportent le visiteur au sein d'une expérience scientifique, ce qui peut faciliter la compréhension non seulement des d'œuvres exposées, mais aussi de leurs histoires, de leurs fonctionnements ou de leurs utilisations.

En pratique, ce découplage entre espace réel et espace virtuel est rendu possible grâce à un audioguide équipant le visiteur et incluant plus particulièrement un casque audio. Ce casque offre la flexibilité nécessaire à la création d'un paysage sonore n'obéissant pas forcément aux lois de la nature et non soumis à des distorsions sonores, et au bruit d'environnement. Le contenu audio est ainsi contrôlable et personnalisable. D'un point de vue architectural, et comme proposé par les modèles de *Hips*, et *Listen* le paysage sonore peut être décomposé en plusieurs niveaux.

Dans notre modèle, nous décomposant la scène sonore en trois couches interdépendantes :

La couche physique décrit l'environnement physique où le visiteur peut déambuler et interagir avec le système. Elle contient toutes les informations géométriques sur l'espace physique et les objets qui le composent.

La couche virtuelle (ou paysage sonore) est organisée sous une forme hiérarchique qui décrit des zones virtuelles sonores.

La couche sémantique correspond à la description sémantique des zones et du son, et à leur annotations qui sont organisées dans un thésaurus.

4.5.1 Couche physique

La couche physique est le premier niveau de description de l'environnement du musée. Elle permet de représenter le modèle géométrique du musée, localiser les salles d'exposition, leurs accès et leurs fenêtres. Elle permet également de positionner les objets exposés en spécifiant leurs coordonnées, leurs dimensions et leurs formes. Elle correspond à la couche dite *Word model* dans les projets *Hips* et *Listen*.

4.5.2 Couche virtuelle

La notion de base du modèle acoustique du soundscape est la source sonore localisée dans l'espace. Lorsque la position du visiteur le permet, ce dernier perçoit les sons émis par la ou les sources sonores qui les entourent. Dans certains cas particuliers, il paraît utile de dépasser le modèle plaquant strictement la réalité physique. En effet, il pourrait être intéressant de créer des espaces sonores dont la distribution et la forme ne reflète pas la propagation sphérique d'une onde sonore issue d'une source ponctuelle.

Pour ceci, et afin de maîtriser l'espace d'écoute du visiteur, nous introduisons la notion de *Zone d'Audibilité (ZA)* qui délimite l'espace dans lequel un son particulier est audible. L'intérêt de cette approche par *ZA* est la possibilité de personnaliser l'espace d'audibilité lié à une source sonore selon le besoin de l'application. Ainsi, il est possible de créer des zones d'audibilité non conventionnelles (sphériques) telles que des espaces de formes parallélogrammes ou cylindriques qui s'adaptent au mieux à la géométrie du musée. L'autre intérêt de cette approche c'est qu'il peut a priori ne pas y avoir de lien entre les artefacts exposés dans le musée et les *ZA*, celles-ci peuvent ne correspondre à aucun objet du musée, ou plutôt à des objets "banals" présents dans le musée (porte de secours, escaliers, vestiaire, etc). La disposition des sources sonores au sein du musée peut être libre et aléatoire ou suivant une logique définie. Plusieurs objets exposés peuvent être réunis dans un seul groupe et être représentés par une seule *ZA* qui peut décrire la fonction ou le thème du groupe. Cette *ZA* peut également être divisée en plusieurs *ZA* plus petites représentant chacune un objet ou une partie d'un objet particulier. L'immersion dans un environnement de RAS peut être mieux appréciée lorsque les auditeurs sont capables d'apercevoir simultanément des sons provenant de diverses sources sonores. Ce qui se résume en la création d'une architecture de *ZA* dans laquelle plusieurs zones sonores se recouvrent. Ces exigences nous orientent vers l'adoption d'un modèle virtuel acoustique hiérarchique sous forme d'arbre pour la représentation des zones d'audibilité. Une zone d'audibilité se caractérise par deux éléments distincts : le son (**soundtoplay**) et l'espace (**area**). Cette dualité se retrouve au niveau de la spécification d'une zone d'audibilité exprimée dans le langage et formalisée à l'aide de schémas XML (listing 4.1).

Listing 4.1 – Schémas XML d'une zone d'audibilité

```

1 <xs:element name="audibilityzone">
2 <xs:complexType>
3 <xs:sequence>
4 <xs:element name="area" type="areatype" .../>
5 <xs:element name="soundtoplay" type="soundtoplaytype"
6     minOccurs="1" maxOccurs="1"/>
7 <xs:element name="perception" minOccurs="0" maxOccurs="1"
8     type="typeofperception" default="shared"/>
9 </xs:sequence>
10 <xs:attribute name="id" use="required" type="xs:ID"/>
11 <xs:attribute name="name" use="optional" type="xs:string"/>
12 <xs:attribute name="description" use="optional" type="xs:string"/>

```

```

13 <xs:attribute name="parent" use="optional" type="xs:IDREF"/>
14 </xs:complexType>
15 </xs:element>

```

4.5.2.1 Perception sonore

L'objectif d'associer un des objets du musée des ZA peut être différent selon la nature des objets exposés et le lieu dans lequel se positionne le visiteur. Dans certains cas, le but du concepteur de systèmes peut être d'immerger le visiteur dans un paysage de sons spatialisés émanant de divers objets sonorisés. Ceci peut être particulièrement intéressant en vue de recréer pour le visiteur l'impression d'être entouré par des machines en marche. Dans d'autres cas de figure, l'objectif pourrait être d'attirer l'attention du visiteur vers un objet en particulier et éviter qu'il soit attiré par d'autres sources sonores moins intéressantes. Ainsi, afin de gérer ces différentes configurations de perceptibilité des sources sonores, nous introduisons la notion de type de perception de la ZA (élément **perception** du listing 4.2 et de la ligne 6 du listing 4.1).

Nous définissons deux types de ZA :

- **zone partagée** (*shared*) : C'est une zone qui tolère les recouvrements sonores. Si cette zone se recouvre avec une autre zone du même type, les sons provenant des deux zones seront mixés dans leur espace de recouvrement. Le visiteur au sein de l'intersection des deux zones pourra percevoir les contenus audio provenant des deux sources sonores différentes.
- **zone exclusive** (*standalone*) : le son de cette zone ne peut être mixé avec un autre son provenant d'une autre source sonore.

Listing 4.2 – Perception

```

1 <xs:simpleType name="typeofperception">
2 <xs:restriction base="xs:string">
3 <xs:enumeration value="shared" />
4 <xs:enumeration value="standalone" />
5 </xs:restriction>
6 </xs:simpleType>

```

4.5.2.2 Géométrie de la zone d'audibilité

Une zone d'audibilité peut avoir une géométrie particulière au sein de laquelle un contenu audio spécifique est joué. L'espace (**area**) occupé par cette zone peut prendre différentes formes (**shape**). Plus généralement, comme le montre son schéma XML (listing 4.3), une surface (de type **areatype**) peut être dépeinte à partir d'un ensemble (lignes 1-6) de formes simples (de type **shapetype**). Une ZA prend la forme d'un sphéroïde, d'un prisme, d'un cylindre, etc (*spheroïde, prism, cylindroïde*).

Listing 4.3 – Schémas XML d'une surface

```
1 <xs:complexType name="areatype">
2   <xs:sequence>
3     <xs:element name="shape" type="shapetype" minOccurs="1"
4     minOccurs="unbounded"/>
5   </xs:sequence>
6 </xs:complexType>
7
8 <xs:complexType name="shapetype">
9   <xs:attribute name="idref" use="required" type="xs:IDREF"/>
10  <xs:attribute name="type" use="required" type="typeofshape"/>
11 </xs:complexType>
12
13 <xs:simpleType name="typeofshape">
14   <xs:restriction base="xs:string">
15     <xs:enumeration value="spheroïde"/>
16     <xs:enumeration value="prism" />
17     <xs:enumeration value="cylindroïde"/>
18   </xs:restriction>
19 </xs:simpleType>
```

4.5.2.3 Organisation des zones d'audibilité

Pour chaque objet, on peut attacher une zone dans laquelle un son associé à cet objet est perceptible par un visiteur dans la zone. Une zone peut aussi être associée à un groupe d'objets, le contenu sonore dans ce cas peut être la description de la fonction globale ou du thème de ce groupe. En outre, il peut être utile de pouvoir créer des ZA élémentaires à l'intérieur d'une ZA de groupe où chacune correspondant à une source sonore particulière. Il serait ainsi possible de décomposer la ZA du soundscape du musée en sous éléments correspondant aux salles de ce musée, puis aux îlots d'objets présentés dans les salles, puis éventuellement aux parties de ces objets, si ils sont volumineux et complexes. L'ensemble défini ainsi un arbre, semblable aux graphes de scènes utilisés en synthèse d'image.

L'ensemble des zones d'audibilité forment le paysage sonore dans lequel évolue le visiteur. Les zones d'audibilité peuvent se juxtaposer les unes avec les autres. Dans l'absolu, ces zones pourraient être organisées de différentes manières.

Chacune des zones est contenue dans une zone mère, à l'exception de la racine de l'arbre. La figure 4.5, illustre un exemple du modèle hiérarchique. Le Fardier est associé à une zone mère (*fardier*), qui contient des zones filles qui délimitent les parties du Fardier, notamment la chaudière, le piston, et les roulements. Chacune des zones produit un son particulier, et par conséquent mérite d'être traitée à part.

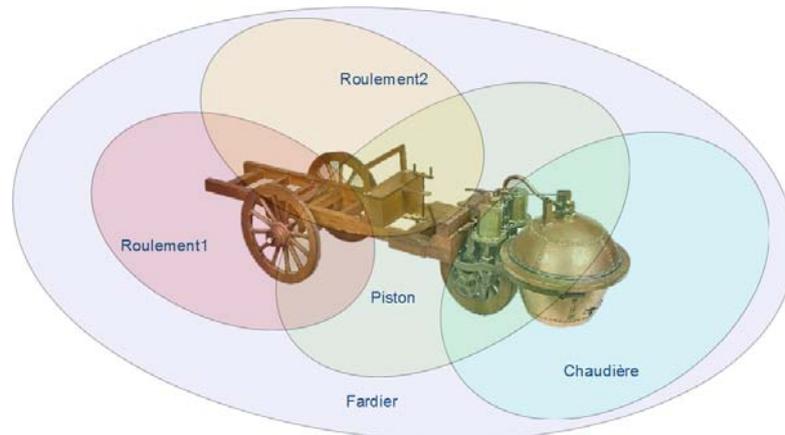


FIGURE 4.5 – Représentation hiérarchique des zones d'audibilité sur le fardier de Cugnot

4.5.2.4 Diffusion du son

Fondamentalement, la perception spatiale d'un son implique la présence d'un axe de référence qui dans notre cas correspond au visiteur : la spatialisation du son est adaptée en fonction de sa position et de son orientation.

Le point de référence à partir duquel sont calculées la direction (les angles) et la distance, se situe au niveau de la tête du visiteur. La source du son est quant à elle, localisée en fonction de la distance la séparant du visiteur (distance perçue) et des angles d'Euler calculés par rapport à son vecteur visuel : *lacet*, *élévation*, *roulis*. La perception d'une source est en cela multidimensionnelle. En pratique, un son se caractérise par une source², un médium dans lequel le son se propage et une destination (*listener*) correspondant ici au visiteur. Ce médium sous-entend un (voir plusieurs) chemin(s) de propagation entre la source et le visiteur. Le son est de nature omnidirectionnel : il se propage indépendamment de sa direction et tend à former un champ sphérique (le son étant identique sur tous les plans perpendiculaires à sa direction de propagation et s'atténue en fonction de la distance), et cela, à moins qu'une surface réfléchissante ou qu'un bruit n'altère sa propagation.

Listing 4.4 – Types de son

```

1 <xs:simpleType name="typeofsource">
2 <xs:restriction base="xs:string">
3 <xs:enumeration value="uniform"/>
4 <xs:enumeration value="spacialised"/>
5 </xs:restriction >
6 </xs:simpleType>

```

2. Bien que plusieurs sources puissent être à l'origine d'un son, il est plus simple de le décrire comme résultant d'une seule source.

Le type de diffusion du contenu audio à l'intérieur de la zone définit si le son est spatialisé ou non (valeurs *spatialised* ou *uniform* qualifiant le type de "sound" dans le listing 4.4). L'option spatialisé signifie que le contenu est délivré de manière binaurale. Il est caractérisé par une source dont la position est donnée (élément **location** exprimant les coordonnées cartésiennes de la source), ce qui signifie que les paramètres psycho-acoustiques du contenu sonore sont adaptés suivant la position de cette source.

Listing 4.5 – Types de son

```

1 <xs:complexType name="soundfromtype">
2 <xs:attribute name="name" use="optional" type="xs:string"/>
3 <xs:attribute name="id" use="required" type="xs:ID"/>
4 <xs:attribute name="type" use="required" type="typeofsource"/>
5 <xs:attribute name="location" use="optional"
6     type="locationtype" default="0 0 0"/>
7 </xs:complexType>

```

Au contraire, un son uniforme est omnidirectionnel, non localisé, et n'est donc pas délivré au format binaural. Cette diffusion est appropriée dans le cas d'un contenu audio lié à un ensemble d'objets ou à une salle du musée.

D'autres caractéristiques du son sont ici définies (comme par exemple, la répétition du son, et la latence entre deux répétitions).

L'ensemble des sons forme une collection sonore exploitée pour former un paysage sonore correspondant ici à une visite augmentée. Ces collections composées de zones d'audibilités peuvent être annotées de façon à former un catalogue. Une fois ses collections annotées, classifiées par thème et catégorisées, la description résultante pourra être exploitée pour identifier facilement des groupes distincts de zones d'audibilité. Et ainsi les classifier selon le choix du visiteur.

4.5.3 Couche sémantique

L'espace du musée se décline par sujet et thématique plutôt que par une organisation géométrique ou architecturale.

Chacune des zones d'audibilité peut être décrite par un ensemble de descripteurs correspondant à la fois à l'objet (aux objets) se trouvant dans cette zone, mais également au son diffusé dans cette zone. Certaines ZA peuvent se partager les mêmes descripteurs, surtout dans le cas où les objets associés à ces ZA assurent les mêmes fonctionnalités, ou quand les sons associés ont les mêmes caractéristiques. Il est ainsi essentiel d'assurer la compatibilité et l'interopérabilité du point de vue de la sémantique de ces descripteurs et de dépasser l'ambiguïté du langage naturel dans la description des ZA.

Un thésaurus constitue un élément utile pour décrire de façon structurée des informations sous la forme d'un ensemble de concepts représentés par des termes. Les termes d'un thésaurus sont organisés hiérarchiquement, ce qui permet d'adapter la pré-

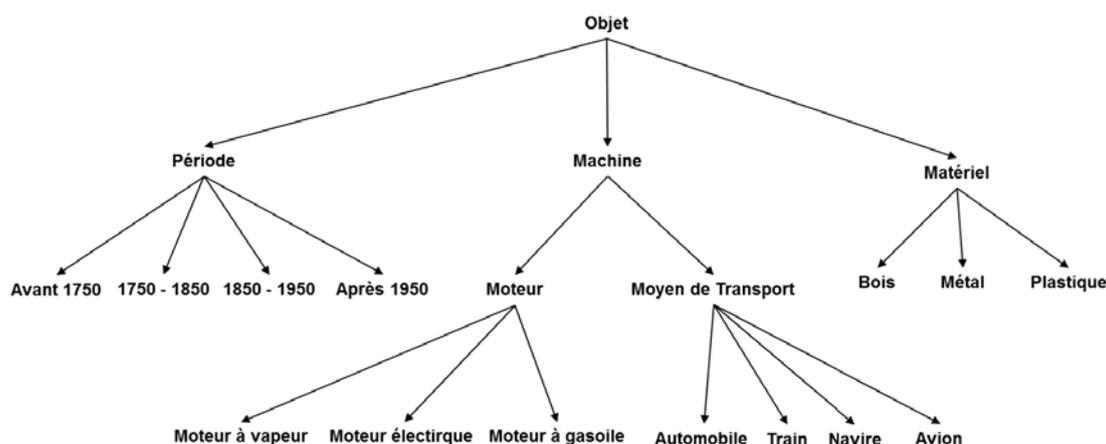


FIGURE 4.6 – Exemple d'un thésaurus

cision lors de l'indexation et de la recherche. Les thésaurus sont souvent utilisés pour indexer des documents ou des objets dans une banque de données bibliographiques. Se référer à un thésaurus existant permet d'homogénéiser l'indexation puisque celle-ci ne dépend plus de la culture ou du domaine de l'indexeur. De nombreux thésaurus ont été établis. Parmi les plus connus, nous pouvons citer Motbis³, un thésaurus édité par le Ministère de l'Éducation nationale en France et portant sur le domaine de l'enseignement général. Trois types de relations unissent les termes d'un thésaurus :

- une relation hiérarchique *stricto sensu* reliant un terme générique à un terme spécifique,
- une relation d'équivalence reliant deux termes synonymes,
- une relation d'association reliant deux termes sémantiquement connexes.

A des fins de simplification, et puisque les thésaurus utilisés seront développés pour un nombre limité d'objets, nous utilisons uniquement des relations hiérarchiques pour relier les termes. Nous utilisons la notion de thésaurus pour décrire la structure des concepts décrivant les objets des ZA et aussi les concepts décrivant les contenus audio. La figure 4.6 illustre un exemple de thésaurus composé d'un ensemble de concepts correspondant à une partie des objets exposés dans le musée, et la figure 4.7 représente un thésaurus permettant de donner une idée sur les concepts qui décrivent les différents contenus sonores proposés aux visiteurs.

Comme le montre le listing 4.6, un thésaurus portant sur les sujets ou les sons (**thesaurus**) est organisé hiérarchiquement ; chaque concept dispose donc d'un père (attribut **parent** du listing 4.7).

Listing 4.6 – Schémas XML d'un thésaurus

```
1 | <xs:complexType name="thesaurus">
```

3. <http://www.cndp.fr/thesaurus-motbis/site/index.php>

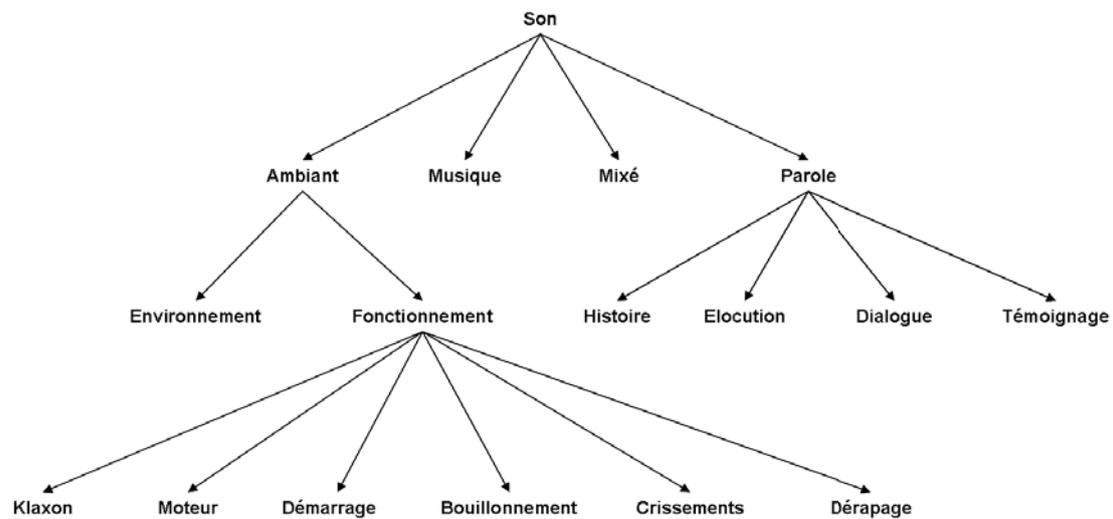


FIGURE 4.7 – Exemple d'un thésaurus du contenu sonore

```

2 <xs:sequence>
3 <xs:element name="concept" type="concepttype" minOccurs="1"
4   maxOccurs="unbounded"/>
5 </xs:sequence>
6 </xs:complexType>

```

Listing 4.7 – Schémas XML d'un concept

```

1 <xs:complexType name="concepttype">
2   <xs:attribute name="id" use="required" type="xs:ID"/>
3   <xs:attribute name="name" use="required" type="xs:string"/>
4   <xs:attribute name="parent" use="optional" type="xs:IDREF"/>
5   <xs:attribute name="weight" use="optional" type="xs:int"/>
6 </xs:complexType>

```

A titre d'illustration (listing 4.8), la zone d'audibilité correspondant au fardier de Cugnot est décrite par les concepts **automobile**, **transport**, **moteur à vapeur**, **1750-1850**.

Listing 4.8 – Description XML

```

1 <concepts>
2   <concept id="t1" name="automobile" parent="o2" weight=5/>
3   <concept id="mo1" name="moteur à vapeur" parent="ma1" weight=9/>
4   <concept id="p1" name="1750-1850" parent="o1" weight=10/>
5   <concept id="o2" name="Moyen de transport" parent="ma2" weight=9/>
6 </concepts>

```

La couche physique proposée dans notre modèle de visite est similaire au *Word model* dans les projets *Hips* et *Listen*, quant à la couche virtuelle, elle est basée sur la notion de zone d'audibilité (voir la section 4.5.2), qui constitue le noyau de cette couche. Une

ZA est définie comme le moyen pour délimiter, dans un espace 3D, la diffusion d'un son. C'est plus sa géométrie en 3D, la capacité d'organisation de ZA (intersection, et d'hierarchie), ses caractéristiques de diffusion du son (centralisé ou pas autour d'une source sonore), son type (partagée ou exclusive) qui fait de cette ZA une notion novatrice qui n'a pas été traitée auparavant dans la littérature. Par ailleurs, une fois annotées et classifiées par thèmes, ces collections sonores formées par l'ensemble des zones d'audibilités et des sons affiliés peuvent être exploitées sous la forme de catalogues (au travers d'un thésaurus). Ainsi, leur exploitation s'en trouve-t-elle facilitée.

4.6 Visiteur

Bien qu'il soit impossible de modéliser un visiteur dans son contexte général, il est nécessaire de proposer un modèle comportemental traçant son évolution au sein du musée.

Tout comme le projet *ec(h)o*, le modèle du visiteur dans le projet *Listen* se compose de trois parties : l'historique de la visite, le type du visiteur analysé à partir de son style de mouvement, et les intérêts du visiteur. Dans notre cas, afin d'étudier le comportement du visiteur, il est nécessaire de traquer son activité au sein du musée. C'est à dire son déplacement, ses arrêts, son focus (les objets qu'il regarde), et ses gestes pour pouvoir construire son profil.

Nous procédons à la création du modèle du visiteur suivant différents étapes et niveaux. Chaque étape se conclue par la définition de différents descripteurs qui permettent de distinguer les caractéristiques du visiteur (Figure 4.8) :

Profil du visiteur. A son arrivée au musée, le visiteur est invité à communiquer des informations le concernant (§4.6.1). Seules quelques informations sont demandées de façon à ce que ce feedback soit limité. A partir de ces informations, qui constituent le profil du visiteur, des hypothèses concernant les caractéristiques de celui-ci sont faites de façon à aiguiller le visiteur vers des parcours de visite qui pourront l'intéresser.

Captation spatiotemporelle. La position du visiteur au sein du musée est continuellement analysée sur la base de son déplacement, de son orientation et de sa vitesse du mouvement.

Captation gestuelle et acoustique. Sur la base d'un modèle formel de descriptions de gestes, les mesures issues du niveau de la captation spatiotemporelle sont analysées pour inférer les gestes que le visiteur réalise pendant son déplacement. Ses mesures sont également utilisées pour déduire la ZA dans laquelle se trouve le visiteur, ainsi que l'objet que le visiteur regarde. Ceci est réalisé en se basant sur le modèle virtuel acoustique de la section (§4.5.2)

Signature d'intérêt. A partir du profil du visiteur et des informations acquises dynamiquement à partir de son comportement, des hypothèses concernant les préfé-

rences du visiteur peuvent être inférées. En pratique, ces hypothèses permettent d'établir un modèle de visiteur, appelé signature d'intérêt, (§4.6.4) à partir de laquelle la personnalisation ou l'adaptation du parcours est effectuée au fur et à mesure que le visiteur découvre de nouvelles œuvres (section §4.7).

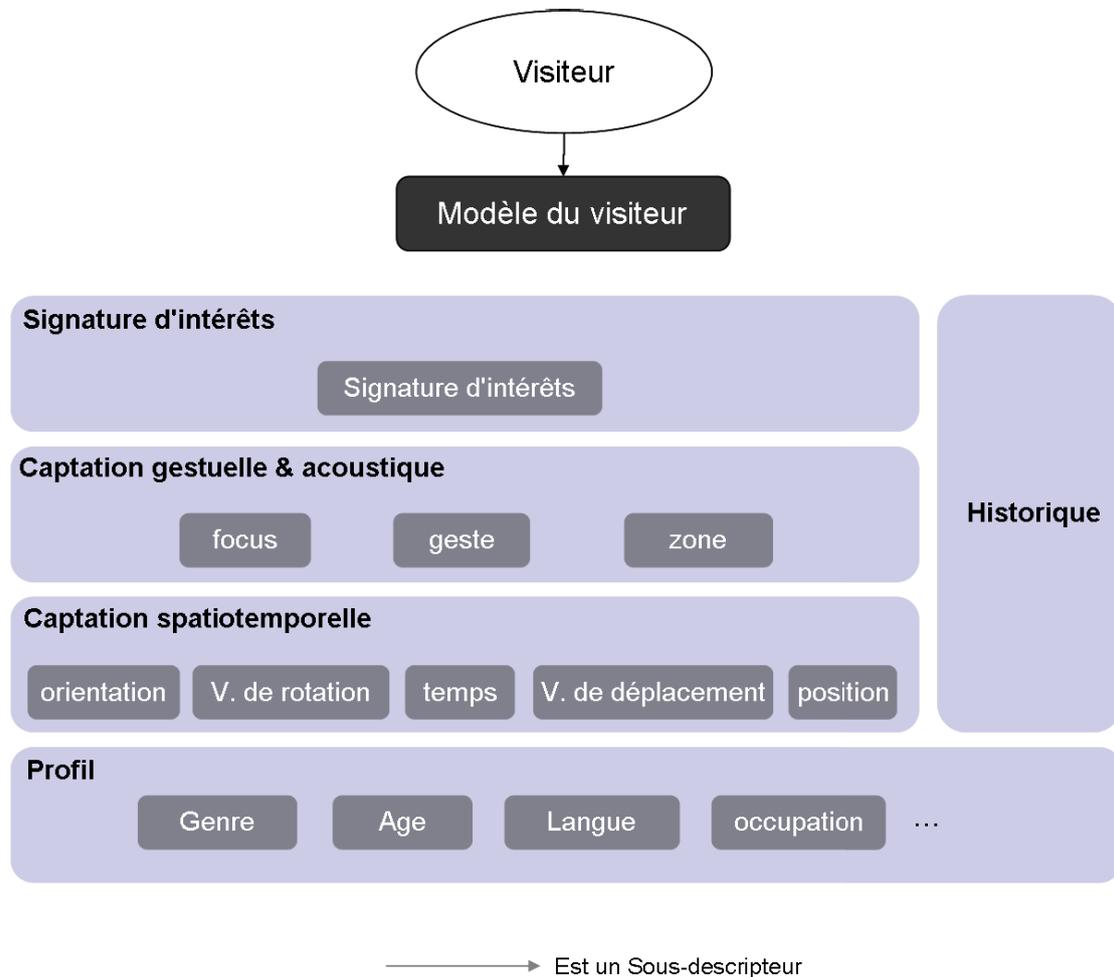


FIGURE 4.8 – Le modèle du visiteur.

En outre la couche gestuelle est ajoutée, ce qui permet de capturer et interpréter les gestes des visiteurs pour une meilleure personnalisation du contenu envoyé à chaque visiteur. Par ailleurs, si l'historique de visite dans *Listen* a comme objectif d'empêcher la lecture répétitive du même contenu audio. Dans notre modèle, celui-ci a deux objectifs : recalculer la signature des intérêts du visiteur (voir section 4.6.4) qui permet d'adapter son parcours pendant la visite, puis de garder des traces pour les visites postérieures. Dans une deuxième visite, le visiteur aura le choix entre prendre en considération ses

préférence durant la précédente visite, soit de partir sur une nouvelle expérience de visite qui va recalculer à nouveau une nouvelle signature d'intérêts. En outre, dans *Listen* l'historique est un simple log (fichier) qui trace l'identificateur de l'objet, l'identificateur du son et le temps. Dans notre modèle du visiteur, l'historique constitue une base de donnée qui sauvegarde les comportements du visiteur durant sa visite.

4.6.1 Profil

En amont de la visite, le visiteur est invité à communiquer des informations le concernant, telles que son âge, sa langue, ses centres d'intérêts, les domaines qu'ils l'intéressent parmi ceux du musée (comme exprimé au sein du thésaurus), et s'il s'agit ou pas de sa première visite du musée. Ces données sont utilisées afin de sélectionner le parcours le plus adapté comme par exemple le parcours ludique destiné aux enfants, parcours pour adulte, amateur ou expert. Ces informations permettent également de configurer l'audio-guide en fonction de la langue par exemple, et du fait qu'il s'agisse ou non d'une première visite. Notons que le fait qu'un utilisateur ne soit pas porté à effectuer des actions pour lesquelles il ne perçoit pas d'intérêt ou de bénéfice immédiat est un phénomène connu depuis plusieurs décennies dans le domaine des interfaces homme-machine [Carroll87]. Ainsi, la quantité d'informations demandée à l'utilisateur doit-elle être réduite au minimum ou obtenue par des méthodes indirectes. Par ailleurs, les intérêts d'un visiteur sont amenés à varier durant sa visite du fait de sa navigation dans le musée et de ses découvertes. Or, demander un feedback au visiteur durant sa visite, ou encore, requérir qu'il attribue des scores, négatifs ou positifs, aux œuvres l'intéressant ou non, ou encore qu'il mette à jour ses centres d'intérêts ne constitue pas une solution viable. Il est ainsi plus approprié d'observer les interactions, et sa navigation dans le musée, sans pour autant être (ou paraître) intrusif ou encore entraver sa liberté de mouvement, de façon à en inférer dynamiquement les intérêts dudit visiteur.

4.6.2 Captation spatiotemporelle

Traquer le visiteur consiste à définir sa position et son orientation de façon à en déduire son vecteur visuel, sa vitesse de déplacement et de rotation. Plus simplement, il s'agit de répondre aux questions suivantes : où est le visiteur ? dans quelle direction regarde-t-il ? à quelle vitesse il se déplace ? et à quelle vitesse il tourne sa tête ? La position du visiteur à l'instant t est exprimée par les coordonnées cartésiennes $(x(t), y(t), z(t))$. L'orientation quant à elle est définie par les trois angles d'Euler : le lacet, le tangage et le roulis

A partir de la position et de l'orientation, sont établies la vitesse de déplacement $V(t)$ et la vitesse angulaire $\omega(t)$.

4.6.3 Captation gestuelle et acoustique

Les différentes mesures calculées au niveau de la captation spatiotemporelle sont analysées pour déduire des informations concernant les gestes que le visiteur réalise avec sa tête. Ces gestes peuvent être utilisés comme moyen d'interaction avec la scène sonore pour contrôler le flux audio ou pour répondre aux questions posées à travers l'audioguide (voir section 4.7.2).

L'administrateur du système a ensuite la possibilité de définir autant de gestes qu'il le souhaite en utilisant les descripteurs élémentaires qui sont le temps, et les trois angles d'Euler. Dans notre cas, la définition se fait à travers des diagrammes UML exploitables par le système. Un exemple de définition de gestes est décrit par la figure 4.9.

Les gestes sont définis à travers un langage à base de règles qui permet de décrire un geste en se basant sur des indicateurs du niveau inférieur tels que l'orientation, la vitesse, le temps, etc. Par exemple, les gestes suivants peuvent être définis :

Réponse négative (*dire_non*) (Figure 4.9) : mouvement⁴ rapide allant de gauche à droite révélé par un changement prononcé du lacet.

Réponse affirmative (*dire_oui*) de haut en bas révélé par un changement de l'élévation.

Expectative (ou peut être) (*balancerTte*) correspondant à une variation du roulis.

En plus des informations de captation gestuelle, les informations sur le positionnement du visiteur au sein de l'environnement acoustique virtuel sont collectées. Deux informations importantes sont requises, la (ou les) zone(s) d'audibilité dans laquelle (lesquelles) se trouve le visiteur, et l'objet le plus proche du vecteur visuel du visiteur (le focus). Ceci est fait en se basant sur le modèle acoustique de la scène sonore défini précédemment. Une fois collectées, ces informations de captation sont sauvegardées dans un historique en vue d'une exploitation *on-line* c'est-à-dire effectuée au cours de la visite ou de la conception ou encore d'une exploitation *off-line* c'est-à-dire ultérieure et permettant d'établir des statistiques. Dans les deux cas, cet historique est utile à l'établissement des centres d'intérêt de chaque visiteur.

4.6.4 Signature d'intérêt

La signature d'intérêt est une notion introduite pour repérer les préférences du visiteur par rapport aux différentes catégories et classes d'objets et de contenus sonores présentés à travers la couche sémantique du modèle de la scène sonore. En effet, après la constitution d'une première idée sur les intérêts du visiteur lors de la création de son profil, un parcours prédéfini est proposé au visiteur. Cependant, ce parcours peut contenir des classes d'objets ou de contenus sonores qui n'intéressaient guère le visiteur, et sur lesquels le visiteur n'aimerait pas s'attarder. De la même manière, le par-

4. Une adaptation par rapport à la langue de visite est envisageable afin d'adapter la signification du mouvement à la culture du visiteur.

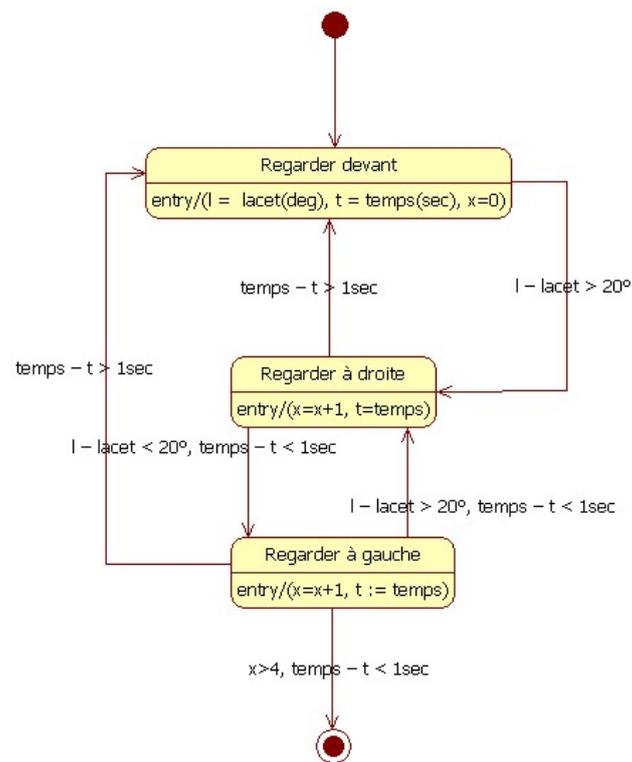


FIGURE 4.9 – Diagramme d'états-transitions du geste "réponse négative"

cours pourrait contenir un petit nombre d'objets appartenant à une classe précise qui intéresserait le visiteur et pour laquelle il aimerait découvrir plus d'objets.

L'objectif est ainsi d'établir dynamiquement, et pour chaque utilisateur, une liste de préférences détaillant pour chaque classe d'objets ou de contenus sonores un score d'intérêt reflétant l'importance qu'elle représente pour le visiteur.

C'est ce que nous avons appelé par la suite une signature d'intérêt. Cette liste est inférée à partir des informations de captation : zone d'audibilité activée, le focus, la durée et aussi les gestes collectés durant la visite (voir la section 4.6.5).

Signature d'intérêt = {(tag,valeur) où tag ∈ {concepts du thésaurus} et val ∈ N}

La figure 4.10 montre dans un diagramme de classe UML la liaison entre le visiteur et les tags du thésaurus de la couche sémantique de la scène sonore.

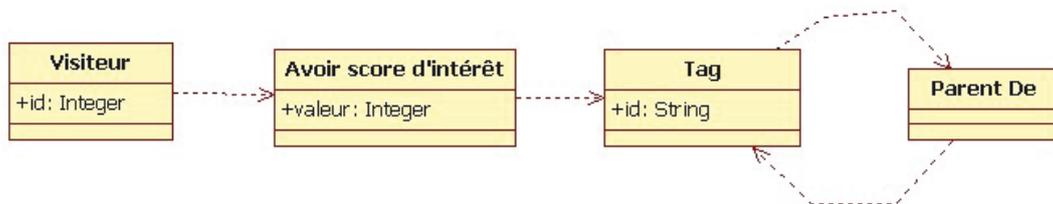


FIGURE 4.10 – Liaison entre le visiteur et les tags du thésaurus.

Nous nous inspirons de l'approche de Pazzani [Pazzani97] qui a observé que l'intérêt que porte un visiteur à un objet est déterminé par le temps passé à le regarder. Le fait de fixer un seuil de temps permet de discrétiser les œuvres particulièrement intéressantes. En pratique, si le visiteur se trouve dans une zone pendant une durée supérieure à un seuil donné, ou encore si le visiteur fixe la zone (pour une durée supérieure au seuil), alors le visiteur est supposé apprécier le contenu de cette zone. Au niveau de la dépréciation, plusieurs choix subsistent et sont sujets à discussion et évaluation. Le fait de ne pas passer de temps devant une œuvre ou quitter l'œuvre avant que le contenu audio associé soit terminé ou d'éviter une salle peut être considéré comme une dépréciation, ainsi que le geste d'interrompre le contenu audio. En tant que telles, ces informations ne peuvent être exploitées afin de déterminer les centres d'intérêts de l'utilisateur. Elles doivent être confrontées à des domaines, et des types de contenus audio.

Dans cette optique, nous nous basons sur les concepts définis à priori, organisés dans le thésaurus sous une forme hiérarchique en annotant les zones d'audibilité. L'annotation est effectuée non pas uniquement au niveau des sujets des zones mais aussi au niveau des contenus sonores associés. En effet, le visiteur peut manifester son intérêt ou désintérêt pour une zone sans que le système sache si le visiteur juge le sujet de la

zone ou le son l'accompagnant. C'est pour cette raison qu'un traitement supplémentaire est effectué pour distinguer l'intérêt/désintérêt du visiteur par rapport au sujet de celui par rapport au son (figure 4.11). Ainsi si l'algorithme déduit que l'utilisateur apprécie un sujet (resp. un son) particulier associé à une ZA, les scores d'intérêt des tags décrivant ce sujet (resp. ce son) seront tous relevés de 1. Au contraire, si l'algorithme déduit que l'utilisateur montre un désintérêt, les scores des tags correspondants seront diminués de 1.

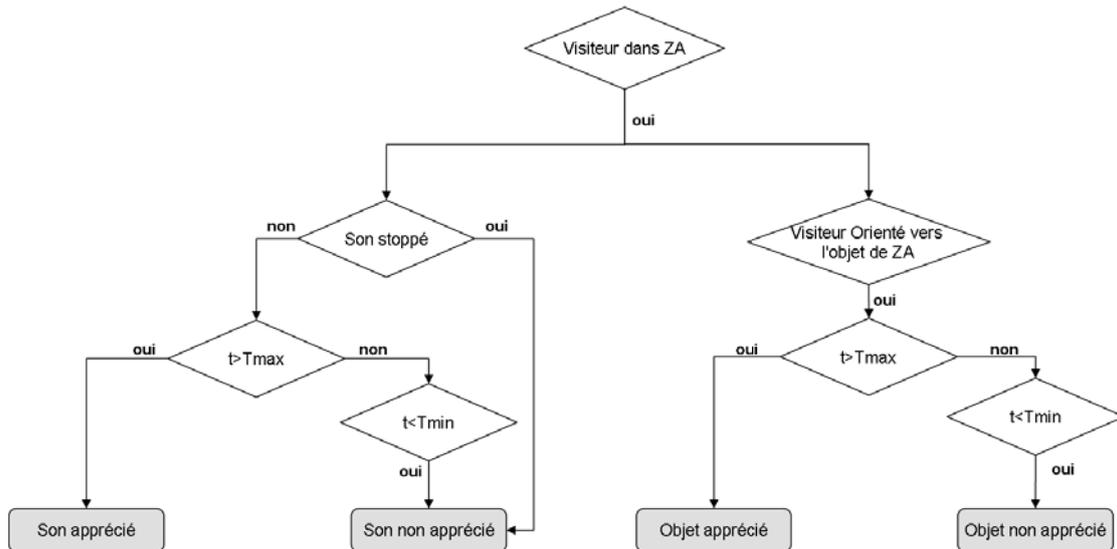


FIGURE 4.11 – Analyser les préférences.

4.6.5 Historique

L'historique du visiteur est analysé et stocké durant et après la visite. Il dépend des paramètres zone, focus et position. L'historique s'alimente à partir des différents descripteurs du visiteur. Comme le profil, qui à son tour, s'alimente dynamiquement de l'historique (relation de va et vient).

L'historique de la visite qui n'est rien d'autre qu'une base de données, qui est renseignée au fur et à mesure de la visite.

Un démon, qui est exécuté en arrière plan, analyse en permanence les positions et orientations du visiteur. Il se base sur les la couche acoustique de la scène sonore ainsi que sur l'ensemble des définitions de gestes pour calculer directement les descripteurs de la deuxième couche du modèle du visiteur, qui sont la zone ayant le focus, la zone dans laquelle se trouve le visiteur, et le geste effectué par le visiteur. La figure 4.12 représente un diagramme de classe modélisant les classes qui participent dans la création de l'historique. A chaque opération de mise à jour des tables *Focus_Vers_Zone*,

ET_Dans_Zone, et *Effectuer_Geste* des préférences du visiteur, le même démon réévalue les valeurs de la table *Avoir_score_dInteret* suivant l'algorithme de l'analyse des préférence décrit par la figure 4.11.

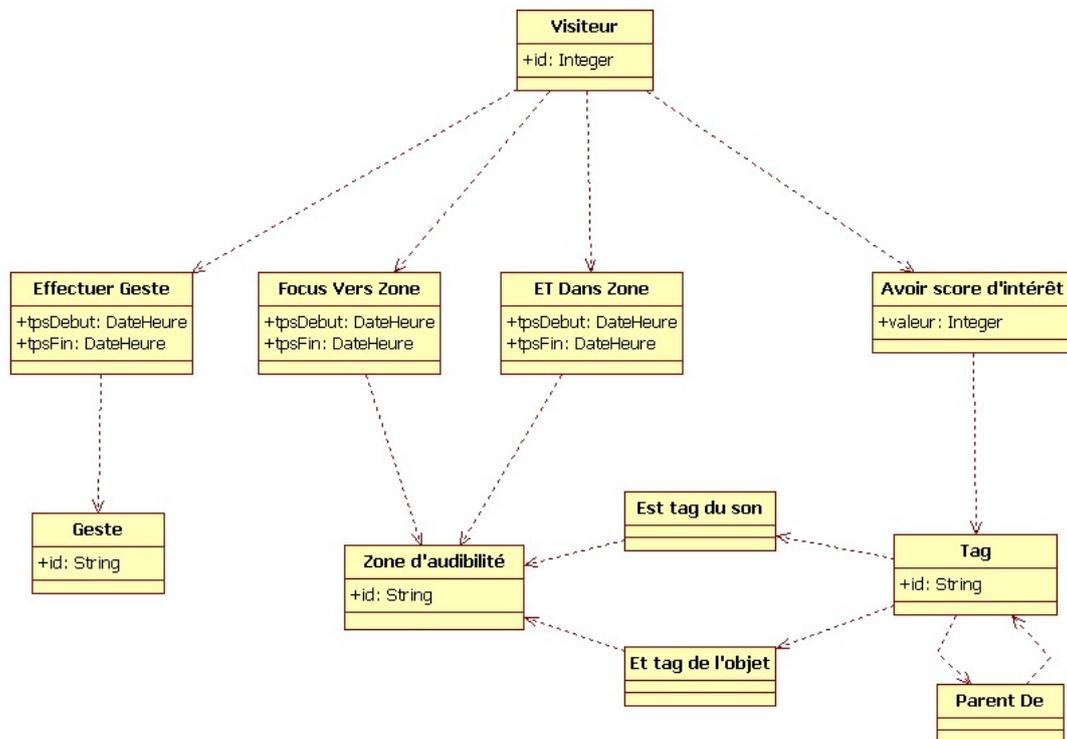


FIGURE 4.12 – Diagramme de classe de l'historique.

Ainsi le modèle du visiteur (Figure 4.8) organisé par couches dont les niveaux de descripteurs vont du plus basique (captation, profil) aux niveaux les plus élevés (signature d'intérêt) permet d'adapter et de personnaliser à la fois le paysage sonore mais aussi plus généralement le parcours et la visite.

4.7 Navigation

Le modèle de navigation décrit les traitements réalisés en réponse aux comportements du visiteur lors de sa visite du musée. Il vise à interpréter les actions du visiteur pour interagir avec la scène sonore.

La navigation du visiteur dans les 2 projets *ec(h)o* et *Listen* est structurée en deux niveaux :

Niveau 0 : appelé aussi macro chez *ec(h)o* et contrôles continus chez *Listen*, c'est le niveau où le visiteur interagit avec le paysage sonore uniquement par son mouvement, notamment sa position et/ou son orientation qui sont traquées automatiquement.

Niveau 1 : appelé micro chez *ec(h)o* et événement chez *Listen*, il s'agit du niveau le visiteur interagit avec le paysage sonore par la sélection directe(*ec(h)o*), ou passer du temps devant le même objet(*Listen*).

Dans notre modèle de navigation, le flux audio diffusé au visiteur, le paysage sonore dans lequel il évolue, et le parcours de visite qu'il emprunte sont modifiés suivant son déplacement, sa gestuelle et ses préférences.

Trois modes d'interaction sont alors définis dans ce modèle :

Traque : c'est la mise à jour continue du contenu audio spatialisé diffusé au visiteur qui se base sur la position et l'orientation de ce dernier au sein de la scène sonore.

Contrôle : il s'agit du fait de donner à l'utilisateur la possibilité de contrôler, la diffusion sonore indépendamment du fonctionnement prévu par la spatialisation, en lui permettant de contrôler le volume, d'activer ou désactiver des objets de son parcours, ou de stopper la diffusion d'un contenu audio, à travers des gestes.

Adaptation : il s'agit de l'exploitation de la signature des intérêts pour adapter le parcours du visiteur sur la base de ses préférences inférés au cours de son parcours.

Les trois modes sont organisés de manière concentrique de façon à ce que chaque mode satisfait également la fonction de celui qu'il contient.

4.7.1 Traque

La traque est le premier noyau de navigation où le visiteur interagit avec le paysage sonore uniquement à travers son déplacement dans les salles du musée. En effet, la position du visiteur et son orientation, lui permettent de recevoir le contenu spatialisé adapté à son déplacement. L'intérêt de ce mode est d'immerger le visiteur dans le paysage sonore du musée et créer un environnement sonore qui interagit avec le mouvement du visiteur.

Le protocole du mode Traque est donné par la figure 4.13 et le diagramme d'activités (figure 4.14). Ce protocole se base sur la notion de zone d'audibilité qui délimite l'espace dans lequel le son est entendu par le visiteur. Si un visiteur se trouve à l'intérieur de l'espace géographique délimité par une zone d'audibilité, alors cette dernière est activée. Rappelons que les zones d'audibilité sont appelées à se juxtaposer les unes avec les autres, ce qui requiert de composer les sons en résultant. Ainsi, un visiteur peut se trouver dans une zone géographique à l'intersection de deux zones d'audibilité, A_j et A_k ($j = k$), qui se juxtaposent et partagent toutes deux la même zone mère A_i , c'est-à-

dire que *visiteur* A_i A_j A_k . Dans ce cas (listing 4.9), nous devons différencier deux cas :

- A_j et A_k sont de type partagées (lignes 1-3 du listing 4.9).
En d'autres termes, $A_j.perception = shared$ et $A_k.perception = shared$. Dans ce cas, les sons (mixés) correspondant aux deux zones sont entendus par le visiteur ($A_j \text{ } A_k.sountoplay$),
- A_j et A_k sont toutes deux exclusives (lignes 5-7 du listing 4.9).
Ainsi, $A_j.perception = standalone$ et $A_k.perception = standalone$. Un seul son est entendu par le visiteur ($A_j \text{ } A_k.sountoplay$), soit celui correspondant à la zone A_j soit celui de A_k : $A_j \text{ } A_k.sountoplay$. Dans les faits, il s'agit de celui regardé par le visiteur.
- Dans le cas où l'une des zones serait exclusive (lignes 10-12 du listing 4.9) - notons⁵ cette zone A_j - et l'autre zone A_k serait partagée, alors comme dans le cas précédent seule la zone regardée par le visiteur est entendue de façon à se prémunir d'un son non audible provenant par exemple du mixage de deux dialogues.

Listing 4.9 – Contrôle des zones d'audibilités

1	IF (visiteur A_j A_k)	$A_j.perception = shared$
2		$A_k.perception = shared$
3	DO	$A_j \text{ } A_k.sountoplay$
4		
5	IF (visiteur A_j A_k)	$A_j.perception = standalone$
6		$A_k.perception = standalone$
7	DO	$A_j \text{ } A_k.sountoplay$
8		
9	IF (visiteur A_j A_k)	$A_j.perception = standalone$
10		$A_k.perception = shared$
11	DO	$A_j.sountoplay$

Les relations existantes entre une zone parent A_i et une zone fille A_j sont assez proches.

- A_i et A_j sont partagées (lignes 1-3 du listing 4.10). Dans ce cas, le son de la zone fille est en premier plan (in Background $A_i.sountoplay$) alors que celui de la zone mère est joué en arrière plan (in foreground $A_i.sountoplay$),
- A_i et A_j sont toutes deux exclusives (lignes 5-7 du listing 4.10), le son dont la zone est regardée est entendu.
- Si une zone est exclusive (lignes 9-11 du listing 4.10) et l'autre est partagée, seule la zone regardée par le visiteur est entendue.

Listing 4.10 – Contrôle des zones d'audibilités

1	IF (visiteur A_i A_j)	$A_i.perception = shared$
2		$A_j.perception = shared$

5. sans manque de généralité

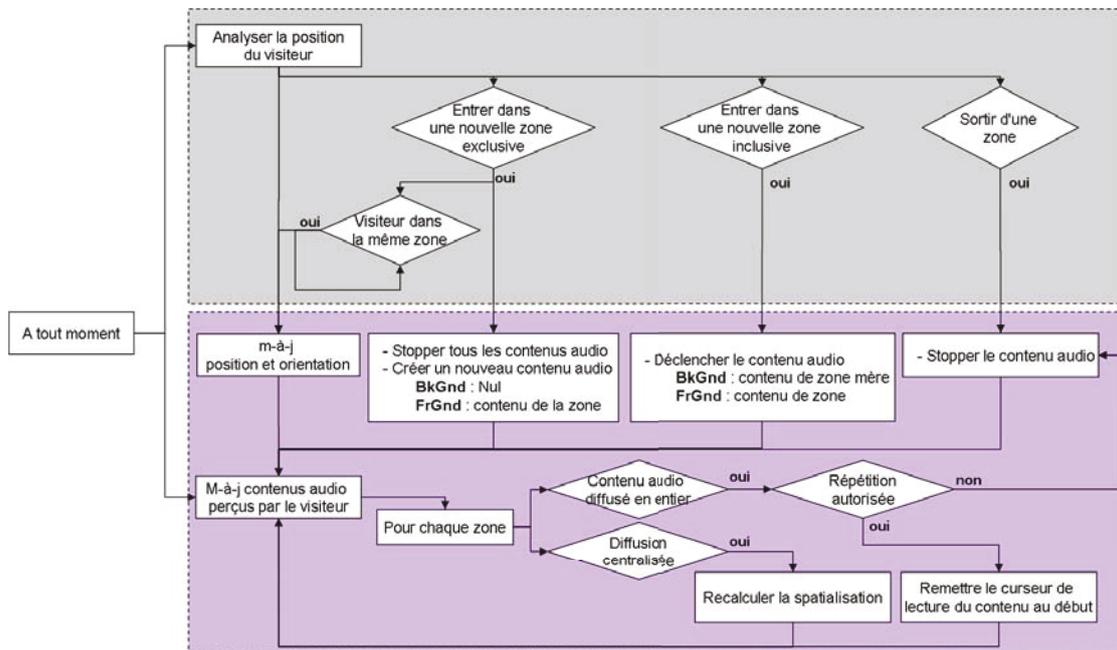


FIGURE 4.13 – Protocole de la navigation en mode Traque dans le paysage sonore

```

3 | DO in Background  $A_i$ .sountoplay in foreground  $A_i$ .sountoplay
4 |
5 | IF (visiteur  $A_i$   $A_j$ )  $A_i$ .perception = standalone
6 |                        $A_j$ .perception = standalone
7 |   DO  $A_i$   $A_j$ .sountoplay
8 |
9 | IF (visiteur  $A_i$   $A_j$ )  $A_i$ .perception = standalone
10 |                       $A_j$ .perception = shared
11 |   DO  $A_i$   $A_j$ .sountoplay

```

4.7.2 Contrôle

Le mode *Contrôle* est un mode qui englobe celui de la *Traque*, mais qui permet un contrôle plus avancé du paysage sonore à travers des interactions volontaires du visiteur. En effet, en plus de réaliser le suivi du déplacement du visiteur et la spatialisation du contenu audio qui lui est diffusé, ce mode autorise le visiteur à contrôler le paysage sonore par des gestes simples de sa tête. Ce mode d'interaction qui est basé sur les gestes de la tête, a été adopté pour sa simplicité et son aspect intuitif. Ainsi, il permet à l'utilisateur de garder ses mains libres sans l'encombrer avec des dispositifs technologiques qui éloigneraient son attention des œuvres exposées et des sons diffusés. Aussi, disposer d'un système qui capterait les trois angles d'Euler correspondant aux diffé-

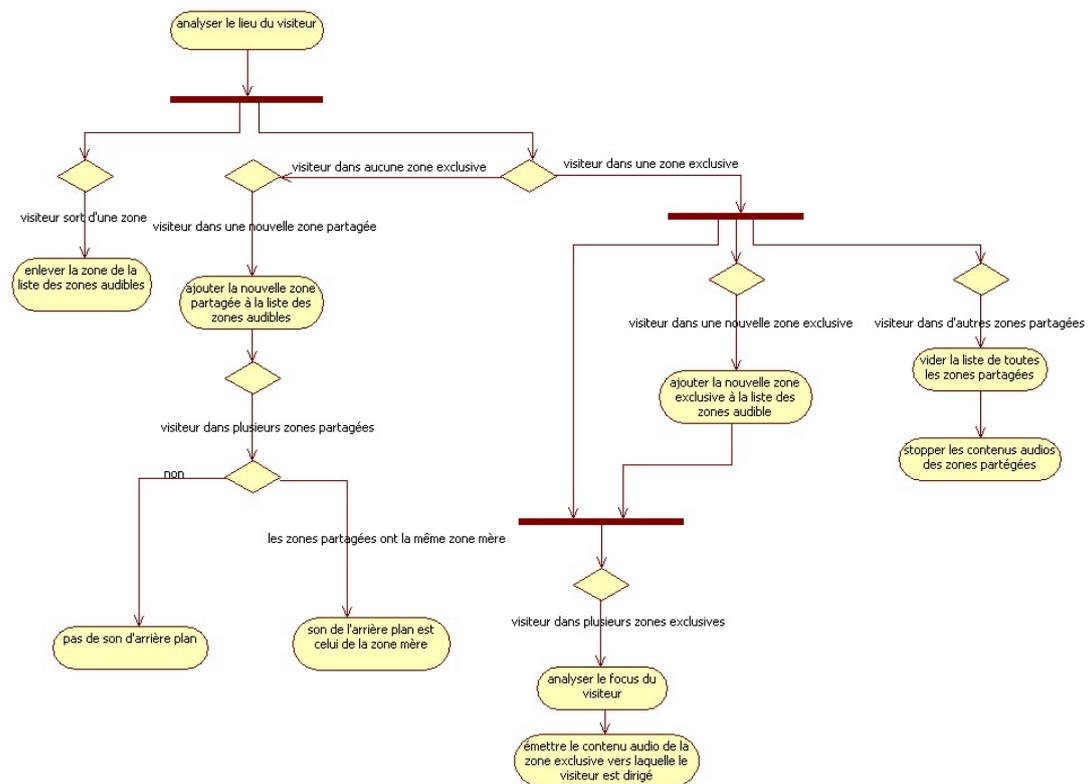


FIGURE 4.14 – Protocole de la navigation dans le paysage sonore en mode traque

rents mouvements de la tête du visiteur offrirait une large expressivité qu'on pourrait exploiter pour définir un pseudo langage d'interaction entre le visiteur et le paysage sonore.

Le mode *Contrôle* est modélisé au travers d'un langage déclaratif (en anglais *declarative rule based language for expressing constraints*) qui se prête à l'expression de règles et de contraintes. Une règle est exprimée sous le format suivant (listing 4.11) : dans un contexte donné (**context**), une condition particulière (**condition**) mène à une action (**action**) possiblement répétitive (**WHILE (parameters)**).

Listing 4.11 – Règle de contrôle

```
1 IN ( context ), IF ( condition ) WHILE ( parameters )
2 DO ( action )
```

Nous souhaitons que la visite reste le plus possible naturelle et que l'interaction se rapproche au maximum de l'implicite que de l'explicite. Nous voulons éviter d'exiger des visiteurs la réalisation des gestes complexes qui pourrions les incommoder pendant la visite.

En effet, l'objectif est de composer un langage assez expressif qui permet d'exprimer un maximum de commandes. Ceci devra être fait en utilisant des gestes simples et élémentaires que les hommes emploient en général pour exprimer rapidement des prises de positions. Ces gestes, comme expliqué dans la section 4.6.3, peuvent être : *dire_oui* en levant la tête et baissant la tête, *dire_non* en tournant la tête de gauche à droite, *balancer_Tte* en la faisant pencher la tête de droite à gauche. Des arrêts prolongés devant les objets peuvent être également interprétés comme un geste et pourra être utilisé dans l'interaction avec l'environnement sonore.

Le concept de *contexte* est ainsi introduit comme étant un contexte d'interactivité avec l'environnement sonore. Dans chaque *contexte* l'interprétation des gestes réalisés par le visiteur peut être différente. Le visiteur pourra naviguer par un geste simple et unique *balancer la tête*. Il peut ensuite sélectionner un *contexte* par un geste de *dire oui* après une éventuelle demande de confirmation audio. Il accède ainsi à une interaction par gestes pour contrôler l'environnement sonore.

La figure 4.15 montre un exemple d'une carte de navigation par les gestes que le concepteur peut définir à travers les règles de contrôle.

Nous définissons 4 *contextes* :

- Le *contexte Normal_visit* qui renvoie vers le mode par défaut d'une visite. Dans ce *contexte*, trois gestes du visiteur vont être interprétés comme interactions avec la scène sonore. Le geste *dire_Oui* permet d'activer la ZA associée à un objet devant lequel est arrêté le visiteur et qui correspond le plus au profil du visiteur. Le geste *dire_Non* permet de stopper la diffusion du son de la ZA courante. Le geste *Arrêt_Prolongé* permet d'activer la ZA courante et de l'inclure dans le parcours du visiteur si elle est initialement inactive.

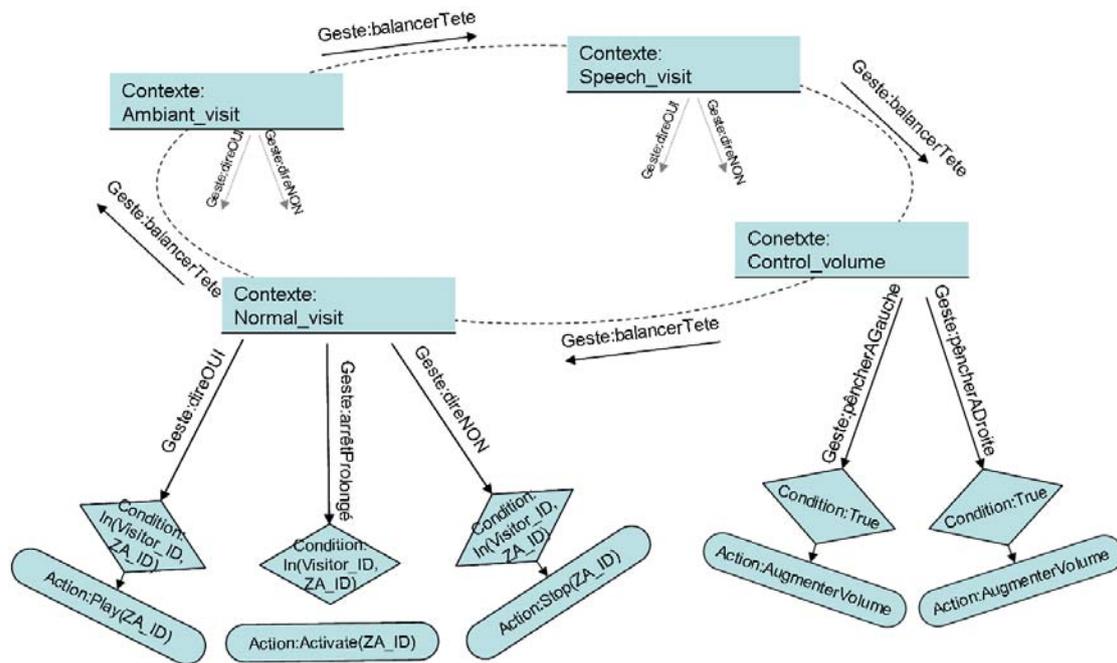


FIGURE 4.15 – Exemple d’une carte de navigation par les gestes

- Le *contexte Control_volume* qui renvoie vers le mode de contrôle de volume de diffusion. Deux gestes permettent de contrôler le volume : le geste *pencher_A_Droite* qui permet d’augmenter le volume, et le geste *pencher_A_Gauche* qui permet de le diminuer.
- Le *contexte Ambient_visit* qui renvoie vers un mode dans lequel toutes les ZA dont le son à un autre type que l’ambient sont désactivées. Les gestes interprétables restent les mêmes que ceux du *contexte Normal_visit*, et les actions en découlant sont les mêmes également.
- Le *contexte Speech_visit* qui renvoie vers un mode dans lequel toutes les ZA dont le son à un autre type que l’ambient sont désactivées. Les gestes interprétables restent les mêmes que ceux du *contexte Normal_visit*, et les actions en découlant sont les mêmes également.

Le concepteur a ensuite le choix de définir d’autres modes d’interaction.

Les actions pouvant être utilisées sont principalement celles qui permettent l’arrêt de la diffusion d’un son associé à une ZA (action 1), le déclenchement d’un son associé à une ZA (action 2), l’ajout d’une ZA dans le parcours d’un visiteur (action 3) ou sa suppression du parcours du visiteur (action 4), augmenter le volume de la lecture des contenus audio (action 5) ou le baisser (action 6), le changement du contexte d’interactivité (action 7), et finalement la création d’un nouveau contexte (action 8).

Listing 4.12 – Actions entreprises sur le paysage sonore

```

1 1. setStoped(Visitor.tour, ZA)
2 2. setPlaying(Visitor.tour, ZA)
3 3. add(Visitor.Tour, ZA)
4 4. remove(Visitor.Tour, ZA)
5 5. volume_up(Visitor.tour)
6 6. volume_down(Visitor.tour)
7 7. setMode(Control_volume)
8 8. createMode(new_Mode)

```

Les exemples suivant décrivent la définition de règles de différents types. Le listing 4.13 sont les règles qui permettent au visiteur de contrôler le volume du son diffusé. Le listing 4.14 décrit les règles qui permettent l'arrêt et le déclenchement de sons associés au ZA visités par le visiteur suite à ses gestes, et éventuellement le rajout des ZA dans son parcours de visite. Finalement, le listing 4.15 décrit les règles qui permettent de naviguer entre les différents contextes d'interactivité tel que définit dans la figure 4.15.

Listing 4.13 – Règles de contrôle pour régler le volume

```

1 IN Control\_volume, IF (TiltRight(Visitor, time))
2   WHILE TRUE DO (volume\_up(Visitor.tour))
3 IN Control\_volume, IF (TiltRight(Visitor, time))
4   WHILE TRUE DO (volume\_down(Visitor.tour))

```

Listing 4.14 – Règles de contrôle pour contrôler la diffusion des sons liés aux ZA

```

1 IN Normal\_visit, IF (Say\_No(Visitor, timeStart, timeEnd))
2   WHILE (IN(Visitor, ZA, timeStart, timeEnd))
3     DO (setStoped(Visitor.tour, ZA))
4 IN Normal\_visit, IF (Say\_Yes(Visitor, timeStart, timeEnd)
5   OR Long\_Stand(Visitor, timeStart, timeEnd))
6   WHILE (IN(Visitor, ZA, timeStart, timeEnd))
7     DO (
8       IF(IN(Visitor.Tour, ZA)) DO(
9         IF(isStoped(Visitor.tour, ZA))
10          DO setPlaying(Visitor.tour, ZA)
11       )
12       ELSE DO (Add(Visitor.Tour, ZA))
13     )

```

Listing 4.15 – Règles de contrôle pour naviguer entre les contextes d'interactivité

```

1 IN Normal\_visit, IF (Say\_maybe(Visitor, time))
2   WHILE TRUE DO (setMode(Ambiant\_visit))
3 IN Ambient\_visit, IF (Say\_maybe(Visitor, time))
4   WHILE TRUE DO (setMode(Speech\_visit))
5 IN Speech\_visit, IF (Say\_maybe(Visitor, time))
6   WHILE TRUE DO (setMode(Control\_volume))
7 IN Control\_volume, IF (Say\_maybe(Visitor, time))
8   WHILE TRUE DO (setMode(Normal\_visit))

```

4.7.3 Adaptation

L'adaptation est le troisième mode d'interaction. Superposant les deux autres modes, l'adaptation permet en plus de réaliser la *Traque* et de permettre le *Contrôle*, d'adapter le parcours de la visite en analysant le comportement antécédent du visiteur et en y déduisant ses préférences. Ce niveau exploite les descripteurs du visiteur fournis par la signature des intérêts.

L'algorithme d'adaptation est responsable de décider si les ZAs incluses dans le parcours du visiteur sont toujours intéressantes pour lui et ainsi devront toujours être présentes dans son parcours. Il est aussi responsable de décider si des ZAs qui ne sont pas incluses dans le parcours du visiteur doivent l'être en vue de ses intérêts exprimés dans la signature des intérêts.

L'adaptation consiste ainsi à s'inspirer des scores d'intérêt des tags de chaque objet ou son associée à un ZA de la scène sonore, pour inférer si elle est dépréciée ou appréciée et ainsi la supprimer ou la rajouter au parcours du visiteur.

Soit une zone d'audibilité ZA qui correspondent un objet d'art O est un son S . Soit $Tags(O) = to_1, \dots, to_n$ les tags associés à l'objet O , et soit $Tags(S) = ts_1, \dots, ts_m$ les tags associés au son S . Soit $N_{to} = |Tags(O)|$ le nombre de tags d'objet, et $N_{ts} = |Tags(S)|$ le nombre de tags de son associés à cette ZA .

Nous définissons :

- $score_interest(V, ZA)$ est le score d'intérêt du visiteur V à la zone ZA ,
- $score_interest(V, O)$ est son score d'intérêt du visiteur V à l'objet O ,
- $score_interest(V, S)$ est son score d'intérêt du visiteur V au son S
- $score_interest(V, t)$ est son score d'intérêt du visiteur V au tag t .

Nous considérons que l'objet et le sujet participe chacun à 50% dans la création de l'intérêt du visiteur à une ZA . Ainsi :

$$Score_interest(V, ZA) = \begin{cases} \frac{score_interest(V, O) + score_interest(V, S)}{2} & \text{si } N_{to} = 0 \text{ et } N_{ts} = 0 \\ score_interest(V, O) & \text{si } N_{to} = 0 \text{ et } N_{ts} = 0 \\ score_interest(V, S) & \text{si } N_{to} = 0 \text{ et } N_{ts} = 0 \\ 0 & \text{si } N_{to} = 0 \text{ et } N_{ts} = 0 \end{cases}$$

A tout moment pendant le parcours du visiteur, le score d'intérêt à un tag objet s'exprime de la manière suivante :

A tout moment pendant le parcours du visiteur, si un objet (resp. un son) d'une ZA est apprécié, les scores d'intérêt à ses tags d'objets (resp. de son) sont incrémentés. De même s'il est déprécié, les scores d'intérêt à ses tags sont dévalués.

En cas d'appréciation, on incrémente le score d'intérêt à un tag de la manière suivante :

$$score_interest_i(V, t) = score_interest_{i-1}(V, t) + \frac{1 - score_interest_0(V, t)}{\text{nombre de ZA taguées par } t} \quad (4.1)$$

En cas de dépréciation, le score est dévalué de la manière suivante :

$$score_interest_i(V, t) = score_interest_{i-1}(V, t) - \frac{score_interest_0(V, t)}{\text{nombre de ZA taguées par } t} \quad (4.2)$$

avec $score_interest_0(V, t)$ le score initial attribué au tag t suivant le profil initial du visiteur. Les valeurs minimales et maximales des scores d'intérêts sont fixées à 0 et 1.

L'appréciation de l'objet ou du son est déduite à partir du processus d'analyse des préférences décrit dans la figure 4.11.

Le score d'intérêt qu'un visiteur V porte à un objet O (resp. à un son (S)) d'une ZA est défini comme la moyenne des scores d'intérêts que ce visiteur porte à l'ensemble des tags d'objets (tags de son) de cette ZA.

$$Score_interest(V, O) = \frac{\sum_{i=1}^{N_{to}} Score_interest(V, to_i)}{N_{to}} \quad (4.3)$$

$$Score_interest(V, S) = \frac{\sum_{i=1}^{N_{ts}} Score_interest(V, ts_i)}{N_{ts}} \quad (4.4)$$

Ensuite, on définit deux seuils liés à l'intérêt à une ZA : un seuil de désintérêt $minInterestT$ et un seuil d'intérêt $maxInterestT$. L'objectif est de définir trois plages d'intérêt. $[0, minInterestT]$ où les ZA seront inintéressants pour le visiteur, $[minInterestT, maxInterestT]$ où les ZA seront neutres pour le visiteur, et $[maxInterestT, 1]$ où les ZA seront intéressants pour le visiteur.

Le score d'intérêt à une ZA sera d'après les équations précédentes entre 0 et 1 :

$$0 \leq score_interest(V, ZA) \leq 1.$$

Le choix des deux seuils devra respecter l'inégalité suivante : $0 \leq minInterestT \leq maxInterestT \leq 1$.

Pour initialiser le processus d'adaptation, il faudra définir un seuil d'appréciation T_a , ainsi que les deux seuils $minInterestT$ et $maxInterestT$.

4.8 Exemple développé

Nous prenons comme exemple la salle des communications du musée des arts et métiers (figures 4.20,4.21), qui abrite plusieurs outils de communication et de transmission de l'information. Le nombre impressionnant d'œuvres exposées dans cette salle et

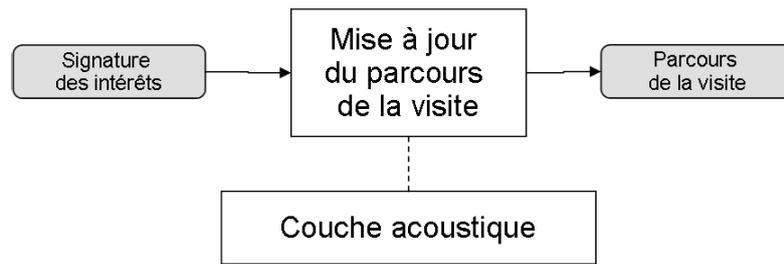


FIGURE 4.16 – Processus d’adaptation du contenu audio

appartenant à différents domaines peut facilement désorienter le visiteur. Il serait ainsi judicieux de porter l’attention du visiteur sur un nombre restreint d’œuvres représentatives, puis d’enrichir son parcours avec plus d’œuvres qui rentrent dans son centre d’intérêt.

La salle de communication sur laquelle va porter notre cas d’usage comporte des œuvres (des machines) classées par 3 catégories : l’imagerie, l’imprimerie et la télégraphie.

Le modèle de scène sonore développé pour cet exemple est constitué de 3 couches :

- La couche physique reproduit la répartition des objets exposés dans l’espace réel tel que c’est présenté dans la figure 4.17. Dans le tableau 4.22, les machines 10 qui rentrent dans l’exemple sont identifiées de A à J. Les machines correspondant à l’*imprimerie* sont de A à D. La machine A (Machine à écrire de Remington) est considérée comme représentative de cette catégorie. Les machines correspondant à la *télégraphie* sont E et F, F (Télégraphe aérien de Chappe) étant l’œuvre représentative. Les œuvres liés à l’*imagerie* sont identifiés de G à J. H (Chambre de touriste Enjalbert) et J (Fantasmagories de Robertson) sont considérées comme représentatives.
- La couche acoustique décrit les ZAs composant la scène sonore. Vingt ZAs sont créées dans cet exemple. A chaque objet du musée en associe une ZA de type partagé dans laquelle un son Ambient est joué, et une ZA de type exclusif dans laquelle un son de description est joué. Un exemple la ZA ambiante associée à la machine A (Machine à écrire de Remington) est présenté dans le listing 4.16.
- La couche sémantique est choisie d’une telle manière à simplifier les calculs. En effet pour simplifier l’adaptation aux préférences du visiteur, nous allons nous baser sur un thésaurus réduit (figure 4.18) comportant trois descripteurs : *imagerie*, *imprimerie* et *télégraphie*. Chacune des œuvres est décrite par un seul descripteur à la fois.

La navigation dans cet exemple se fera selon les trois niveaux ; la traque, le contrôle et l’adaptation. Pour le contrôle, les règles définies dans le listing 4.14 vont être utilisées. Pour l’adaptation nous utilisons le même algorithme défini précédemment pour l’adaptation avec $minInterestT = 0.4$ et $maxInterestT = 0.7$.

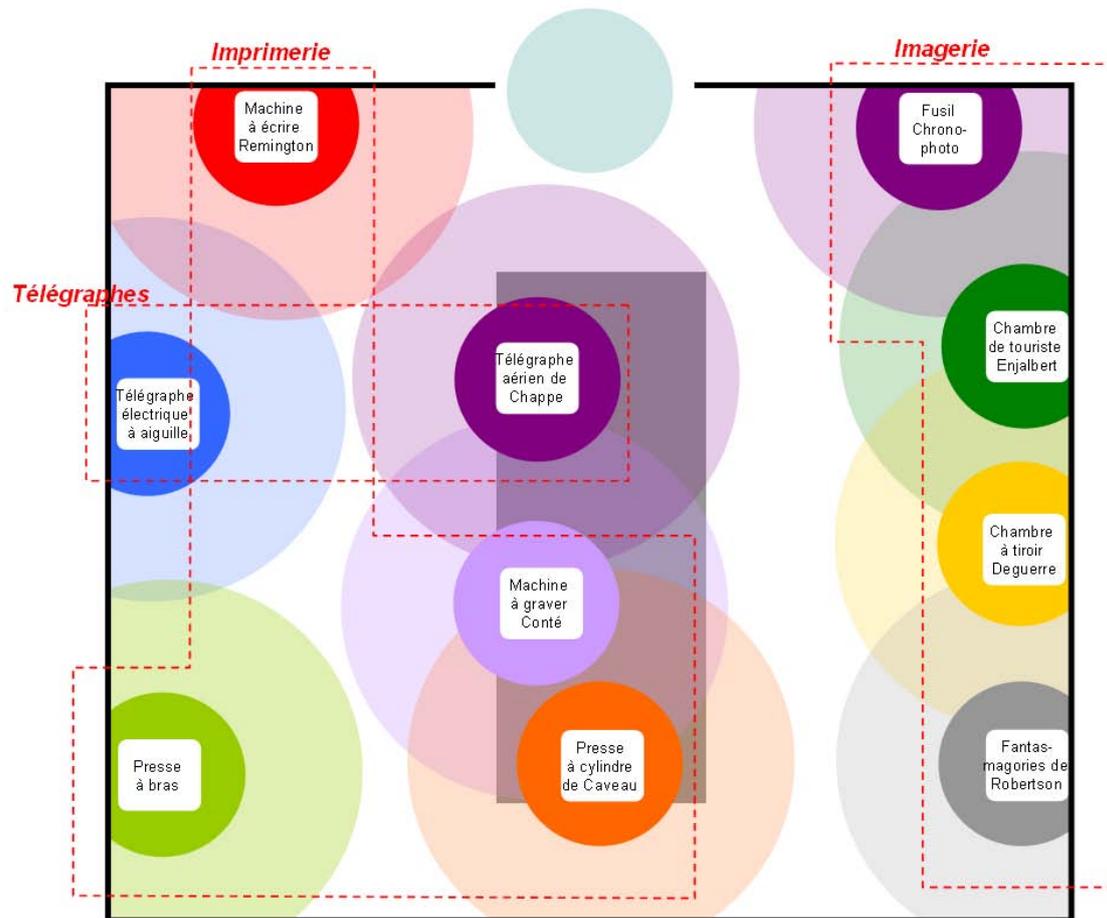


FIGURE 4.17 – Scène sonore de l'exemple.

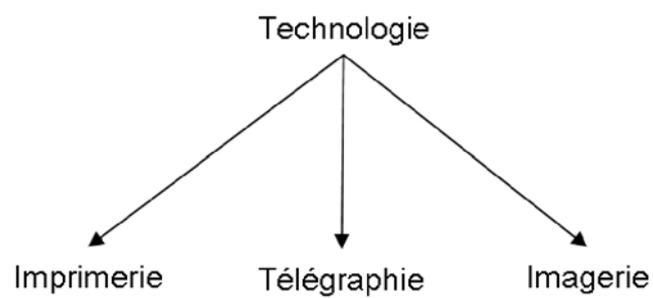


FIGURE 4.18 – Thesaurus de l'exemple.

A son arrivée au musée, le visiteur répond à un questionnaire évaluant ces centres d'intérêts. Les réponses à ce questionnaire serviront à choisir les œuvres qui constitueront son parcours, et qui seront ainsi sonorisées. Dans notre cas de figure, ce questionnaire lui demande son avis par rapport aux 3 catégories du thésaurus *imagerie*, *imprimerie* et *télégraphie*. Prenons le cas d'un visiteur dont le profil montre un intérêt pour l'imagerie, et un désintérêt vis à vis de la télégraphie. Ces réponses au questionnaire sont présentées dans la figure 4.19. Sur la base de ce profil, le système marque les machines décrites par *imagerie* comme étant intéressantes pour le visiteur et les inclut toutes à son parcours. De la même manière, les œuvres décrites par *télégraphie* sont marquées comme inintéressantes pour le visiteur et sont écartées de son parcours. Le visiteur s'exprime neutre concernant l'imprimerie, le système inclut uniquement son élément représentatif, la machine A, au parcours du visiteur.

$$\text{Visitor.tour}(t_0) = A, G, H, I, J$$

	Pas intéressé	Neutre	Intéressé
• Télégraphie	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Imagerie	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
• Imprimerie	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

FIGURE 4.19 – Questionnaire présenté au visiteur et ses réponses.

Le tableau 4.22 décrit les différents états du parcours du visiteur dans la salle, et le changement de l'état de sonorisation des machines en concordance avec le comportement de l'utilisateur.

A partir des réponses du visiteur, nous inférons les scores d'intérêts suivants pour les 3 catégories :

- $\text{Score_interest}(\text{Visiteur}, "Imagerie")=1$
- $\text{Score_interest}(\text{Visiteur}, "Imprimerie")=0,5$
- $\text{Score_interest}(\text{Visiteur}, "Télégraphie")=0$

Dès son entrée à la salle, le visiteur reçoit un contenu sonore lui présentant le thème de la salle (Etat 0). Puis, en se déplaçant à travers les couloirs de la salle, le visiteur accède aux différentes zones d'audibilité liées aux objets exposés et sonorisés. Le visiteur commence sa visite par le couloir de droite et passe devant la machine à écrire "Remington" (ZA A). Ce dernier choisit de s'arrêter devant cette machine le temps d'écouter le contenu audio (Etat 1). Ceci ne change rien en terme de score d'intérêt au tag Imprimerie $\text{Score_interest}(\text{Visiteur}, "Imprimerie")=0,5+(1-0,5)/4=0,625$ (Formule 1.1).

Puis il avance et s'arrête devant le télégraphe aérien de "Chappe" (ZA F). Bien que ce télégraphe n'est pas sonorisé, le visiteur décide d'activer le contenu audio lié à cet objet en utilisant le moyen d'interaction offert par le niveau d'interactivité "Contrôle". Le visiteur lève sa tête de haut en bas, en effectuant un geste *direOui*, ce qui provoque le déclenchement du contenu audio de la ZA (Etat 2). Le score d'intérêt du visiteur



FIGURE 4.20 – La salle des communications du musée des arts et métiers (vue 1)



FIGURE 4.21 – La salle des communications du musée des arts et métiers (vue 2)

Interaction	Etat Initial (0)	Etat 1	Etat 2	Etat 3	Etat 4	Etat 5	Etat 6	Etat 7	Etat 8
ZA		Ecouter suffisante de A	Dire_Oui devant F → activer F	Adaptation: score intérêt des ZA minI et maxI → activer les ref	Arrêt prolongé devant C → activer C	Adaptation: score d'intérêt des ZA Imprimerie > max, → activer les	Ecoute insuffisante de J	Dire_Non devant H → activer H	Adaptation: score intérêt des ZA Imagerie entre minI et maxI → activer les ref
Imprimerie	Neutre (0,5)	Neutre 0,62	Neutre 0,62	Neutre 0,62	Intéressant 0,75	Intéressant 0,75	Intéressant 0,75	Intéressant 0,75	Intéressant 0,75
A Machine à écrire Remington	Active (0)	Active 0,62	Active 0,62	Active 0,62	Active 0,75	Active 0,75	Active 0,75	Active 0,75	Active 0,75
B Presse à Bras	Inactive (0)	Inactive 0,62	Inactive 0,62	Inactive 0,62	Inactive 0,75	Active 0,75	Active 0,75	Active 0,75	Active 0,75
C Machine à graver Conté	Inactive (0)	Inactive 0,62	Inactive 0,62	Inactive 0,62	Active 0,75	Active 0,75	Active 0,75	Active 0,75	Active 0,75
D Presse à cylindre de Ceveau	Inactive (0)	Inactive 0,62	Inactive 0,62	Inactive 0,62	Inactive 0,75	Active 0,75	Active 0,75	Active 0,75	Active 0,75
Télégraphie	Inintéressant (0)	Inintéressant 0	Neutre 0,5	Neutre 0,5	Neutre 0,5	Neutre 0,5	Neutre 0,5	Neutre 0,5	Neutre 0,5
E Télégraphie électrique à aiguille	Inactive (0)	Inactive 0	Inactive 0,5	Inactive 0,5	Inactive 0,5	Inactive 0,5	Inactive 0,5	Inactive 0,5	Inactive 0,5
E Télégraphie aérien de Chadda	Inactive (0)	Inactive 0	Active 0,5	Active 0,5	Active 0,5	Active 0,5	Active 0,5	Active 0,5	Active 0,5
Imagerie	Intéressant (1)	Intéressant 1	Intéressant 1	Intéressant 1	Intéressant 1	Intéressant 1	Intéressant 0,75	Neutre 0,5	Neutre 0
G Fusil Chrono Photo	Active (1)	Active 1	Active 1	Active 1	Active 1	Active 1	Active 0,75	Active 0,5	Inactive 0,5
H Chambre de touristes Enjalbert	Active (1)	Active 1	Active 1	Active 1	Active 1	Active 1	Active 0,75	Active 0,5	Active 0,5
I Chambre à tiroir Daguerre	Active (1)	Active 1	Active 1	Active 1	Active 1	Active 1	Active 0,75	Inactive 0,5	Inactive 0,5
J Fantasmagories de Robertson	Active (1)	Active 1	Active 1	Active 1	Active 1	Active 1	Active 0,75	Active 0,5	Active 0,5

FIGURE 4.22 – Etats de l'exemple.

au tag "Télégraphie" augmente : alors $Score_interest(Visiteur, "Télégraphie")=0+1/2=0,5$. Ainsi, pour toute ZA annotée Télégraphie $Score_interest(Visiteur, ZA) = 0,5$, ce qui les fait passer de la plage inintéressant [0, 0.4] à la plage Neutre [0.4, 0.7]. Ceci implique l'activation par *Adaptation* de ces ZA si elles sont de référence. C'est le cas de la ZA F mais pas celui de la ZA E. Ainsi on inclut la ZA F dans le parcours du visiteur (Etat 3).

$$Visitor.tour(t_1) = A, F, G, H, I, J$$

Le visiteur continue sa visite et s'arrête devant la machine à graver "Conté" (ZA C). Cette machine initialement écartée du parcours y sera ajoutée puisque le visiteur reste le temps que le son de la ZA s'active automatiquement par processus du Contrôle (Etat 4). Le score d'intérêt du visiteur au tag "Imprimerie" ainsi augmente :

$$Score_interest(Visiteur, "Imprimerie")=0,62+(1-0,5)/4=0,75.$$

Ainsi toutes les ZA annotées "Imprimerie" auront un score d'intérêt égale à 0,75. Ce qui les fait passer de la plage Neutre [0, 0.0,7] à la plage Intéressant [0.7, 1]

Ainsi on inclue toutes les ZA annotées "Imprimerie" dans le parcours du visiteur (Etat 5).

$$Visitor.tour(t_2)=\{A,B,C,D,F,G,H,I,J\}$$

Le visiteur poursuit sa visite de la salle, et passe devant la ZA de *Fantasmagories de Robertson*. Bien que cette ZA soit sonorisée car l'imagerie a été marquée comme intéressante par le visiteur, ce dernier passe devant l'objet, mais ne reste pas suffisamment de temps pour écouter les contenus audio qui lui sont liés (Etat 6). Le score d'intérêt du visiteur à l'imagerie est ainsi dévalué : $Score_interest(Visiteur, "Imagerie")=1-1/4=0,75$ (Formule 1.2).

A son passage devant la deuxième machine d'imagerie (*Chambre de touriste Enjalbert*), le visiteur décide d'arrêter le contenu audio en secouant la tête de droite à gauche (geste *DireNon*) (état 7), puis quitte la ZA correspondante. Ce qui provoque la dégradation du score d'intérêt au tag Imagerie : $Score_interest(Visiteur, "Imagerie")=0,75-1/4=0,5$. Ainsi toutes les ZA annotées Imagerie se voient dégradées en intérêt pour le visiteur. En effet, pour chaque ZA annotée imagerie : $score_interest(V,ZA)=0,5$. Ainsi ces ZA deviennent neutres pour le visiteur, et resterons sonorisés uniquement les éléments représentatifs, en l'occurrence les ZA J et H (état 8).

$$Visitor.tour(t_2)=\{A,B,C,D,F,H,J\}$$

Listing 4.16 – Schémas XML de notre cas d'usage

```

1 <audibilityzone id="411" name="machine à écrire Remington" parent="MUSEUM">
2   <area>
3     <shape idref="spher411" type="spheroide" />
4     <spheroideCoords centre="10 5 1" xRadius="5" yRadius="5" zRadius="1" />
5   </area>
6   <soundtoplay>
7     <sound name="RemingtonAmbiant.wav" type="Ambiant">
8       <volume min="0" max="1">0.5</volume>
9

```

```
10     <loop enabled="true" offset="1" />
11   </sound>
12   <soundfrom name="machine à écrire Remington"
13     type="spacialized" locationType="110 57 1" />
14 </soundtoplay>
15 <perception>shared</perception>
16 <ObjectConcepts>
17   <Concepts>
18     <concept id="t1" name="Imprimerie"/>
19   </Concepts>
20 </ObjectConcepts>
21 </audibilityzone>
```

4.9 Extension du modèle pour les visiteurs malvoyants

Si la dimension sonore est importante pour identifier et comprendre le fonctionnement des œuvres du musée, cette dimension sonore est d'autant plus pertinente si le visiteur ne voit pas l'œuvre, ou la voit mal. Entendre le bruit d'une machine permet au public à déficience visuelle de se faire une idée plus complète de l'œuvre. Malheureusement, de nombreuses contraintes empêchent la mise en marche de ces machines au sein du musée, ce qui prive les visiteurs des informations pertinentes sur le fonctionnement de la machine. Ce défaut est encore plus ressenti chez les visiteurs malvoyant, pour qui, le son pourrait jouer un rôle fondamental facilitant la reconnaissance et la découverte des objets.

4.9.1 Scénario de la visite : objectifs

L'objectif d'une visite assistée par un système de RAS pour un malvoyant peut être différent selon la nature des objets qui l'entourent et le lieu dans lequel il se positionne. Le premier objectif est d'accompagner le malvoyant dans un environnement non initialement conçu pour la réception de personnes malvoyantes. L'audioguide est ainsi un accompagnateur du visiteur qui le guidera pour se déplacer entre les différentes salles du musée. Par exemple, il lui permettra d'éviter les obstacles (accompagnement) et émettra des alertes en cas de risque quelconque (signalétique).

Le deuxième objectif d'un tel système, est la découverte des œuvres exposées à travers leurs dimensions sonores au lieu de se limiter à la dimension visuelle. Ceci peut se faire par l'augmentation de l'objet avec un son ambiant qui dépeint l'environnement. Il peut correspondre au son fonctionnel produit par l'œuvre. Il s'agit alors de la bande-son originale de l'œuvre ou d'un enregistrement effectué pendant que l'œuvre est en marche. Il peut aussi s'agir d'un son de synthèse imitant celui que produirait une machine en état de marche avec la possibilité de mixer avec d'autres sons. Le contenu sonore ambiant peut être soit de la musique pour les œuvres dont l'objectif est de pro-

duire un son (Joueuse de tympanon, boîte à musique,...). Il peut aussi être un son de fonctionnement pour les objets qui produisent des sons fonctionnels (horloge, télévision, phonographe), ou un bruit produit lors du fonctionnement d'un objet (moteurs, outils, instruments scientifiques). L'utilisation de la spatialisation de ces contenus audio émanant des diverses œuvres entourant le visiteur procurera à ce dernier un sentiment d'immersion dans le MAM et rendra la visite du musée similaire à celle d'une usine à machine.

Le troisième objectif est un objectif informatif. Le système d'audioguide, en fonction de la position et de l'orientation du visiteur, permet l'envoi de la description vocale (parole ou élocution) de l'objet le plus proche du visiteur. Elle peut aussi englober une histoire, un témoignage, un dialogue, etc.

4.9.2 Modèle de la scène sonore : les types de zones d'audibilité

Afin d'atteindre les différents objectifs décrits dans la section précédente, nous introduisons quatre types de ZA :

Zone d'accompagnement : c'est une ZA qui entoure le visiteur et qui se déplace avec lui au cours de sa visite. Son objectif est de diriger le visiteur malvoyant pour arriver à sa destination.

Zone partagée : c'est une ZA qui est associée à chaque œuvre du musée que l'on souhaite intégrer dans le parcours du visiteur. Dans cette zone, le visiteur percevra un son ambiant émanant d'une source sonore assimilée au centre de la ZA. Plusieurs zones ambiantes peuvent se recouvrir. Ces zones peuvent être soit liées à des objets différents, soit à plusieurs composants du même objet. Le visiteur qui se retrouve au sein d'un lieu de recouvrement de deux ou plusieurs zones ambiantes pourra percevoir simultanément les sons spatialisés provenant des sources sonores différentes.

Zone exclusive : ce type de ZA est essentiellement dédié à la description des objets exposés. Elle se trouve généralement au plus proche d'un objet. Dans cette zone, le visiteur percevra de la parole décrivant l'œuvre concernée. La particularité de cette zone est que son contenu audio ne peut pas se mixer avec d'autres zones même en cas de recouvrement. Cette condition est volontairement imposée pour garder l'attention du visiteur sur l'objet à décrire.

Zone d'alerte : c'est une ZA qui se positionne dans les endroits particulièrement risqués pour un malvoyant ou un aveugle, comme les escaliers, les rampes, etc... Elle entoure également toute œuvre, afin de signaler au visiteur, qu'il s'en approche trop. Le son de cette zone d'alerte est prioritaire sur tous les sons émanant d'une autre source sonore.

4.9.3 Modèle de la scène sonore : algorithme de suivi

Le scénario du visiteur que nous proposons est un scénario qui combine l'immersion audio dans le bruit ambiant et la description orale des œuvres exposées. Lorsque le visiteur pénètre dans l'environnement sonore, la position et l'orientation de sa tête sont automatiquement traquées par le système de positionnement et envoyées au système de spatialisation sonore. Ce dernier récupère le contenu audio adapté à la position et l'orientation de l'utilisateur, puis l'envoie au visiteur à travers son casque binaural. Les règles suivantes régissent le scénario de la visite et la diffusion des contenus audio :

Règle 1 : à chaque accès à une salle d'exposition, le visiteur reçoit un contenu audio lui présentant la thématique de l'exposition et le type d'œuvre qui rentre dans son parcours. Ainsi, il pourra avoir une première idée sur le contenu de la salle.

Règle 2 : lorsque le visiteur pénètre dans une zone partagée, il commence à entendre un son ambiant binaural spatialisé lié à l'objet contenu dans la zone. L'intensité et l'orientation du son sont respectivement proportionnelles à la distance séparant le visiteur de la source sonore, et à l'angle entre le vecteur visuel du visiteur et le vecteur 'visiteur-source sonore'.

Règle 3 : lorsque le visiteur se retrouve dans le recouvrement de plusieurs zones partagées, il entendra simultanément les sons ambiants spatialisés provenant des différentes sources sonores.

Règle 4 : lorsque le visiteur pénètre dans une zone exclusive liée à un objet du musée, le son ambiant lié au même objet est renvoyé progressivement en arrière plan et avec une faible intensité. Ensuite, le texte de description de cet objet commence à s'entendre. Il est également envoyé en binaural.

Règle 5 : lorsque le visiteur pénètre une zone d'alerte, tous les autres sons sont arrêtés et le visiteur commence à entendre une voix qui l'incite à s'éloigner de la zone à risque en lui indiquant ce qu'il faut faire (reculer, aller à droite, etc...). Un son d'alarme retentit et augmente en volume au fur et à mesure que le visiteur se rapproche de la zone à risque (centre de l'objet, mur, escalier, etc...).

Règle 6 : à l'extérieur de toute autre zone, et selon la conception de la scène sonore, une zone d'accompagnement suit le visiteur et le dirige pour naviguer au sein du musée. Elle est notamment utilisée pour se déplacer entre les salles du musée, et entre les espaces de recouvrement des sons ambiants. Les transitions d'une zone à l'autre sont faites d'une manière fluide et progressive, de telle sorte que le visiteur ne sente pas de coupures dans la diffusion des contenus sonores.

4.10 Conclusion

Dans ce chapitre, nous avons présenté dans un premier temps quelques exemples de scénarios dans le but de proposer un modèle de visites le plus général possible. Afin

de modéliser la visite augmentée, nous avons commencé par la structure du paysage sonore en trois couches : la couche physique qui contient les informations géométriques sur l'espace physique et les objets qui le composent, la couche virtuelle et acoustique et la couche sémantique

Dans la couche virtuelle et acoustique, nous avons introduit la notion de zone d'audibilité qui délimite l'espace d'écoute de la source sonore et qui permet au concepteur de gérer l'espace d'écoute en le délimitant par des zones d'audibilité de différents formats. L'ensemble des zones d'audibilités forment le paysage sonore dans lequel évolue le visiteur. Les zones d'audibilité peuvent se juxtaposer les unes avec les autres. Ces zones pourraient être organisées de différentes manières (structures p2p ou hiérarchique par exemple). L'organisation choisie a un impact sur l'efficacité du rendu sonore. L'intérêt d'une approche hiérarchique est qu'elle permet de représenter des structures complexes suivant différents niveaux de résolution. Ainsi le visiteur peut être immergé dans un paysage sonore composé de sons ambiants, de musiques et de commentaires en adéquation avec sa position et son orientation. Par ailleurs, dans la couche sémantique, ces ZA une fois annotées et classifiées par concepts sous forme d'un thésaurus, peuvent être facilement exploités pour déterminer ses centres d'intérêt. Le visiteur quant à lui est indexé par différents types d'informations allant des informations de profils, données par le visiteur au début de sa visite aux informations de tracking permettent de donner sa localisation et son focus, aux signatures d'intérêts qui permettent de définir les intérêts des visiteurs.

Puisqu'une visite est une navigation dans un paysage sonore, nous avons distingué trois niveaux de navigation du visiteur dans le paysage sonore :

- La navigation par le déplacement où le visiteur interagit avec le contenu sonore uniquement par son déplacement et son orientation.
- Le contrôle, où le visiteur contrôle le paysage sonore par ses gestes, il peut contrôler la lecture du contenu sonore comme il peut activer ou désactiver des ZA.
- L'adaptation où le système adapte le parcours du visiteur en cours de sa visite selon ses intérêts.

Dans le futur, ce modèle peut être étendu et amélioré dans plusieurs aspects. Un premier aspect concerne celui de la description des zones d'audibilité. D'autres types de ZA autres que exclusives et partagées peuvent voir le jour suivant l'application. En outre, alors que ces ZA sont définies d'une manière statique dans notre modèle, nous pouvons imaginer des ZA créées dynamiquement par le visiteur durant sa visite. L'association des contenus audio aux ZA peut être repensée, nous pouvons proposer un modèle qui associerait plusieurs sons alternatifs à la même ZA, et que le choix du son diffusé soit établi par le visiteur lui-même ou bien désigné automatiquement selon son profil. En plus, un nouveau protocole de navigation peut être défini à travers de nouveaux canaux d'interaction autres que la position et l'orientation du visiteur. En effet, nous pouvons étendre le modèle pour intégrer d'autres types de contrôle de la scène sonore comme, par exemple, de la parole émise par le visiteur ou aussi des gestes qu'il

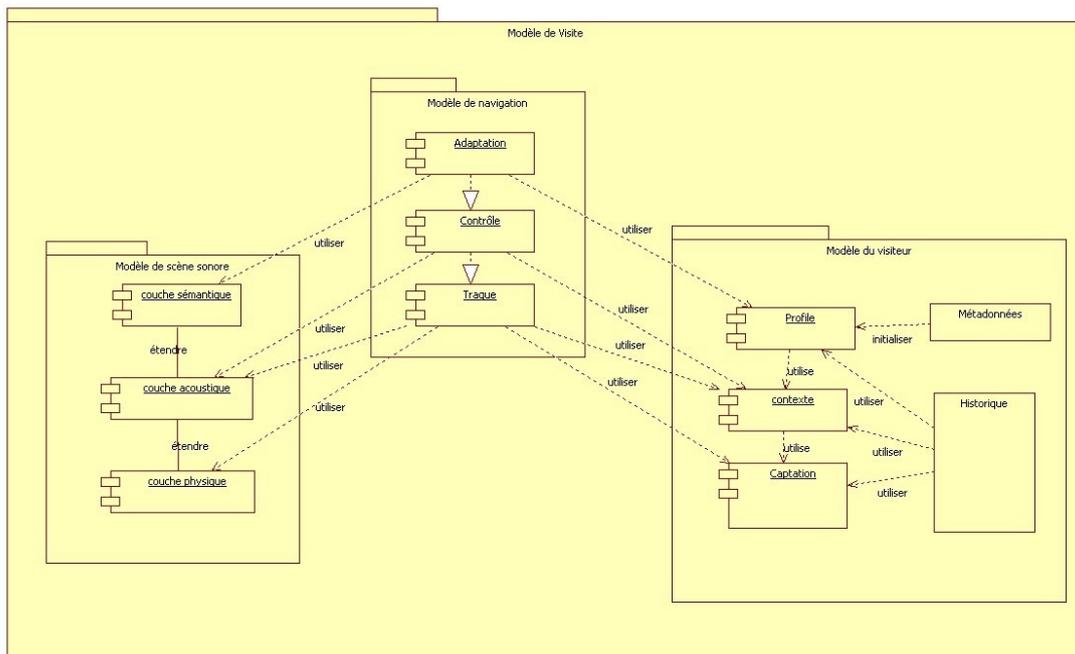


FIGURE 4.23 – Diagramme de composants du modèle de la visite

peut exprimer avec son corps, mais également à travers d'autres dispositifs multimédias intelligents. Par ailleurs, le modèle du visiteur pourrait être étendu pour prendre en compte la visite augmentée du groupe auquel le visiteur pourrait appartenir.

5

Evaluation de la visite augmentée

5.1 Introduction

Ce chapitre décrit la partie expérimentale de notre travail. Elle a pour objectif la mise en application du modèle, des scénarios et du dispositif de l'utilisateur dans une situation réelle : la visite du musée des arts et métiers de Paris. Nous avons besoin de cette étape d'expérimentation pour valider nos hypothèses théoriques, et pour pouvoir évaluer l'acceptabilité de ce genre de système de RAS, jusqu'alors peu connu par les visiteurs du musée.

Ce chapitre consiste alors à faire une évaluation empirique de SARIM et de ses fonctionnalités. Comme nous ne pouvons pas évaluer toutes les fonctionnalités que permet le modèle, nous nous sommes donnés pour objectif de mettre en œuvre et d'évaluer quelques scénarios de visite permettant de préfigurer de nouveaux usages et de nouveaux champs applicatifs de la réalité augmentée, dans lesquels la composante auditive jouerait un rôle majeur. Nous nous posons la question de la réalité augmentée sonore d'un point de vue ergonomique, donc de la façon dont elle est mise en œuvre.

Dans la première section de ce chapitre, nous aborderons la problématique de l'utilisation d'un dispositif de RAS dans le cadre de la visite de musée. Dans cette section, nous évoquerons les critères évalués dans d'autres systèmes de RA appliqués à la visite de musées. Nous nous positionnons par rapport à ces systèmes, puis nous évoquerons nos questions de recherche. Dans la troisième et la quatrième section, nous décrirons les deux expériences qui ont été menées dans deux salles du musée des arts et métiers.

Nous décrirons les objectifs de chaque expérience et son déroulement, et nous procéderons finalement à l'analyse des résultats.

5.2 Évaluation des dispositifs de réalité augmentée

5.2.1 Les pratiques courantes dans l'évaluation de la RA

Avant de mettre au point une évaluation de SARIM, il s'avère utile d'examiner les pratiques courantes de l'évaluation des systèmes de RA. Nous allons étudier trois travaux que nous estimons importants, et qui décrivent des techniques et des méthodes utilisées dans l'évaluation de la RA. Le premier est celui d'Anastassova [Anastassova07], un état de l'art où sont décrits 48 articles portant sur l'évaluation centrée utilisateur des applications industrielles de la RA. Ses auteurs rapportent que la majorité des articles étudiés (38%) traitent la question de l'utilisabilité et seulement 17% se spécialisent dans l'analyse des besoins de l'utilisateur. En termes de méthodologie, les auteurs qui étudient la question de l'analyse des besoins des utilisateurs, le font souvent à travers des interviews ou des questionnaires. Pour les auteurs qui évaluent l'utilisabilité des systèmes de RA, le but est souvent d'étudier la facilité d'usage des interfaces de RA et leur efficacité, en se servant de la méthode du questionnaire. Une autre remarque très importante est que la majorité des évaluations effectuées sont faites en laboratoire et non pas en conditions réelles.

Gabbard et al. [Swan05, Gabbard08] ont aussi proposé une classification des expériences centrées utilisateurs liées à la RA. Ils définissent trois axes complémentaires suivant le domaine de recherche :

- La perception humaine (profondeur de la perception, coordination entre l'œil et la main).
- La performance de la tâche pour des applications spécifiques de la RA (temps, nombre d'erreurs).
- L'environnement collaboratif et le travail coopératif.

Dunser et al. [Dünser08] ajoutent un autre axe à la classification de Gabbard qu'ils ont nommé : "étude de l'utilisabilité des interfaces des systèmes de RA". Ils proposent aussi une taxonomie des méthodologies d'évaluation des systèmes de RA en cinq catégories :

- Les mesures objectives (temps, taux d'erreurs, précision).
- Les mesures subjectives, ou la perception de l'utilisateur, souvent recueillies par des questionnaires.
- L'analyse qualitative qui inclut l'observation formelle des utilisateurs, l'entretien formel.
- La mesure de l'utilisabilité (méthodes heuristiques, analyse des tâches).

- L'évaluation informelle qui inclut les méthodes d'observation ou les retours informels des utilisateurs.

Notre objectif initial était de déterminer les critères évalués dans les tests d'utilisabilité, et les problématiques posées dans l'évaluation des systèmes de RA en général et RA sonore en particulier. Nous avons tenté dans un premier temps de faire le lien avec les interfaces d'IHM mobiles en général et les systèmes de RA mobiles en particulier. Ceci a été entrepris dans le but d'évaluer la possibilité d'utiliser les critères d'utilisabilité de l'IHM pour évaluer les systèmes de RA. Dans un second temps, nous présentons un aperçu des critères évalués dans les audioguides de musée, et comment les chercheurs ont tenté de trouver le lien entre les systèmes de RA mobiles et les audioguides mobiles de musées.

5.2.1.1 RA et utilisabilité

Dans cette section, nous passons en revue quelques travaux de recherche concernant l'évaluation des applications de la RA dans le domaine de l'IHM. La question que nous poserons est : est-ce que les critères d'utilisabilité de l'IHM en général peuvent être appliqués aux technologies de la RA ?

Afin de répondre à cette question, Gabbard et al. [Swan05, Gabbard08] estiment que la combinaison des critères de l'utilisabilité de l'IHM et les études centrées utilisateur constitue non seulement un défi important, mais aussi une opportunité de recherche. Selon ces auteurs, un bon système doit être à la fois utile et utilisable. Pour faire l'intersection des critères de l'utilisabilité d'une interface homme machine en général et un système de RA mobile, les auteurs définissent les caractéristiques orientées utilisateur suivantes :

- La facilité d'apprentissage.
- La rapidité d'exécution des tâches utilisateur.
- Le taux d'erreurs d'utilisation.
- La satisfaction subjective de l'utilisateur.
- Le temps passé par l'utilisateur à se servir du dispositif.

Ils insistent aussi sur l'étude préalable du domaine de l'application de la RA, ce qui consiste à définir comment doit être l'utilisateur et quel est la tâche qu'il doit accomplir.

Dunser et al. [Dünser07a, Dünser07b], tentent aussi de réduire l'écart existant entre le domaine de l'IHM et celui de la RA en matière d'utilisabilité. Ils suggèrent des critères d'utilisabilité de l'IHM qui peuvent être applicables dans le domaine de la RA :

- L'affordance.
- Réduire l'effort cognitif.
- Réduire l'effort physique.
- La facilité d'apprentissage.
- La satisfaction de l'utilisateur.

- La réactivité et la tolérance aux erreurs du système.

5.2.1.2 RA et audioguides

Les audioguides classiques utilisés pour les visites de musées sont sujets à plusieurs types d'évaluation. Ces évaluations ont permis de tester leur utilisabilité sur plusieurs critères tels que l'amusement, l'utilité, la durée d'usage, la facilité d'usage, la distraction, l'acquisition de connaissances, l'équilibre attentionnel, l'interaction sociale, l'ergonomie, la navigation, le système de positionnement, l'interface, le contenu, la personnalisation, la géolocalisation, l'accessibilité, l'impact cognitif, etc. [Proctor03, Woodruff01, Bellotti02].

Une taxonomie de critères d'évaluation des systèmes de RA a été proposée par Damala et al. [Damala09, Sylaiou05]. Trois groupes de critères d'évaluation ont été définis ; les critères technologiques, les critères d'information et les critères logistiques :

- Les critères technologiques : ce sont les critères liés à la technologie choisie pour l'implémentation de l'application du contenu ou de l'interface, ou pour la géolocalisation.
- Les critères de l'information : ce sont les critères liés au contenu interactif, notamment la qualité du contenu, sa durée, la nature du contenu par objectif (amusement, distraction, interaction sociale, etc.).
- Les critères logistiques : ce sont les critères liés à la logistique et l'administration de l'audioguide, à savoir la maintenabilité, le nombre d'audioguides, la formation du personnel, la distribution de l'audioguide, la durée de vie de la batterie, etc.

En outre, Damala a étudié les corrélations entre les critères évalués pour les audioguides classiques et ceux des audioguides de la RA. Elle conclut que le lien est fort concernant les critères suivants : l'orientation et la navigation dans l'environnement du musée, l'orientation et la navigation dans l'application, la facilité d'alterner entre le guide et l'objet, la facilité de l'usage, la distraction et l'équilibre attentionnel, la réaction affective, la transparence, et la facilité de l'utilisation.

Ensuite, [Damala09] définit deux groupes de critères d'évaluation : les critères liés à la RA et les critères non liés à la RA, de l'ordre de l'utilisabilité dans le domaine de l'IHM.

Le premier groupe lié à la RA, est d'une importance primordiale pour l'acceptation du guide augmenté. Ce groupe inclut la facilité de :

- l'orientation et de la navigation dans l'environnement du musée, et la facilité de l'identification des objets commentés.
- la navigation dans le contenu interactif du guide.
- la compréhension de la manière utilisée dans la métaphore de la RA par les utilisateurs non informaticiens.

Le deuxième groupe de critères concerne les professionnels du musée et les concepteurs des applications :

- L'utilité et l'amusement du visiteur, et la satisfaction générale produite par l'utilisation du guide.
- L'impact de la phase post-visite et l'impact de l'apprentissage et aussi l'impact cognitif.
- La question de la distraction du visiteur et son équilibre attentionnel.
- La préférence entre des médias spécifiques.

5.2.1.3 Évaluation centrée utilisateur

Des critères spécifiques ont été utilisés par les chercheurs pour évaluer les systèmes de la RA mobiles, et en particulier les systèmes de RA utilisés dans le domaine de la visite du patrimoine culturel.

Ayant comme objectif d'offrir aux visiteurs du musée une expérience éducative et ludique améliorée, *ARCO* (Augmented Representation of Cultural Objects) [Sylaiou05] est un système de RAV qui a été évalué en termes de composants et d'interfaces. Pour cela, une évaluation heuristique et des solutions cognitives sont employées. Les auteurs ont évalué deux composants du système *ARCO* : la composante liée au contenu, et l'interface de la RA. Afin d'effectuer une évaluation centrée utilisateur, trois catégories d'évaluation principales sont incluses dans cette étude :

L'utilisabilité technique qui se réfère aux aspects perceptifs et physiques de l'interface homme machine (facilité d'apprentissage et de mémorisation, efficacité de l'utilisation, nombre d'erreurs, satisfaction).

L'adéquation entre le contenu et le domaine qui examine la pertinence du contenu et de sa représentation.

L'acceptabilité du système par l'utilisateur qui signifie l'efficacité du système à supporter les besoins cognitifs de la tâche, et la satisfaction fournie.

En plus de ces critères, le niveau de présence de l'utilisateur a été également évalué à travers un questionnaire.

Sylaiou et al. [Sylaiou08] réévaluent le système *ARCO* en termes de groupe de participants. Ils visent à étudier l'impact du type de participants sur l'utilisation du dispositif. Deux groupes de participants ont expérimenté le dispositif : les conservateurs (experts) et les visiteurs (non-expert) du musée. Ils concluent que la meilleure configuration est de faire une évaluation avec des participants experts et des vrais visiteurs du musée (non experts).

Deux travaux paraissent très intéressants dans ce domaine. La première étude est publiée en 2005 par Swan et Gabbard [Swan05, Gabbard08]. Ils examinent 266 publications sur la RA publiées entre 1998 et 2004. En analysant ces articles, ils trouvent que

seulement 14% des articles abordent la question de l'IHM, et seulement environ 8% décrivent une étude formelle centrée utilisateur. La deuxième étude est publiée en 2008 par HITLab en Nouvelle-Zélande [Dünser08]. Elle cite 557 publications sur la RA qui sont publiées entre 1993 et 2007. Seulement 10% de ces travaux incluent une évaluation centrée utilisateur. D'après ces deux études, nous constatons que les évaluations centrées utilisateurs sont peu présentes dans les évaluations des systèmes de RA. Cependant en 2013 ce constat peut ne plus être vrai.

5.2.2 Critères évalués par la RAS

Les auteurs qui ont évalué des applications de RAS dans le cadre de la visite du musée définissent des critères supplémentaires liés au contenu audio, au modèle de la scène sonore et à l'expérience auditive. Dans le projet *ech(o)* [Hatala05], les auteurs définissent deux groupes d'objectifs à évaluer : des objectifs conceptuels et des objectifs fonctionnels. Les objectifs conceptuels portent sur l'amusement, la quantité de données à transmettre et la nature du contenu audio (devinette, jeu de mots,...). Les objectifs fonctionnels définissent les critères de l'utilisabilité du système RAS selon les critères suivants :

- La satisfaction générale.
- Le confort lié à l'utilisation de l'interface tangible de l'utilisateur.
- Le confort lié à l'utilisation du casque de RAS.
- La facilité d'apprentissage.
- La perception de la performance du système (temps de réponse).
- La qualité du contenu (informatif/non informatif, surprenant/prévisible, généralisé/ personnalisé, rigide/ludique).
- La qualité de l'expérience audio (confusion/clair, expérience mécanique/humaine, inutile/précieuse).
- La navigation et le contrôle (capacité de naviguer, se perdre dans le système, certitude de l'état du système).

Dans un autre contexte, Paterson et al. [Paterson10], évaluent un jeu en RAS. Ils évaluent :

- L'engagement et la présence et l'immersion.
- Le contrôle .
- L'interaction.
- L'utilisabilité.
- L'impact de la localisation sur l'engagement et sur l'immersion.

Dans le projet *Ontrack* [Jones08], les auteurs proposent une nouvelle approche pour aider les utilisateurs mobiles dans les tâches de recherche d'itinéraire en intégrant des indices de navigation dans la musique. Ils ont fait passer un test d'utilisabilité en laboratoire, dont les résultats étaient encourageants pour développer la version mobile du prototype. L'étude de l'utilisabilité a porté sur l'efficacité du système, ç.à.d. si les

utilisateurs sont capables de compléter les tâches dans un délai raisonnable, ainsi que sur le retour subjectif des utilisateurs.

A travers la littérature, on constate un manque au niveau de l'évaluation des systèmes de RA. Les évaluations qui sont faites se passent souvent en laboratoire et pas en situations réelles, ce qui rend difficile la vérification des objectifs de l'expérience en dehors du contexte (musée par exemple), surtout s'il s'agit d'un contexte culturel ou ludique.

5.2.3 Questions de recherche

Après avoir passé en revue les questions et critères évalués dans les systèmes de RA étudiés, l'étape suivante consiste à définir les questions de recherche qui vont être étudiées dans la phase d'évaluation. Nous définissons trois groupes de questions de recherche, le premier groupe traite l'apport de la RAS dans la visite de musées. Le deuxième groupe porte sur la manière de l'utilisation de la RAS dans la visite du musée. Le troisième groupe porte sur la conception d'un dispositif de RAS accessible et utilisable.

Le premier groupe de questions traite l'apport de la RAS dans la visite de musées, cela inclut :

- De quelle manière l'interaction dans la RAS du musée peut être différente des autres interactions homme machine ?
- Est ce que la technologie de la RAS peut constituer une alternative intéressante pour l'interaction et la navigation dans les audio guides des musées ?
- Est ce que la RAS permet de faciliter la navigation dans l'environnement du musée ?
- Est ce que la RAS permet de faciliter l'accès à l'information dans l'environnement du musée ?
- Est ce que la RAS permet de :
 - faciliter l'identification et la localisation des objets sonorisés ?
 - rendre la visite plus ludique ?
 - mieux valoriser les objets sonorisés ?
 - accroître l'intérêt porté aux objets sonorisés ?

Le deuxième groupe porte sur l'exploitation de la RAS pour améliorer la visite :

- Comment rendre la technologie de la RAS facile et intuitivement compréhensible même par des utilisateurs novices ?
- Quels genres de scénarios de visite permettront au visiteur d'interagir intuitivement avec les objets du musée ?
- Quel est l'impact de l'interaction directe entre l'utilisateur et les objets sonores sur la visite du musée ?
 - Est ce qu'elle permet d'accéder à plus d'objets ?

- Est ce qu'elle permet d'accroître l'intérêt porté aux objets ? (se sentir acteur de la scène plutôt qu'observateur)
- Quel impact a l'ajout des sons ambiants à la visite ?
- Quel impact pourrait avoir la sonorisation des objets exposés sur la visite du musée et sur le parcours de la visite ?
- Quel impact pourrait avoir la sonorisation des objets exposés sur l'intérêt des utilisateurs envers les objets observés ?
- Quel est l'impact de la sonorisation du patrimoine sur la charge émotionnelle et nostalgique de la visite du musée ?

Le troisième groupe porte sur la conception d'un dispositif de RAS accessible et utilisable. En plus des critères d'utilisabilité traditionnels : la compréhension, la facilité d'apprentissage, la satisfaction, le besoin de flexibilité et de personnalisation, la réactivité, l'amusement de l'utilisateur et l'évaluation subjective de la durée de la visite. Plusieurs questions liées à la RAS sont à aborder :

- Quelle doit être la qualité du contenu sonore pour que les utilisateurs acceptent le casque de RAS pendant la durée de la visite ?
- Est ce que le système de RAS fournit les mêmes repères spatiaux comme le fait la situation de l'audition normale ?

La qualité sonore est un facteur très important dans l'acceptabilité d'un système de RAS. La qualité sonore devrait être assez bonne, pour que les gens puissent porter le casque de RAS pendant de longues périodes sans désagrément ni cacophonie. Ainsi, la représentation pseudoacoustique doit être aussi similaire à l'environnement sonore réel que possible.

L'audition spatiale est un autre facteur qui joue un rôle important dans l'acceptabilité du système de RAS. Le système de la RAS doit fournir les mêmes repères spatiaux comme dans une situation d'audition normale.

En outre, la portabilité du matériel RAS affecte la facilité d'utilisation. Un système RAS devrait exiger peu d'équipements supplémentaires, et doit être facile à porter. De nos jours, les gens deviennent plus tolérants pour se déplacer au quotidien avec des appareils électroniques, comme les téléphones cellulaires et les lecteurs MP3. Beaucoup d'entre eux sont également habitués à porter des écouteurs assez fréquemment.

L'utilisabilité du système a une grande influence sur la compréhension de l'utilisation du système, la facilité d'apprentissage, la satisfaction, le besoin de flexibilité et de personnalisation.

Nous proposons une évaluation empirique avec un protocole d'évaluation qui inclut principalement les méthodes quantitatives, basées notamment des questionnaires et des logs. Les méthodes d'évaluation utilisées, notamment le questionnaire et les logs, sont définies plus en détail dans les sections suivantes. Avant la phase d'évaluation, des tests pilotes ont été menés auprès des visiteurs du musée pour vérifier le bon fonctionnement du matériel et du logiciel à expérimenter.

Notre processus d'évaluation comporte deux expériences dans deux salles différentes du musée, avec des objectifs et questions différents pour chaque expérience. Nous allons dans ce qui suit décrire chaque expérience, les hypothèses et les variables à vérifier, la procédure suivie et les résultats trouvés.

5.3 Expérience 1 : Étude comparative de deux dispositifs dans la salle des communications

L'expérience 1 a pour objectif d'étudier l'utilisabilité de SARIM et son apport à la visite, en le comparant avec l'audioguide actuellement utilisé dans le musée. L'expérience s'est déroulée dans la salle des communications du musée. Le choix s'est porté sur cette salle, parce qu'elle est considérée comme représentative de ce musée. Elle contient des objets qui produisent des sons fonctionnels comme les imprimantes, les téléphones et les télévisions, et d'autres qui, en état de marche, produisent un bruit comme les imprimantes et les appareils photos. C'est une salle rectangulaire qui fait 35m*10m. Les objets qui y sont exposés sont de différentes tailles : les petits objets sont exposés dans des vitrines en verres situées sur les bordures de la salle, et les gros objets sont exposés par terre au centre de la salle.

L'audioguide du musée des arts et métiers¹ se compose d'un boîtier bleu, qui comporte 10 boutons noirs chiffrés de 0 à 9, d'un bouton vert pour les fonctions "mise en marche" et "arrêt", d'un bouton rouge pour la fonction "pause", de deux boutons noirs pour la lecture rapide et un bouton bleu de la fonction "en savoir plus". Le boîtier est équipé également d'un haut parleur permettant d'écouter les commentaires audio et d'une prise jack pour pouvoir utiliser ses propres écouteurs ou casque d'écoute. Il est muni également d'un cordon pour pouvoir l'accrocher autour du cou. Le baladeur est à porter à la main et l'écoute se fait en rapprochant le baladeur à son oreille (sauf si on possède des écouteurs qui ne sont pas fournis).

Les objets du musée qui sont commentés, se trouvent munis d'étiquettes de couleurs avec un nombre écrit par dessus. Ces couleurs désignent le parcours. En effet, trois parcours sont définis : un parcours pour les enfants, deux parcours pour les adultes et les adolescents qui sont :

- Parcours d'1 h : découvrir les objets incontournables du musée en suivant les étiquettes rouges.
- Parcours de 2 h 30 : s'initier à l'histoire des sciences et des techniques en suivant les étiquettes rouges et jaunes.
- Parcours "Liberté" : découvrir un domaine entier du musée ou déambuler à sa guise en suivant les étiquettes rouges, jaunes et bleues.

1. <http://www.arts-et-metiers.net/musee.php?P=229>

Dans notre expérience, nous n'avons pas fait de parcours basés sur la couleur des étiquettes, nous avons expérimenté un seul parcours où le visiteur pouvait saisir le numéro correspondant à toutes les étiquettes associées aux objets de la salle des communications quelle que soit leur couleur.

5.3.1 Méthode

5.3.1.1 Tâche des participants

La tâche des participants à cette expérience consiste à trouver les objets commentés ou/et sonorisés et les découvrir pendant une durée prédéterminée. Nous nous sommes servi des vignettes colorées associées aux objets du musée pour identifier les objets commentés. Les vignettes sont surtout indispensables pour l'utilisation de l'audioguide du musée, et le sont moins dans le cas où le participant expérimente SARIM. Ce dernier sera surtout guidé par le son qui émane de l'objet. Le participant doit visiter librement la salle et découvrir les informations associées aux objets exposés pour une durée déterminée à 10 minutes. Le participant n'est pas sensé connaître la durée de la visite à l'avance. Lorsque le temps imparti est écoulé, l'expérimentateur met fin à la visite, et invite le visiteur à répondre au questionnaire.

5.3.1.2 Contenu sonore

Le son ambiant correspondant à un objet peut être : soit le son produit par cet objet (télévision, phonographe), soit le son de fonctionnement de cet objet (imprimante, appareil photo), soit le son de l'environnement ambiant de l'objet (applaudissements, cris), ou encore, le son d'ambiance correspondant à l'époque ou à l'environnement de cet objet. Les commentaires audio se constituent de descriptions vocales des artefacts. Elles sont racontées par les artefacts eux mêmes ou par leurs inventeurs ou encore, par leurs utilisateurs de l'époque qui expliquent leurs fonctionnements. Les commentaires audio peuvent se présenter aussi sous forme d'une narration, d'un témoignage ou d'un dialogue.

Comme la plupart des objets exposés dans cette salle sont vieux et ne marchent plus, nous n'avons pas pu enregistrer les sons qu'ils produisaient en cas de fonctionnement. Par conséquent, des sons artificiels ont été créés. La création est effectuée soit à l'aide d'un mixage de plusieurs sons, soit en créant des sons similaires, ou des ambiances sonores qui donnent une idée sur l'objet. Pour cela, il a fallu trouver de la documentation sur les objets exposés pour savoir à quoi pourrait ressembler le son qu'ils produisent. Heureusement, le musée des arts et métiers est équipé d'un centre de documentation où nous avons pu trouver les informations nécessaires à la conception des sons. Le contenu sonore créé a bénéficié ensuite d'une phase de traitement des sons, qui est basée sur des sur-échantillonnages, des filtres passe-bas, des retardements des sons, etc.

Tous les sons ambiants sont mis à niveau, ils durent 10 secondes et ils sont joués en boucle.

Les commentaires audio utilisés dans SARIM sont ceux fournis avec l'audioguide du musée. Ce contenu a été créé par la société Sycomore. Le musée des arts et métiers, étant en possession des droits d'utilisation, nous l'a délivré pour faire nos expériences. En utilisant l'audioguide du musée, le visiteur peut écouter deux niveaux de commentaires audio : ceux accessibles en composant le numéro affiché sur l'étiquette posée sur l'objet et ceux accessibles en cliquant sur la touche "en savoir plus" du boîtier de l'audioguide. Nous avons intégré tous les commentaires audio de l'audioguide qui correspondent aux objets sélectionnés dans notre parcours de visite. Tous les commentaires durent 1 min 30. Nous avons créé également un son ambiant correspondant à l'entrée de la salle et un commentaire d'introduction sur l'histoire de ce lieu. Au total, 11 objets de cette salle sont sonorisés, chaque objet est décrit par un son ambiant et un commentaire, soit un total de 24 contenus audio et 25 minutes 30 d'enregistrement.

5.3.1.3 Scène sonore

L'expérimentation s'est déroulée dans la salle des communications du MAM. Dans cette salle (figure 5.1), le choix a été porté sur 10 objets exposés : *la presse à bras*, *la machine à graver de Conté*, *le télégraphe électrique*, *le télégraphe aérien*, *la presse typographique de Marinoni*, *la machine à écrire de Remington*, *le fusil Phonographique*, *la chambre le touriste*, *la chambre de Daguerre*, et *le masque de squelettes*. A chacun de ces objets nous avons associé deux ZA d'une forme sphérique, une ZA partagée et une ZA exclusive (figure 5.1). Les objets sont nécessairement contenus dans leurs propres ZA, mais n'étaient pas nécessairement positionnés au centre de ces dernières. En effet la position des ZA a été soigneusement choisie pour créer des zones partagées qui se croisent, et dans lesquelles le visiteur percevra plusieurs sons émanant de différents objets. Le diamètre de toutes les ZA exclusives était de 3 mètres, sauf pour *la presse typographique de Marinoni* qui était de 5 mètres. Quant aux ZA partagées, le diamètre variait entre 8 et 12 mètres.

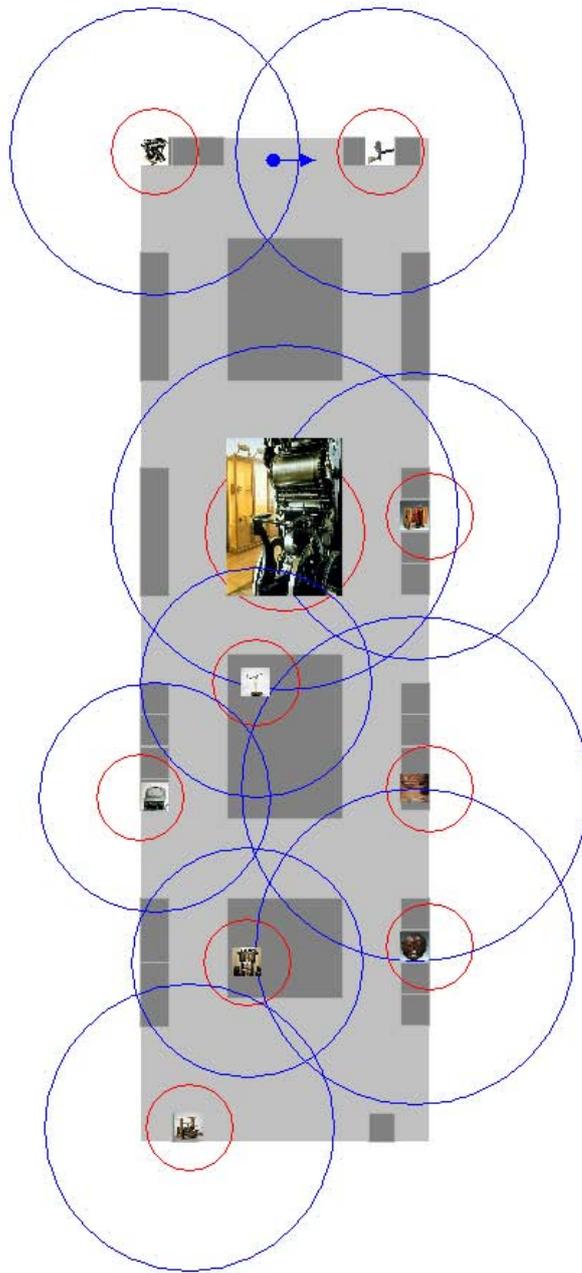


FIGURE 5.1 – Interface de suivi des participants durant l'expérience 1.

5.3.2 Protocole expérimental

5.3.2.1 Participants

Dix huit volontaires non payés ont participé à cette expérience. Ils étaient 9 hommes et 9 femmes. Les participants sont âgés entre 25 et 73 ans avec une moyenne d'âge de 49 ans. Treize participants ont été recrutés sur invitation par mail et 5 participants ont été recrutés sur place parmi les visiteurs du musée. Les participants recrutés au préalable de l'expérience sont invités à passer une condition sur deux pour chaque expérience. Les participants qui sont recrutés sur place sont des visiteurs du musée que l'on a sollicité pour participer à l'expérience. La majorité des visiteurs abordés ont accepté de participer à l'expérience. La moitié des participants avaient déjà visité le musée, mais ils n'avaient pas utilisé auparavant l'audioguide proposé par ce dernier. Aucun des participants recrutés sur place n'avait visité auparavant la salle des communications et tous étaient francophones.

5.3.2.2 Déroulement

Afin de participer à l'expérience, les participants sont invités à remplir un formulaire, disponible en annexe (A.3), où ils communiquent leurs nom et prénom, leur âge, leurs coordonnées (mail ou téléphone), et s'ils ont déjà visité le musée des arts et métiers ou non. Si oui, combien de fois ils l'avaient visité au cours des trois dernières années. Afin d'éviter les parasites et les variables secondaires, les participants qui avaient déjà utilisé l'audioguide du musée des Arts et Métiers ne sont pas sélectionnés pour passer l'expérience. Les participants qui ont visité le musée plus de 3 fois au cours des trois dernières années, ne sont également pas sélectionnés pour passer l'expérience.

L'expérience se déroule en faisant passer un participant à la fois pour des raisons techniques liées au nombre de prototypes que nous possédions. En effet, nous possédons un seul capteur, et pour cela, un seul participant peut passer l'expérience à la fois. Ce qui a été très coûteux en termes de temps.

Pour mener l'expérimentation, aucune suspension des visites dans cette salle n'a été exigée auprès des responsables du musée pendant la période de l'expérience. Il convient de noter tout de même qu'il n'y avait pas énormément de visiteurs présents au moment de notre expérience.



FIGURE 5.2 – Expérience 1 en situation

5.3.2.3 Les variables théoriques

Nous rappelons qu'une variable indépendante (VI) est la variable que nous manipulons directement en tant qu'expérimentateurs. Elle est appelée ainsi parce qu'elle est indépendante de ce que font les participants durant l'expérience, elle est prédéfinie par l'expérimentateur. La variable qui est observée afin de voir si des changements dans la variable indépendante ont un effet sur elle est appelée variable dépendante (VD). Elle est appelée ainsi, par ce que les changements dans celle-ci sont dépendantes des changements de la variable indépendante [Robson94].

Dans cette expérience, la variable indépendante est le dispositif utilisé. Dans la première condition expérimentale, le dispositif étudié est SARIM, dont nous étudions l'uti-

lisabilité et l'apport. Dans la seconde condition expérimentale, le dispositif étudié est l'audioguide actuel proposé par le musée. Le but étant de comparer les deux dispositifs d'un point de vue ergonomique. Nous définissons les variables dépendantes suivantes :

1. L'utilisabilité du dispositif.
2. Le caractère informatif
3. Le caractère ludique.
4. La localisation des objets.

5.3.2.4 Plan expérimental et hypothèses opérationnelles

La description du plan expérimental consiste à énoncer les VI et leurs modalités, tout en indiquant si elles correspondent à des groupes de participants appariés (plan croisé) ou indépendants (plan emboîté). En l'occurrence, notre expérience utilise des groupes indépendants, nous avons donc utilisé des plans expérimentaux emboîtés.

Une hypothèse théorique prévoit une relation de type abstrait entre deux classes de faits, et une hypothèse opérationnelle signifie la traduction de l'hypothèse théorique dans un cas concret précis [Légal07]. Les hypothèses opérationnelles sont les traductions concrètes des hypothèses théoriques. Il s'agit d'essayer de prédire l'influence des VI sur les VD. Le tableau 5.2 présente les hypothèses théoriques et opérationnelles que nous avons formulées et le lien qui les réunit.

5.3.3 Résultats et analyse des données

Dans cette section nous présentons les résultats trouvés qui répondent ou pas aux hypothèses formulées précédemment. Comme nous l'avons évoqué précédemment deux méthodes d'évaluation et de recueil de données ont été utilisées, la méthode des logs et le questionnaire . Nous commençons par présenter les résultats par la méthode des logs, et ensuite, nous présentons les réponses subjectives des participants recueillies à travers le questionnaire.

5.3.3.1 Le suivi avec la méthode des logs

La méthode des logs fait partie des méthodes empiriques que nous avons utilisées dans notre expérience. Elle consiste à observer le comportement en situation d'utilisation du dispositif évalué. Dans notre expérience, nous avons utilisé cette méthode pour capter logiquement, en temps réel et de manière non intrusive, quelques traces de la visite de chaque visiteur. En se basant sur ces logs, nous avons pu découvrir le nombre de ZA que le visiteur a traversé durant son parcours, ainsi que la durée passée dans chacune des ZA traversées.

Hypothèses Théoriques (HT)	Hypothèses Opérationnelles (HO)
SARIM sera plus utilisable que l'audioguide	<ul style="list-style-type: none"> - La compréhension de l'utilisation du dispositif sera plus facile dans la condition expérimentale de SARIM - Les utilisateurs comprendront plus facilement comment utiliser SARIM. - Le nombre d'erreurs de manipulation sera plus important dans la condition expérimentale de SARIM - L'utilisation du dispositif sera perçue plus amusante dans la condition expérimentale de SARIM - SARIM sera perçu moins encombrant. - Le besoin de flexibilité ressenti sera moins important dans la condition expérimentale de SARIM - Le temps de latence entre l'input et le feedback sera moins important dans la condition expérimentale de SARIM.
L'utilisation de SARIM facilitera la localisation et l'identification des artefacts commentés	<ul style="list-style-type: none"> - La localisation des objets sera plus facile en utilisant SARIM. - Les sons de SARIM aideront les visiteurs à se localiser.
SARIM renforcera l'aspect informatif de la visite	<ul style="list-style-type: none"> - La durée d'observation des objets cibles sera plus importante dans la condition expérimentale de SARIM. - L'intérêt déclaré aux objets observés sera plus important dans la condition expérimentale de SARIM. - Le nombre d'objets visités sera plus important dans la condition expérimentale de SARIM.
SARIM renforcera l'aspect ludique de la visite	<ul style="list-style-type: none"> - L'amusement du visiteur sera plus important dans la condition expérimentale de SARIM.

TABLE 5.2 – Corrélation entre les hypothèses théoriques et les hypothèses opérationnelles

Nous constatons qu'avec SARIM, la durée moyenne passée dans les zones exclusives est égale à $7min15 - 17sec$, avec un intervalle de confiance (IC 95%). Dans le cas de l'audioguide classique du MAM, cette durée n'a pas dépassé $5min32 - 40sec$ (figure 5.3). Nous pouvons déduire que le visiteur passe plus du temps à écouter le contenu sonore avec SARIM qu'avec l'audioguide. En effet, dans le cas de l'audioguide, le son n'est déclenché qu'à partir du moment où le visiteur arrive à lire le numéro marqué sur la petite vignette posée sur l'objet. En général les visiteurs, n'arrivent à lire ce numéro qu'en étant très proches de l'objet et donc dans son périmètre de ZA exclusive.

Nous constatons aussi que la durée moyenne passée dans les zones partagées avec SARIM est égale à $2min33 - 15sec$, avec un intervalle de confiance (IC 95%). Dans le cas de l'audioguide classique du MAM, cette durée est égale à $4min8 - 38sec$ (figure 5.4). Ce constat peut être interprété par le fait que le SARIM arrive à attirer les visiteurs à se rapprocher des objets et passer plus de temps à écouter leurs descriptions. Ce qui contribue à leur donner plus d'informations sur les objets exposés et leur permet de visualiser l'objet de plus près et d'apercevoir ainsi plus ses détails.

Concernant le nombre d'objets visités, nous ne trouvons pas de différences significatives entre les deux dispositifs. En effet, la moyenne du nombre de zones visitées en utilisant SARIM est $5,8 - 1,4$ contre $4,8 - 0,7$ pour l'audioguide. Cela est dû probablement au scénario de la visite. En revanche, nous pouvons déduire que la simulation de la localisation n'a pas entravé le bon déroulement de l'expérience.

Analyse des données

Pour chacune des deux systèmes ; Audioguide et SARIM, nous avons effectué trois mesures pour chaque participants : le temps total passé en zones exclusives, le temps total passé en zones partagées et le nombre d'objets visités. Pour mesurer la pertinence de nos résultats, nous avons effectué des analyses de variances (ANOVA) pour chacune des deux systèmes.

En comparant le temps total passé en zones exclusives entre les deux systèmes, nous remarquons une différence significative entre les deux systèmes ($F=5,090$, $P=0,041$). De même, la comparaison du temps total passé en zones partagées, nous remarquons une différence significative ($F=4,762$, $P=0,048$). Par contre, la différence n'est pas significative entre le nombre d'objets visités dans les deux systèmes ($F=1,475$, $P=0,246$).

Ensuite, nous avons voulu vérifier l'influence sur les résultats des tests effectués, de l'âge des participants et de leurs genres.

Pour l'âge, nous avons répartis les participants dans chacun des systèmes en deux groupes : âgés de plus de 30ans et âgés de moins de 30ans. Pour l'audioguide, la différence entre les deux groupes était significative pour le temps passé en zones exclusives ($F=11,494$, $P=0,014$), mais également en zones partagées ($F=10,852$, $P=0,016$). Ceci peut soit présager une réelle différence de comportement entre les deux groupes d'âge, soit être causé uniquement par le fait que les deux groupes ne sont pas équilibrés en nombres de participants. En effet le groupe des moins de 30ans et le tiers du

groupe des plus de 30ans. Pour le nombre d'objet, la différence entre les groupes d'âge n'est pas significative ($F=0,036$, $P=0,854$).

Pour SARIM, par contre, les différences entre les deux groupes d'âge ne sont pas significatives que ça soit pour le temps passé en zones exclusives ($F=1,737$, $P=0,244$), ni pour le temps passé en zones partagés ($F=2,360$, $P=0,185$), ni pour le nombre d'objets visités ($F=2,184$, $P=0,199$).

Pour le genre, nous avons répartis les participants dans chacun des systèmes en deux groupes : hommes et femmes. Pour l'audioguide, la différence entre les deux groupes était non significative que ça soit pour le temps passés en zones exclusives ($F=0,003$, $P=0,955$), pour le temps passés en zones partagés ($F=0,013$, $P=0,911$), ou pour le nombre d'objet ($F=4,668$, $P=0,074$). Pour SARIM, le même constat est fait, puisque la différence entre les deux groupes était non significative que ça soit pour le temps passés en zones exclusives ($F=4,437$, $P=0,089$), pour le temps passés en zones partagés ($F=4,794$, $P=0,080$), ou pour le nombre d'objet ($F=4,263$, $P=0,093$).

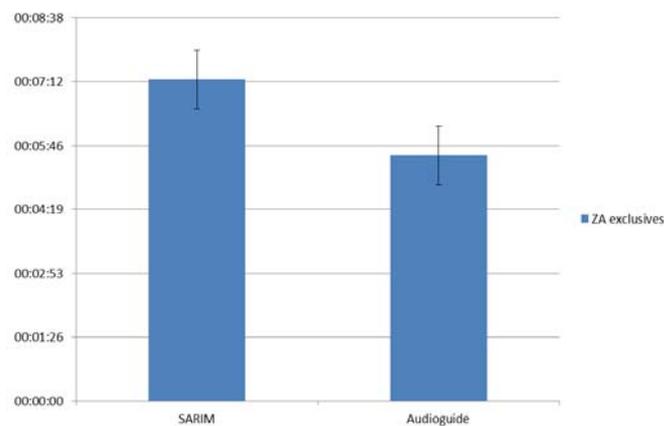


FIGURE 5.3 – Temps passé dans les zones exclusives pour les deux dispositifs (IC95%).

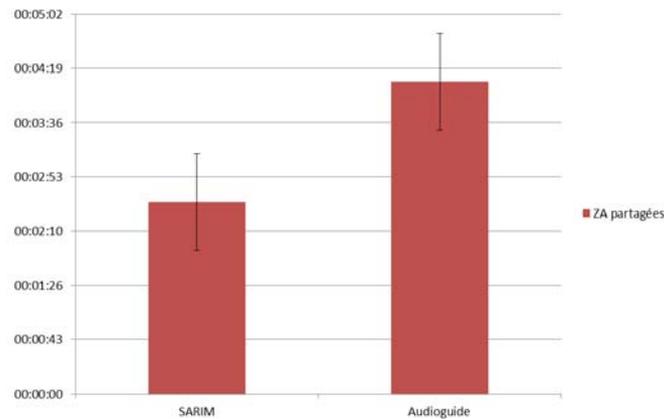


FIGURE 5.4 – Temps passé dans les zones partagées pour les deux dispositifs (IC95%)

5.3.3.2 Les données descriptives du questionnaire

Les questionnaires et entretiens permettent le recueil de données subjectives relatives aux attitudes, aux opinions des utilisateurs, et à leur satisfaction [Scapin97]. Le questionnaire constitue la deuxième méthode empirique d'évaluation que nous avons utilisée dans cette expérience. Le questionnaire est important pour récolter les données descriptives. Il n'est pas coûteux à mettre en place et permet de donner une vue générale sur les résultats, mais manque cependant de profondeur. C'est pour cette raison que dans notre expérience nous ne nous basons pas uniquement sur cette méthode, mais nous utilisons aussi la méthode des logs.

Après avoir passé l'expérience, les participants sont invités à répondre à un questionnaire, disponible en annexe (A.4.1), composé de 25 questions. La plupart des questions sont fermées, où le participant doit choisir entre quatre possibilités de réponse. D'autres questions sont ouvertes. Elles permettent aux participants de s'exprimer librement sur le dispositif évalué. Le questionnaire comporte aussi des questions conditionnelles, des questions dichotomiques dont la réponse est oui ou non, ainsi que des questions directes. Tous les participants doivent répondre aux mêmes questions dans le même ordre et posées par la même personne. Tous les participants répondent au questionnaire immédiatement après l'expérience pour que tous les participants soient dans le même état d'esprit. L'expérimentateur n'est en mesure de répondre aux questions des participants qu'une fois que ces derniers ont répondu au questionnaire.

Pour les questions fermées, nous avons choisi d'utiliser l'échelle de Likert à 4 points (choix), et non pas celle à 5 points. En effet, cette dernière contient une réponse neutre qui peut créer une confusion entre les réponses " je suis neutre" et "je n'ai pas d'opinion" [Albaum97]. Notre échelle comporte les réponses suivantes : "Pas du tout d'accord", "Pas d'accord", "Plutôt d'accord", "Tout à fait d'accord". Nous avons utilisé aussi l'échelle de Likert à 3 points, pour répondre à certaines questions, et donc les ré-

ponses sont de l'ordre de "Pas du tout", "Je pense que oui", et "Absolument". Avant de choisir les questions qui doivent figurer dans le questionnaire, nous avons défini une série de 25 questions (tableau 5.4) qui nous permettront de répondre à nos VDT.

L'utilisabilité de SARIM

Nous avons évalué l'utilisabilité des deux dispositifs en termes de compréhension, de facilité d'apprentissage et d'utilisation, de satisfaction et de réactivité. En termes de compréhension, 77% des participants ont réussi à comprendre l'utilisation des deux dispositifs avec un avantage pour SARIM. En effet, 23% des participants ont été tout à fait d'accord que le fonctionnement de SARIM est compréhensible. En termes de satisfaction, l'avantage est très marqué pour SARIM. En effet, alors que 15% des participants ont été insatisfaits de l'audioguide, ils n'étaient que 4% pour SARIM. De plus, 65% sont très satisfaits de SARIM contre 20% pour l'audioguide, et 4% des participants ont répondu qu'ils ne sont pas du tout satisfaits de l'audioguide contre 0% pour SARIM. La réactivité est également un grand avantage pour SARIM. En effet, la totalité des participants ont trouvé que SARIM est réactif, contre uniquement 56% pour l'audioguide. De plus, 22% répondent que l'audioguide n'est pas du tout réactif. De la même manière, la totalité des participants ont jugé SARIM facile à apprendre. Parmi eux, 63% ont même été tout à fait d'accord pour l'affirmer. Dans le cas de l'audioguide, 26% des participants ne l'ont pas trouvé facile à apprendre.

En additionnant ces 4 critères d'utilisabilité des deux dispositifs, nous remarquons que SARIM est jugé plus utilisable que l'audioguide du musée (figure 5.5).

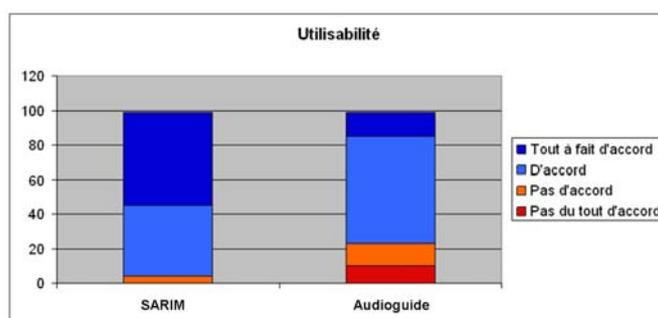


FIGURE 5.5 – Résultats du questionnaire sur l'utilisabilité des deux dispositifs, SARIM et l'audioguide du musée

Le caractère ludique de SARIM

Tous les participants qui ont testé SARIM l'ont trouvé amusant. 44% des participants l'ont même trouvé très amusant. Pour l'audioguide, les réponses étaient miti-

gées. Alors que 55% des participants ne l'ont pas trouvé amusant, 11% l'ont trouvé très amusant.

Le caractère informatif de SARIM

Nous entendons par caractère informatif, l'apport informatif que procure le dispositif sur les artefacts observés, ainsi que l'intérêt porté envers ces artefacts. L'objectif est ainsi de comparer la quantité d'informations apportée par SARIM et l'audioguide du musée, puis d'évaluer l'intérêt que procure chacun des dispositifs aux artefacts sonorisés. Cette VD est déjà évaluée dans la méthode des logs. L'objectif de poser des questions sur cette variable dépendante dans le questionnaire, est de valider (ou pas) les résultats des logs. Le nombre de participants ayant trouvé SARIM plus informatif est légèrement plus élevé que celui de l'audioguide. Ils étaient 88% dans le cas de SARIM et 83% dans le cas de l'audioguide. Par contre, 11% des participants ayant testé l'audioguide ne l'ont trouvé pas du tout informatif, alors qu'ils étaient 0% dans le cas de SARIM. Ce qui donne un léger avantage à ce dernier.

La localisation des objets du musée par SARIM

L'objectif ici est de mesurer la capacité des participants à localiser les objets plus ou moins facilement dans les deux conditions. La totalité des participants ayant testé SARIM ont réussi à localiser les objets facilement. 66% d'entre eux ont même trouvé cela très facile. Dans le cas de l'audioguide, 22% n'ont pas réussi à trouver les objets sonorisés facilement. 11% ont même trouvé cela très difficile. Et, ils étaient 33% à estimer que c'était très facile de localiser les objets. Par ailleurs, la question Q21 (tableau 5.4) a été posée uniquement aux participants qui ont expérimenté SARIM. Elle a pour objectif de mesurer l'impact des sons ambiants spatialisés sur la facilité de la localisation des artefacts. Tous ont trouvé que ce type de contenu sonore les a aidé à localiser les artefacts. Ils étaient 89% à être tout à fait d'accord.

5.3.4 Discussion

Dans cette expérience, nous avons voulu évaluer SARIM en le comparant avec l'audioguide du musée. L'évaluation avait pour objectif de vérifier l'utilisabilité des deux dispositifs, leur apport sur le plan informatif et ludique ainsi que sur la facilité de localisation des artefacts commentés.

Au niveau de l'utilisabilité, nous avons déduit à partir des réponses du questionnaire que les deux dispositifs sont utilisables avec un avantage pour SARIM sur l'audioguide du musée. Durant la séance du questionnaire, quelques participants ont exprimé leurs avis par rapport à l'utilisabilité de SARIM : *''C est pratique d avoir les mains libres''*. D'autres ont formulé des remarques, comme une participante qui a trouvé que le ni-

veau de son n'était pas le même pour tous les artefacts "*Pour certaines machines le son est moins fort que pour d'autres*", ou encore des remarques sur le parcours "*Plus de zones d'intersection serait intéressant*". Quelques participants qui ont testé l'audioguide du musée se sont exprimés sur leur expérience, en le comparant avec la visite classique sans l'audioguide. Un participant a dit : "*Ca donne plus de liberté qu'une visite classique de choisir ce qu'on veut vraiment voir*". Ils ont également exprimé leurs remarques sur l'audioguide, un participant a dit : "*J'avais testé l'audioguide géolocalisé du Louvre la semaine précédente. Je suis mécontent. Je n'ai pas aimé la forme des commentaires*". Un autre participant a dit : "*C'est plus intéressant avec les images projetées sur la borne vidéo*", ou encore sur le contenu "*Les commentaires sont longs*", et aussi sur la compréhension du fonctionnement de l'audioguide, un participant a dit "*Je n'ai pas compris le code de couleurs*" et un autre a dit : "*Le bouton bleu ne fonctionnait pas*".

En ce qui concerne la facilité de localisation, les résultats du suivi avec les logs ont été consolidés par le questionnaire pour déduire que l'utilisation de SARIM facilite la navigation dans le musée et facilite la localisation des artefacts commentés. Les participants ayant testés SARIM ont été séduits par la manière de navigation qu'il procure. Une participante a dit : "*C'est le son qui oriente la visite, et non pas l'inverse!*", une autre évoque la navigation en disant : "*Ne pas avoir à taper les numéros des objets est une bonne idée. La visite nous dirige!*". Les résultats prouvent également, que cette manière de navigation a nettement facilité la localisation des artefacts commentés. En effet, l'utilisateur trouve plus rapidement les artefacts commentés vu qu'il est guidé et orienté par les sons. Ce qui n'est pas le cas avec l'audioguide. De là, nous pouvons valider les hypothèses théoriques que nous avons énoncé. Notamment, que l'utilisation de SARIM facilite la localisation et l'identification des artefacts commentés.

Concernant l'aspect informatif, les résultats obtenus par la méthode des logs démontrent que les participants manifestent plus d'intérêts et passent plus de temps devant les artefacts commentés en utilisant SARIM. Les résultats montrent également que les participants visitent plus de zones en utilisant SARIM qu'en utilisant l'audioguide. Étant donné que chaque zone porte une information sonore particulière, nous déduisons que les participants qui ont utilisé SARIM ont eu accès à plus d'informations. Ces résultats ont été confirmés par les réponses des participants à travers le questionnaire. En effet les participants manifestent plus d'intérêts envers les artefacts observés dans la condition du SARIM. Ils manifestent également plus d'implication en écoutant plus attentivement les commentaires audio en utilisant SARIM. Ce qui nous permet de valider les hypothèses théoriques et opérationnelles liées à l'apport informatif du SARIM.

Et enfin, concernant l'aspect ludique, les résultats du questionnaire ont démontré que les participants ont trouvé SARIM nettement plus amusant que l'audioguide. Et par conséquent, l'hypothèse théorique liée à l'aspect ludique de SARIM peut être validée.

Numéro de la question	Question
Q1	Êtes-vous habituellement intéressé(e) par ce genre d'exposition ?
Q2	La visite vous a-t-elle paru intéressante ?
Q3	D'après vous, combien de temps a-t-elle duré ?
Q4	La visite vous a-t-elle paru amusante ?
Q5	Dans l'ensemble, les objets que vous avez observés vous ont semblé intéressants ?
Q6	D'après vous, quel a été l'impact de la presse rotative située au centre de la salle dans l'histoire de l'imprimerie ?
Q7	Vous êtes vous senti(e) projeté(e) dans les époques correspondant aux objets que vous avez observés ?
Q8	Vous êtes-vous senti(e) à l'aise avec l'utilisation du dispositif ?
Q9	Comment avez-vous trouvé la réactivité du système ?
Q10	A quelle vitesse pensez-vous que la plupart des gens apprennent à utiliser ce système ?
Q11	Personnellement, avez-vous trouvé le dispositif facile à utilisé ?
Q12	Avez-vous compris comment écouter les commentaires audio liées aux objets observés ?
Q13	Si oui, comment ?
Q14	Vous est-il arrivé de faire des erreurs de manipulation du dispositif durant la visite ?
Q15	Si oui, vous avez pu corriger vos actions ?
Q16	Vous est-il arrivé de déclencher des sons que vous ne vouliez pas écouter ?
Q17	Si oui, vous avez pu corriger vos actions ?
Q18	Ces erreurs de manipulation vous ont-elles aidé(e) à mieux comprendre le fonctionnement du dispositif ?
Q19	Dans l'ensemble, vous avez vous compris comment utiliser le dispositif ?
Q20	Vous avez réussi à localiser les objets à observer ?
Q21	Les sons du casque vous-ont-ils aidé(e) à trouver les objets à observer ?
Q22	Dans l'ensemble, Vous avez trouvé le dispositif ?
Q23	Durant la visite, avez-vous ressenti le besoin de pouvoir contrôler le dispositif autrement que par vos déplacements ? Pourquoi et comment ?
Q24	D'une manière générale, êtes vous satisfait(e) du dispositif ?
Q25	Si un tel système se généralisait, aimeriez-vous pouvoir l'utiliser à nouveau lors de vos futures visites de musées ?

TABLE 5.4 – Les Questions du questionnaire de la première expérience

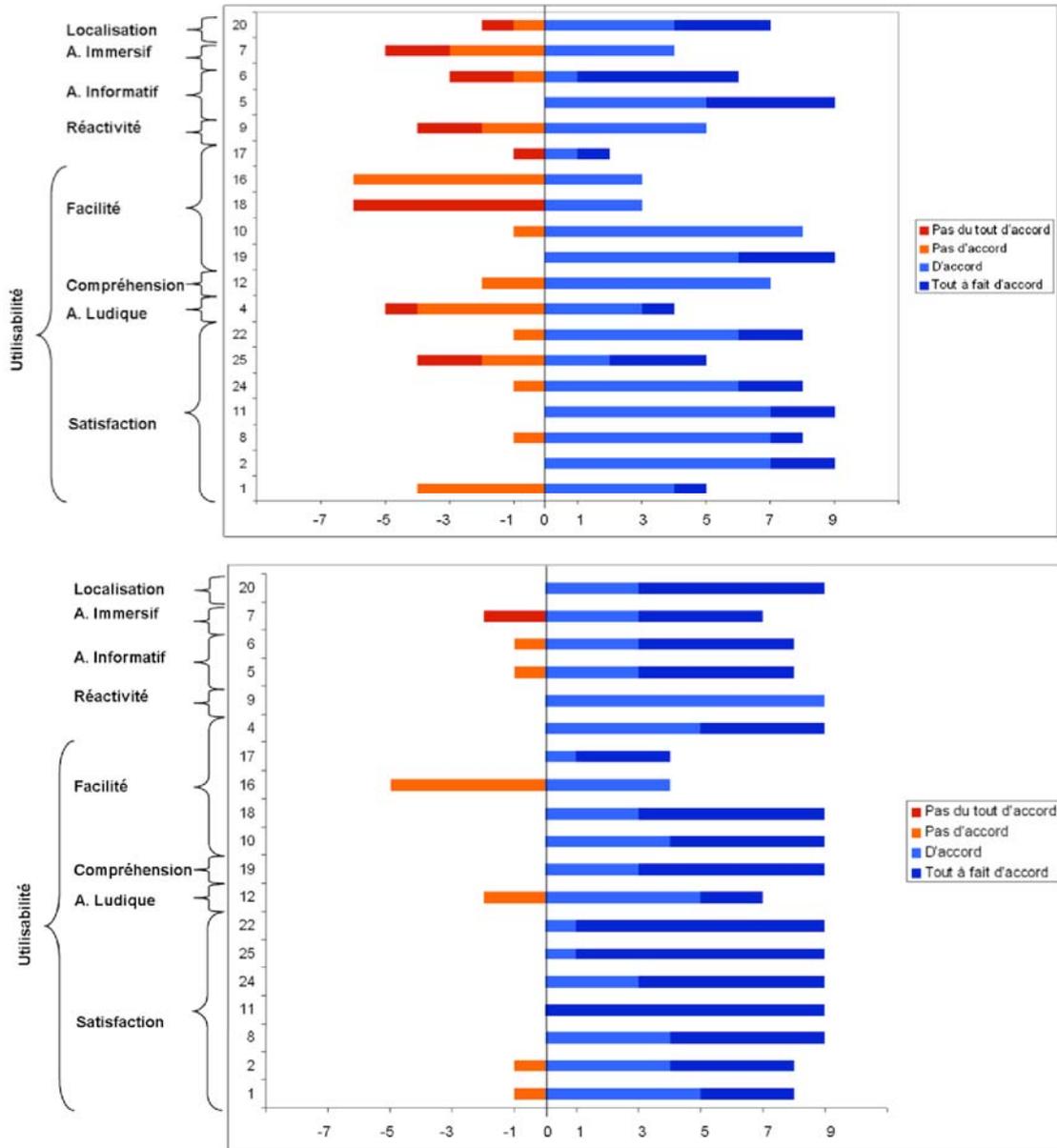


FIGURE 5.6 – Résultats du questionnaire pour les deux conditions : l'audioguide (en haut) et SARIM (en bas)

5.4 Expérience 2 : Étude de deux configurations de scènes sonores dans la salle des automates

Cette expérience vise à étudier l'impact de deux configurations de zones d'audibilité (ZA) dans la scène sonore en utilisant SARIM. La première condition expérimentale est celle des zones croisées, où le visiteur se situe dans l'intersection de toutes les ZAs. Il est donc invité à écouter les sons correspondants à toutes les zones qui se croisent. Si un son retient son attention, il n'a qu'à fixer l'artefact qui le produit pendant quelques secondes pour que tous les autres sons s'arrêtent s'il manifeste plus d'intérêt pour cet artefact, une bande sonore correspondant à la description vocale de cette artefact lui sera envoyée. Dans la deuxième condition, les zones sont disjointes. Le visiteur se trouve à l'extérieur de toutes ses zones. Il ne pourra entendre qu'un seul son à la fois. Pour ce faire, il est invité à fixer chaque objet quelques secondes pour déclencher le son associé. En s'attardant plus sur le même objet, le contenu audio de sa description lui sera transmis.

Les artefacts qui sont exposés dans cette salle sont des automates à musique. Le visiteur ne devant avoir aucune idée préalable du nombre, de la forme ou du nom des automates sonorisés parmi ceux exposés dans cette salle, il aura pour tâche de désigner les automates qui émettent de la musique, de les localiser, et d'écouter les descriptions vocales associées. Les musiques sont envoyées au visiteur d'une manière spatialisée pour l'aider à les localiser, et pour qu'il soit immergé dans l'environnement de cette salle. La salle des automates étant relativement petite, nous n'avons pas eu le besoin de capter la position du visiteur. Les automates sont également disposés d'une manière circulaire. Pour cela, nous avons positionné le visiteur au centre de la salle afin qu'il puisse percevoir tous les automates simplement en tournant sa tête. Le choix a été porté sur cette salle pour plusieurs raisons :

- Les sources sonores sont proches géographiquement.
- Les automates produisent de la musique.
- Les sons produits par les automates sont similaires.
- Les sons produits par les automates ne sont pas connus d'avance par les visiteurs.

5.4.1 Méthode

Le matériel utilisé est identique à celui de la première expérience (voir image 3.9). Le participant est assis sur les escaliers au centre de la salle, et navigue dans la salle en tournant la tête sur le plan horizontal. L'ordinateur qui servait pour appliquer la technique du magicien d'Oz dans la première expérience est utilisé par l'expérimentateur pour s'assurer du bon déroulement de l'expérience. En effet, un programme a été développé pour simuler la navigation du participant dans la salle. Un avatar représentant le participant change de direction si le participant tourne sa tête.

Tâche

La tâche consiste à trouver les objets correspondant aux musiques écoutées et à les observer pendant une durée prédéterminée. Le participant va entendre des musiques provenant de certains des automates exposés. S'il repère un objet qui l'intéresse, il suffira d'orienter sa tête dans sa direction et de le fixer quelques secondes pour écouter le son qu'il émet. Après avoir écouté le morceau de musique d'un automate, il pourrait écouter le commentaire audio qui lui est relatif, s'il continue de regarder dans sa direction. Cependant, s'il en a envie, il peut à tout moment passer à l'observation d'un autre objet en tournant la tête vers celui-ci. Le participant est libre d'explorer la scène comme il le voudra. Au bout de cinq minutes l'expérience s'arrête, pour lui poser des questions sur ce qu'il vient de vivre. Les consignes données aux participants en amont de leurs visite sont disponibles en annexe (A.2).

Contenu sonore

Les contenus sonores utilisés dans cette expérience sont de deux types : la musique des automates et les commentaires. Les musiques des automates nous ont été transmises par le service communication du musée des arts et métiers. Ces musiques ont été enregistrées par des professionnels en faisant marcher les automates. Il s'agit de 5 enregistrements musicaux correspondants aux cinq automates exposés dans cette salle. Les commentaires audio, ont été préparés et enregistrés par l'expérimentateur en se basant sur les informations fournies par le centre de documentation du musée des arts et métiers.

Scène sonore

L'expérimentation s'est déroulée dans la salle des automates. Le choix s'est porté sur 4 automates musicaux dont nous possédons les musiques : *la joueuse de Tympanon*, *la boîte à musique*, *l horloge avec tympanon*, et *l horloge hollandaise*.

Dans cette expérience, nous nous intéressons à l'orientation du visiteur envers ces objets. Nous distinguons deux situations : la première est "En focus de l'automate X", qui signifie que le lacet de X par rapport à l'axe de vue du visiteur est inférieur à un seuil θ choisi par le concepteur. On dit dans ce cas que le visiteur est dans la zone de l'automate X. La deuxième situation est "hors focus" dans laquelle le visiteur n'est orienté vers aucun des automates sonorisés. On dit dans ce cas le visiteur est à l'extérieur de toutes les zones. Nous avons choisi pour cette expérience $\theta = \pi/9$. Deux configurations de sonorisation sont alors comparées. La première est celle appelée "Intersection". Elle stipule qu'en situation de "hors focus", l'utilisateur écoutera un mixage de tous les sons émanant des différents objets sonores, et qu'en situation de "En focus de l'automate X", le visiteur écoutera au début un mixage de tous les sons. Il écoutera après un seuil de temps T_1 exclusivement le son ambiant lié à l'automate X, avant d'écouter en-

suite après un seuil de temps T_2 le commentaire de description de cet automate X. La deuxième configuration est celle appelée "indépendance". Elle stipule qu'en situation de "hors focus", le visiteur ne percevra aucun contenu audio. C'est seulement quand il sera "En focus d'un automate X" qu'il percevra le ambiant, puis après un seuil de temps T_2 il percevra le commentaire de description de cet automate X. Les seuils de temps T_1 et T_2 ont été fixés à 4 secondes chacun.

5.4.2 Protocole expérimental

5.4.3 Participants

Quatorze sujets non rémunérés ont participé à cette expérience. Ils étaient 8 hommes et 6 femmes. Les participants sont âgés entre 25 et 63 ans avec une moyenne d'âge de 44 ans. Dix participants ont été recrutés sur invitation par mail, et 4 participants ont été recrutés sur place parmi les visiteurs du musée. Les participants recrutés au préalable de l'expérience sont invités à passer une condition sur deux pour chaque expérience. Dix participants sur quatorze ont déjà visité le musée, mais ils n'ont jamais utilisé l'audioguide proposé par le musée. Tous les participants qui sont recrutés sur place n'avaient pas encore visité la salle des automates, n'utilisent pas l'audioguide du musée, et sont des visiteurs francophones.

5.4.3.1 Déroulement

Comme dans la première expérience, avant de participer à l'expérience, les participants doivent répondre au formulaire de recrutement, disponible en annexe (A.3). Après avoir passé l'expérience, les participants sont invités à répondre à un questionnaire (voir annexe (A.4.2)), composé de 18 questions. Le questionnaire comporte des questions et des choix de réponses utilisant l'échelle de Likert en 4 points.

Par crainte de perturbation de l'expérimentation, les visites ont été suspendues dans la salle des automates pendant l'expérience. L'expérience dure cinq minutes, suivies de cinq minutes pour répondre au questionnaire.

5.4.3.2 Variables et hypothèses

Dans cette expérience, la variable théorique indépendante est la disposition des zones d'audibilité (ZA), ou l'intersection des zones d'audibilités (ZA) dans la scène sonore. Nous allons étudier l'utilisabilité et l'apport de chaque configuration de ZA. Dans la première condition de cette expérience toutes les ZA se croisent, et dans la deuxième condition les ZA sont toutes disjointes. Le but étant de comparer SARIM dans les deux conditions d'un point de vue ergonomique.

Les variables théoriques dépendantes de cette expérience sont :

1. L'utilisabilité de SARIM dans cette salle.
2. Le caractère informatif.
3. Le caractère ludique.
4. La localisation des objets.

Comme dans la première expérience, cette expérience utilise des groupes indépendants de participants, et des plans expérimentaux emboîtés. Le tableau 5.6 présente les hypothèses théoriques et opérationnelles que nous avons définies pour cette expérience.

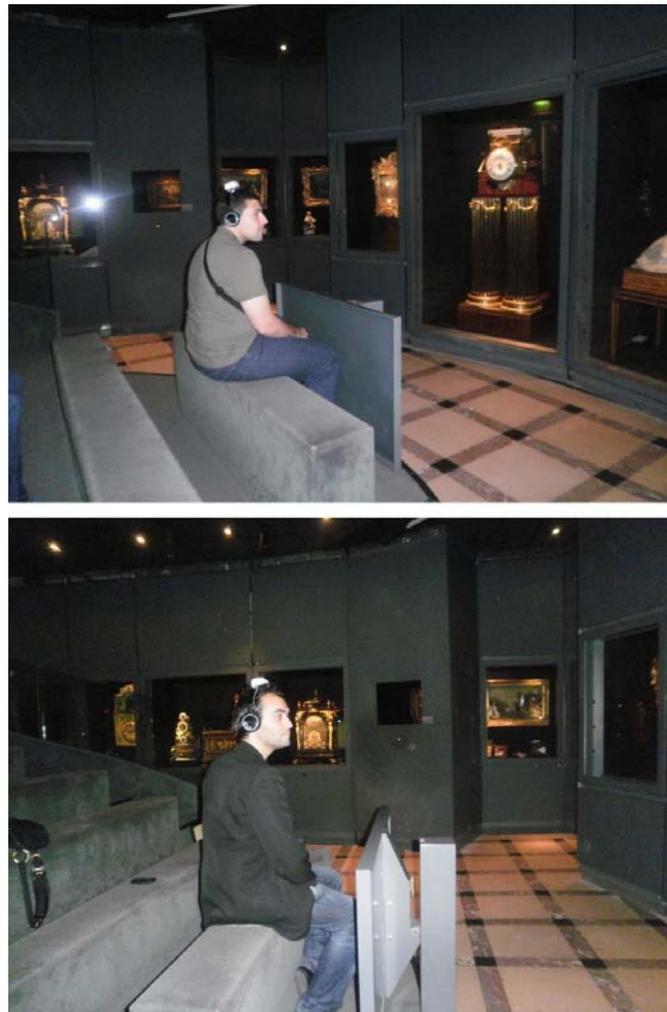


FIGURE 5.7 – Expérience 2 en situation

Hypothèses Théoriques	Hypothèses Opérationnelles
SARIM sera plus utilisable si les ZA se croisent	<ul style="list-style-type: none"> – L’utilisation de SARIM sera plus facile à comprendre si les ZA se croisent – Le nombre d’erreurs de manipulation est moins important si les ZA se croisent – L’utilisation de SARIM paraîtra plus facile si les ZA se croisent – L’utilisation de SARIM paraîtra plus amusante si les ZA se croisent – L’utilisation de SARIM paraîtra moins encombrante si les ZA se croisent. – Le visiteur ressentira moins le besoin de flexibilité si les ZA se croisent – Le visiteur trouvera SARIM plus réactif si les ZA se croisent.
L’intersection des zones permettra de rendre la visite plus instructive	<ul style="list-style-type: none"> – La durée d’observation des objets cibles sera plus importante si les ZA se croisent. – L’intérêt déclaré aux objets observés sera plus important si les ZA se croisent. – Le nombre d’objets visités sera plus important si les ZA se croisent.
L’intersection des zones permettra de rendre la visite plus amusante	L’amusement du visiteur sera plus important si les ZA sont croisées.
L’intersection des ZA facilitera la localisation et l’identification des objets visités	<ul style="list-style-type: none"> – Les sons joués simultanément lorsque les ZA se croisent aideront les visiteurs à se localiser – Les visiteurs arriveront plus facilement à localiser les objets sonorisés si les ZA se croisent.

TABLE 5.6 – Corrélation entre les hypothèses théoriques et les hypothèses opérationnelles

5.4.4 Résultats et analyse de données

5.4.4.1 Résultats des logs

Dans cette expérience, nous traçons le focus du visiteur pour voir s'il est dans la zone de chaque automate ou s'il est hors zone, ça veut dire qu'il n'est dans la zone d'aucun automate. Les intervalles de confiance affichés sur les figures sont calculés pour un degré de confiance de 95%. Nous calculons le moment où le visiteur est orienté vers un objet (en focus) et le moment où il n'est orienté vers aucun objet (hors focus).

La durée moyenne passée hors focus pour les deux conditions expérimentales, montre que les participants passent $2\text{min}5\text{s}$, 16s hors focus dans le cas des zones croisées et seulement 32s 16s dans le cas des zones indépendantes. Cela peut être interprété par une difficulté constatée chez les participants à localiser les œuvres sous cette condition. Dans le cas des zones indépendantes, les participants passent beaucoup plus de temps en focus. Cela signifie qu'ils trouvent plus facilement les automates qui produisent de la musique, et passent plus de temps à les écouter.

Le calcul de la durée moyenne passée en focus montre que les participants passent $4\text{min}28\text{s}$ 8s dans la condition des zones indépendantes et $2\text{min}55\text{s}$ 8s dans la condition des zones croisées. Le fait de passer plus de temps à observer les objets peut être interprété par, premièrement, l'intérêt que portent les participants envers les automates observés dans la condition des zones indépendantes, et deuxièmement, cela confirme le résultat précédent sur la difficulté de localiser les automates dans la condition des zones croisées. Les participants en condition de zones croisées passent moins de temps à observer les automates qu'ils doivent chercher.

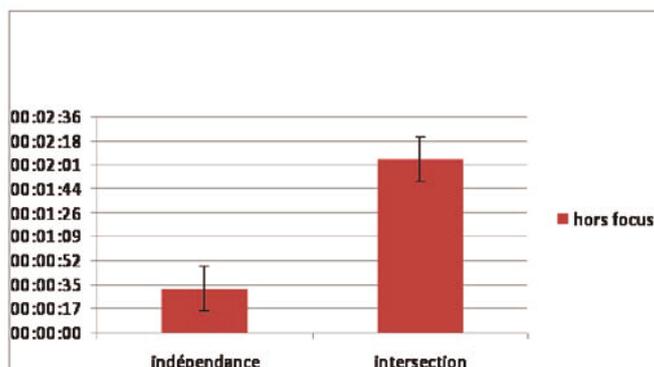


FIGURE 5.8 – Durée de visite quand le visiteur n'est orienté vers aucun objet (hors focus)

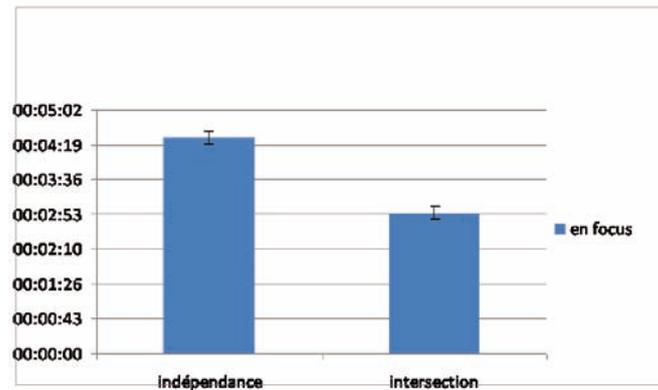


FIGURE 5.9 – Durée de visite quand le visiteur est orienté vers un objet (en focus)

Analyse de données

Pour chacune des deux conditions : intersection et indépendance, nous avons effectué deux mesures : le temps total passé hors focus et le temps passé en focus. Dans un premier temps nous avons vérifié l'influence sur les résultats des tests effectués dans la salle des automates de l'âge des participants, de leurs genres, et du fait que certains participants avaient auparavant expérimenté le système dans la salle des communications. Nous avons ainsi effectué des analyses de variances (ANOVA) pour chacune des deux conditions indépendance et intersection.

Pour l'âge, nous avons répartis les participants dans chacune des conditions en deux groupes : supérieur à 30ans et inférieur à 30ans. Dans la condition d'indépendance, aucune différence significative n'a été détectée entre les deux groupes d'âge, ni pour le temps en focus ($f=0,151$, $p=0,713$), ni pour le temps hors focus ($f=0,151$, $p=0,713$). Le même constat est fait pour la condition d'intersection ; ($f=0,793$, $p=0,413$) pour le temps en focus, ($f=1,147$, $p=0,332$) pour le temps hors focus.

Pour le genre, nous avons répartis les participants dans chacune des conditions en deux groupes : femmes et hommes. Dans la condition d'indépendance, aucune différence significative n'a été détectée entre les femmes et les hommes, ni pour le temps en focus ($f=1,567$, $p=0,265$), ni pour le temps hors focus ($f=1,567$, $p=0,265$). Le même constat est fait pour la condition d'intersection ; ($f=0,174$, $p=0,693$) pour le temps en focus, ($f=0,072$, $p=0,798$) pour le temps hors focus.

Pour l'impact de l'utilisation du système SARIM dans la salle des communications, nous avons répartis les participants dans chacune des conditions en deux groupes : les participants ayant utilisé auparavant SARIM, et ceux ne l'ayant pas utilisé auparavant. Dans la condition d'indépendance, aucune différence significative n'a été détectée entre les femmes et les hommes, ni pour le temps en focus ($f=0,057$, $p=0,820$), ni pour le temps hors focus ($f=0,057$, $p=0,820$). Le même constat est fait pour la condition d'intersection ; ($f=0,002$, $p=0,964$) pour le temps en focus, ($f=0,021$, $p=0,888$) pour le temps hors focus.

Après avoir vérifié que l'influence des trois facteurs précédents et non significatif sur les mesures, nous avons comparé les deux conditions l'une à l'autre. Ainsi, en comparant la durée d'écoute en focus entre les deux conditions indépendance et intersection, nous remarquons une différence significative : ($f=244,781$, $p=2,398E-09$). De la même manière, la différence est significative entre les deux conditions pour la durée hors focus ($f=239,899$, $p=2,692E-09$).

5.4.4.2 Résultats du questionnaire

Le dépouillement du questionnaire pour les deux conditions expérimentales est donné par les diagrammes de la figure 5.11.

L'utilisabilité de l'intersection des zones

Nous avons évalué l'utilisabilité des deux configurations de zones d'audibilité, en termes de compréhension, de facilité d'apprentissage et d'utilisation, de satisfaction, de réactivité et du besoin de flexibilité et de personnalisation. En termes de compréhension, nous considérons que les résultats ont été parasités par l'expérience précédente, donc nous n'allons pas les prendre en compte. En effet, la majorité (11) des participants recrutés commencent par passer la première expérience et 4 parmi eux ont déjà utilisé SARIM dans la salle des communications.

En terme de facilité d'apprentissage, la totalité des visiteurs ayant testé SARIM dans la condition de ZA indépendantes ont trouvé le système facile à apprendre (28% très facile, 72% facile). Alors que 14% de ceux ayant testés SARIM dans la condition de ZA croisées l'ont trouvé difficile à apprendre. Le reste des participants en condition de ZA croisées étaient partagés entre très facile (57%) et facile (14%). Par ailleurs, 57% des participants à la condition de ZA croisées disent qu'il leur est arrivé de déclencher des sons qu'ils ne voulaient pas écouter (Q12), alors que personne n'a déclaré cela dans la condition d'indépendance de ZA. On en déduit que l'utilisation de SARIM dans la condition de l'indépendance des zones est plus facile que dans la condition de ZA croisées.

En terme de satisfaction, alors que tous les participants à la condition d'indépendance ont trouvé l'expérience intéressante (72%) ou très intéressante (28%), 14% des participants à la condition d'intersection l'ont trouvée moyennement intéressante. En outre, la totalité des participants à la condition d'indépendance ont déclaré être très à l'aise (57%) ou à l'aise (43%) dans l'utilisation du système, contre 14% des participants à la condition d'intersection. Toujours en terme de satisfaction, 72% des participants à la condition d'indépendance trouvent SARIM plutôt pratique dans la condition d'indépendance, et aucun participant dans cette condition ne trouve le système moins pratique, alors qu'ils sont 28% des participants à trouver le système plutôt encombrant dans la condition d'intersection. Finalement, 72% des participants à la condition d'in-

dépendance aimeraient beaucoup utiliser un tel système dans leurs prochaines visites dans un musée, et 28% aimeraient assez l'utiliser dans le futur, alors qu'ils n'étaient que 43% dans la condition d'intersection à être très motivés contre 28% qui étaient assez motivés. 14% des participants dans la condition d'intersection se sont déclarés neutres par rapport à cette question.

En combinant les résultats des réponses liées à la satisfaction, nous trouvons que 33% des participants des deux conditions se disent très satisfaits du dispositif. Mais on décèle une préférence pour la condition d'indépendance puisque aucun des participants ne s'est déclaré insatisfait dans cette condition, alors qu'ils étaient 14% dans la condition d'intersection.

Pour la réactivité de SARIM dans les deux conditions expérimentales : 100% des participants aux deux conditions trouvent que le dispositif est très réactif.

Besoin de Flexibilité et de personnalisation

En ce qui concerne le besoin de flexibilité et de personnalisation, 57% des participants qui ont testé SARIM en condition d'intersection des zones ont eu le besoin de le personnaliser contre 72% dans la condition de l'indépendance des zones. Ceux qui ont eu ce besoin de contrôler le dispositif autrement ont exprimé leur besoin en terme de navigation entre les types de son, le réglage du volume, l'arrêt du son. Parmi les réponses des participants à la condition de l'intersection des zones, nous trouvons : *"oui. Pour arrêter la cacophonie, c est stressant , "en appuyant sur une touche sur le casque", "oui. Pouvoir écouter les sons en se déplaçant dans la salle", "oui. Pour déclencher le commentaire avant la n de la musique", "avec un boîtier de commande"*. Les participants de la condition indépendance se sont exprimés en disant : *"oui. Pour mettre en pause. Mais j ai compris plus tard comment faire", "oui. Pour augmenter le volume", "oui. Pour le volume. En touchant au casque", "oui. Pour couper la musique et passer directement aux commentaires. Avec un bouton sur le casque"* A partir des réponses exprimées par les participants dans les deux conditions, nous concluons que 12% des participants à la condition de l'intersection des zones trouvent que SARIM dans cette condition n'est pas très utilisable (Figure 5.10). Alors que la totalité des participants de la condition d'indépendance trouve le système utilisable ou plutôt utilisable. Ce qui montre une préférence pour la condition de l'indépendance en termes d'utilisabilité. Nous pouvons déduire, alors, que SARIM est plus utilisable dans la condition des zones indépendantes (figure 5.10).

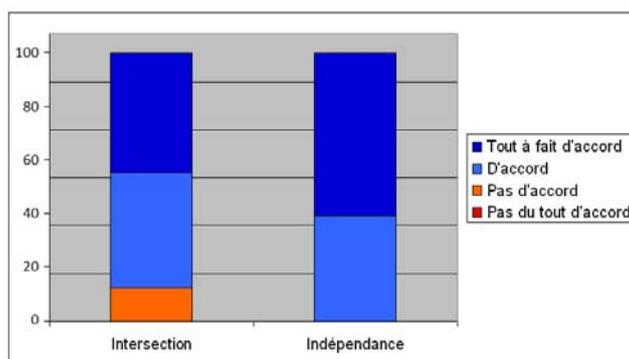


FIGURE 5.10 – L'utilisabilité de SARIM dans les deux conditions expérimentales

L'apport ludique de la disposition des zones

14 % des participants à la condition de l'intersection des zones trouvent que l'expérience n'est pas du tout amusante, un jugement qui n'a pas été approuvé par les participants à la condition des zones indépendantes(0%). 14 % des participants dans les deux conditions trouvent que l'expérience veçue n'est pas très amusante. 43% des participants à la condition de l'intersection des zones trouvent que l'expérience est plutôt amusante, contre 86% chez les participants de la condition de l'indépendance des zones. 28% des participants à la condition de l'intersection des zones trouvent que l'expérience est très amusante, contre 0% chez les participants de la condition de l'indépendance des zones

L'apport informatif de la disposition des zones

14 % des participants dans les deux conditions trouvent que les objets observés sont peu intéressants. 14% des participants à la condition de l'intersection des zones trouvent que trouvent que les objets observés sont plutôt intéressants, contre 72% de chez les participants de la condition de l'indépendance des zones. 72% des participants à la condition de l'intersection des zones trouvent que les objets observés sont très intéressants, contre 14% chez les participants de la condition de l'indépendance des zones

L'apport de la disposition des zones sur la localisation

14 % des participants à la condition de l'intersection des zones répondent qu'ils ont réussi à localiser les objets plutôt difficilement, contre 0% de chez les participants de la condition de l'indépendance des zones. 72% des participants à la condition de l'intersection des zones répondent qu'ils ont réussi à localiser les objets plutôt facilement, contre 43% de chez les participants de la condition de l'indépendance des zones. 14%

des participants à la condition de l'intersection des zones répondent qu'ils ont réussi à localiser les objets très facilement, contre 57% de chez les participants de la condition de l'indépendance des zones .

5.4.5 Discussion

Dans cette expérience, nous avons voulu tester deux configurations de scènes sonores, basées sur la disposition des zones d'audibilité. Nous avons étudié l'utilisabilité et l'apport de deux conditions sonores, la première condition est basée sur des zones d'audibilité qui se croisent, et qui permettent au participant se trouvant dans le recouvrement de plusieurs zones, d'écouter plusieurs sources sonores simultanément. La deuxième condition est basée sur des zones d'audibilité indépendantes qui ne se croisent pas, et dans lesquelles le visiteur entendra une seule source sonore à la fois. L'évaluation avait pour objectif de vérifier l'utilisabilité des deux configurations de scènes sonores, leur apport sur le plan informatif et ludique et sur la facilité de localiser les automates à musique.

Au niveau de l'utilisabilité, nous avons trouvé à partir des réponses du questionnaire que les deux dispositifs sont utilisables avec un avantage pour la configuration de zones indépendantes. Durant la séance du questionnaire, quelques participants ont exprimé leurs avis par rapport à l'utilisabilité de SARIM dans les deux conditions : *"C est agréable de ne pas avoir à appuyer sur des boutons !"*, *"C est bien de ne pas avoir à appuyer sur un clavier"*. Par ailleurs, nous avons eu des avis différents sur le fait d'écouter plusieurs sons à la fois. En effet, il y a des participants à la condition des zones croisées qui ont exprimé leur mécontentement vis à vis de l'écoute simultanée de plusieurs sons : *"Je n ai pas apprécié que les musiques soient jouées en même temps pendant le passage d un objet à un autre !"*, et aussi nous avons eu des retours évoquant le mot "cacophonie". Nous avons eu également des retours opposés dans la condition des zones indépendantes par des participants qui auraient souhaité écouter plusieurs sons en même temps : *"J aurais aimé pouvoir écouter plusieurs sons en même temps !"*. En ce qui concerne la facilité de localisation, les résultats du suivi avec les logs ont été consolidés par le questionnaire pour déduire que la scène sonore avec des zones indépendantes facilite la localisation des automates. Les participants qui ont testé la condition des zones croisées ont eu du mal à localiser les automates : *"C est difficile de différentier les angles correspondant aux sons !"*. Concernant l'aspect informatif et didactique, les résultats obtenus par la méthode des logs montrent que les participants manifestent plus d'intérêts et passent plus de temps devant les automates commentés dans la condition des zones indépendantes. Concernant l'aspect ludique, les résultats du questionnaire sont moins clairs et permettent de dire que l'utilisation de SARIM avec la condition des zones indépendantes est légèrement plus amusante que dans le cas des zones croisées. La figures 5.11 représentent les résultats du questionnaire dans les deux conditions expérimentales.

Numéro de la question	Question
Q1	Êtes-vous habituellement intéressé(e) par ce genre d'exposition ?
Q2	Dans l'ensemble, l'expérience vous a semblé(e) intéressante ?
Q3	D'après vous, combien de temps a-t-elle duré ?
Q4	La visite vous a-t-elle paru amusante ?
Q5	Dans l'ensemble, les objets que vous avez observés vous ont semblé intéressants ?
Q6	Vous êtes vous senti(e) projeté(e) dans les époques correspondant aux objets que vous avez observés ?
Q7	Vous êtes-vous senti(e) à l'aise avec l'utilisation du dispositif ?
Q8	Pourquoi ?
Q9	Comment avez-vous trouvé la réactivité du système ?
Q10	A quelle vitesse pensez-vous que la plupart des gens apprennent à utiliser ce système ?
Q11	Personnellement, avez-vous trouvé le dispositif facile à utilisé ?
Q12	Vous est-il arrivé de faire des erreurs de manipulation du dispositif durant la visite ?
Q13	Si oui, vous avez pu corriger vos actions ?
Q14	Vous avez réussi à localiser les objets à observer ?
Q15	Dans l'ensemble, Vous avez trouvé le dispositif encombrant ou pratique ?
Q16	Durant la visite, avez-vous ressenti le besoin de pouvoir contrôler le dispositif autrement que par vos déplacements ? Pourquoi et comment ?
Q17	D'une manière générale, êtes vous satisfait(e) du dispositif ?
Q18	Si un tel système se généralisait, aimeriez-vous pouvoir l'utiliser à nouveau lors de vos futures visites de musées ?

TABLE 5.8 – Les Questions du questionnaire de la deuxième expérience

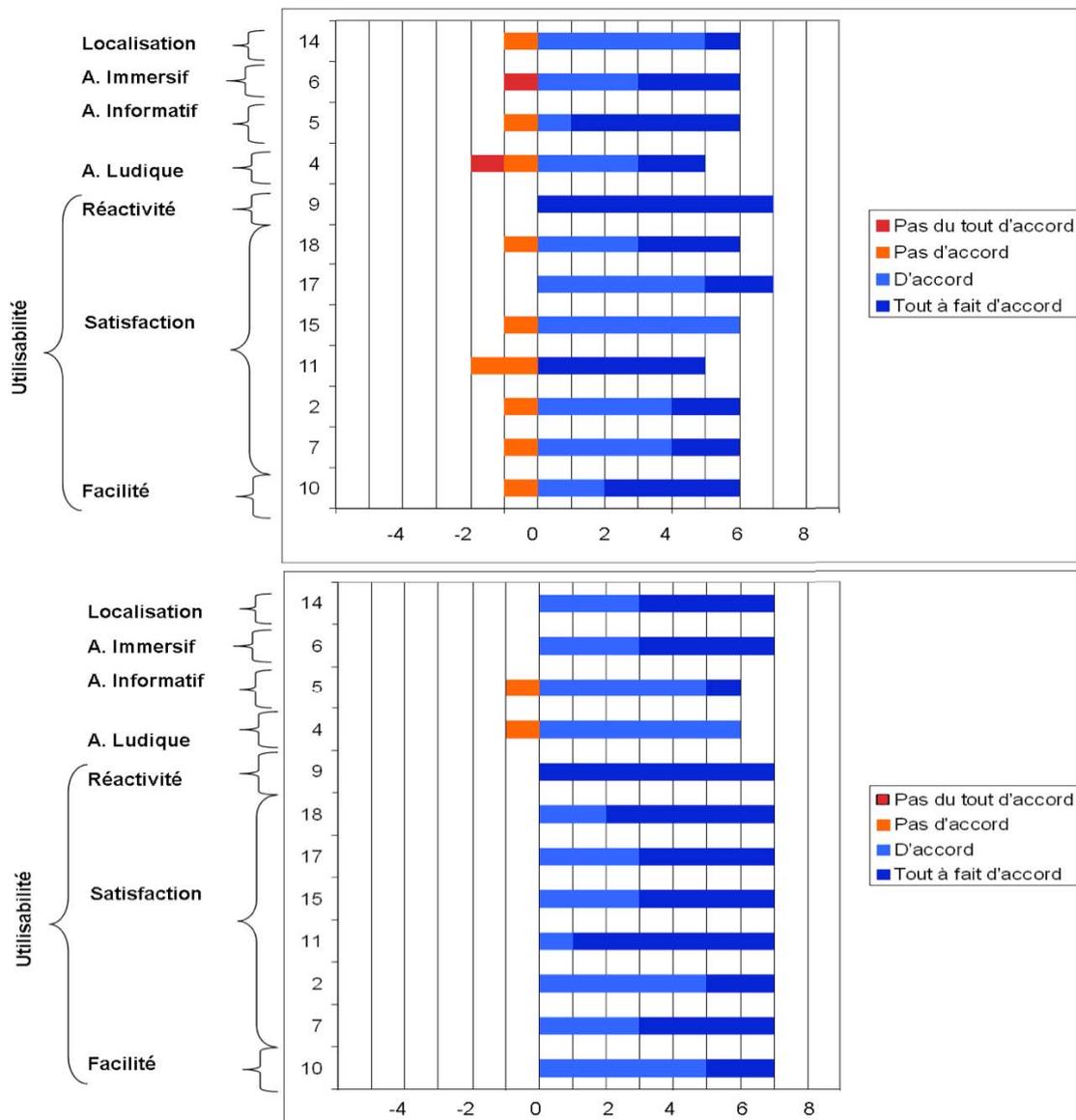


FIGURE 5.11 – Résultats du questionnaire pour les deux conditions : l'intersection des zones (en haut) et indépendance des zones (en bas)

5.5 Conclusion

Dans ce chapitre nous avons décrit la partie expérimentale de notre travail. Cette partie vise à évaluer l'application du modèle de la visite proposée, des scénarios de visite et du dispositif de l'utilisateur dans une situation réelle. Il s'agit d'une évaluation empirique de SARIM et de ses fonctionnalités.

Après avoir défini la problématique de l'utilisation d'un dispositif de RAS dans le cadre de la visite de musée, nous avons décrit les deux expériences qui ont été menées dans deux salles du musée des Arts et métiers.

La première expérience vise à étudier l'acceptabilité et l'apport de SARIM dans la visite du musée en le comparant avec un audioguide classique. Les résultats de l'expérience montrent que SARIM est utilisable et offre un apport à la fois informatif et ludique à la visite du musée. Il permet aussi de faciliter la localisation des œuvres et accroît l'intérêt que porte le visiteur sur les œuvres du musée.

La deuxième expérience consiste à évaluer SARIM dans une petite salle, où les œuvres sont placées les unes à côté des autres, et produisent de la musique. Dans cette expérience, l'objectif était d'évaluer deux architectures de scènes sonores. Dans la première, la scène sonore est composée de zones d'audibilité qui se chevauchent. Ce qui signifie que plusieurs sources sonores seront activées et donc audibles simultanément. La deuxième architecture est basée sur des zones d'audibilité séparées et disjointes. L'idée est de vérifier si l'intersection des zones sera utilisable dans le cas où les œuvres sont très proches et où les contenus sonores consistent en des documents musicaux du même genre. Les résultats ont montré que la configuration de zones croisées a plutôt gêné les participants qu'elle ne les a amusés. Elle les a partiellement empêchés d'admirer la musique produite par ces œuvres à cause de la cacophonie due au mélange sonore. Les participants ont pu plus facilement localiser les œuvres sonores et ont trouvé qu'ils ont appris plus de choses dans la configuration des zones séparées. Cependant, nous avons eu des participants dans la condition des zones séparées qui auraient préféré entendre plusieurs sources sonores à la fois. Ce qui a laissé prétendre qu'une troisième configuration issue du mélange des deux premières en utilisant la notion de clusters sonores pourrait constituer un bon compromis. Malheureusement, cette troisième configuration n'a pas pu être testée à cause du manque de contenus sonores. Cette idée pourrait également être mieux évaluée dans des environnements qui comportent plusieurs objets sonorisés.

Ces expériences nous ont permis de récupérer des premiers retours de visiteurs de SARIM en situation réelle. Nous n'avons pu tester qu'une partie restreinte des scénarios de visites parmi ceux offerts par notre modèle de visite. Toutes les configurations d'intersection des zones d'audibilité n'ont pas pu être évaluées également, surtout la relation d'hierarchie entre les zones d'audibilité que nous estimons importante pour la scène sonore. Nos travaux futurs doivent porter sur l'évaluation des autres configurations possibles avec notre modèle de visite dans un autre cadre expérimental où la dimension sonore serait capable d'augmenter et d'enrichir la réalité.

6

Conclusion et travaux futurs

Dans ce travail, nous avons étudié la question de la réalité augmentée sonore dans la visite des musées, en particulier le musée des arts et métiers, en proposant le système SARIM. Le but est d'enrichir la visite du musée avec une expérience individuelle qui immergerait le visiteur dans un environnement sonore créé par des sons spatialisés qui émanent virtuellement des œuvres exposées. Dans ce chapitre, nous dressons un bilan du travail effectué dans cette thèse. Nous évoquons également un logiciel auteur, que nous avons commencé à concevoir. Cet outil auteur vise à concevoir, créer et éditer des scénarios de visite. Enfin, nous abordons d'autres directions de recherche qui nous paraissent intéressantes pour la suite de ce travail.

6.1 Bilan

Notre but dans ce travail a été de concevoir un système de RA sonore capable d'améliorer et d'enrichir la visite de musées à travers la dimension sonore grâce à la localisation continue du visiteur. Notre objectif a été également de diminuer au maximum l'effort d'attention requis de la part de l'utilisateur pour interagir avec l'environnement sonore. La visite au musée des arts et métiers a été choisie comme cas d'étude, parce que nous avons constaté un manque d'information sonore dans la présentation des œuvres du musée. En l'occurrence, les sons de fonctionnement liés aux machines exposées qui peuvent être attachants et captivants, mais également porteurs d'information pour le visiteur.

Avant de mener une expérimentation grandeur nature, il a été primordial de commencer par le développement d'une preuve de concept pour notre dispositif du visiteur. Ceci est d'autant plus essentiel vu les défis techniques matériels et logiciels nécessaires à la mise en place d'un tel système. Une architecture matérielle et logicielle du prototype a été développée, et plusieurs expérimentations ont été menées avec différents capteurs de mouvement. Ces premières expérimentations ont été accomplies au laboratoire mais aussi au musée avec l'objectif de permettre le passage à l'auto-localisation dans l'environnement du musée. La réalisation de cette preuve de ces premières expérimentations nous a été utile, puisqu'elle a orienté notre recherche vers d'autres problématiques majeures. En effet, un système de spatialisation sonore adapté à une visite de musée ne pourrait être développé sans une mise en place d'un modèle théorique permettant de représenter les différents acteurs de ce système. En outre, la mise en place d'un outil auteur permettant à la fois la conception et la simulation de scénarios de visite serait judicieuse.

C'est ainsi que nous avons étudié la problématique de modélisation de la visite augmentée. Dans cette partie, nous avons proposé un modèle composé de trois entités. La première entité est le paysage sonore présenté selon trois niveaux d'abstraction : physique, virtuelle et sémantique. Le principal apport introduit par notre modèle de paysage sonore est l'introduction de la notion avancée de la zone d'audibilité. Celle-ci, qui représente un espace 3D dans lequel un contenu audio est perceptible, permet de structurer l'espace du musée en un ensemble de zones qui peuvent être croisées, séparées ou encore hiérarchiques. Ces zones peuvent également être partagées ou exclusives. Des parcours de visite libres ou guidés peuvent être créés sur la base de ces zones d'audibilité. Ainsi le visiteur peut être immergé dans un paysage sonore composé de sons ambiants, de musiques et de commentaires en adéquation avec sa position et son orientation. Par ailleurs, ces zones peuvent être annotées et classifiées par concepts sous forme d'un thésaurus dans la couche sémantique du modèle de la scène sonore. Elles peuvent ensuite être facilement exploitées pour déterminer les intérêts des visiteurs.

La deuxième entité du modèle est le visiteur. Nous l'avons modélisé selon plusieurs niveaux. Le premier niveau est le profil initial du visiteur comportant par exemple son âge et ses centres d'intérêt. Le niveau suivant est celui de la captation spatiotemporelle de son mouvement durant sa visite. Ce niveau permet de décrire le niveau au-dessus qui est celui de la position du visiteur par rapport à la scène sonore mais aussi de la captation des gestes qu'il accomplit avec sa tête en guise d'interaction avec l'environnement sonore. À partir de ces niveaux, un quatrième niveau est introduit qui est celui de la signature des intérêts, celle-ci étant une nouvelle notion introduite pour définir les scores d'intérêt du visiteur aux différents descripteurs sémantiques associés aux zones d'audibilité de la scène sonore. Un historique de tous les descripteurs du visiteur à ces différents niveaux est maintenu afin de réaliser une adaptation en temps réel du parcours de la visite aux intérêts des visiteurs, mais également afin de permettre une étude plus détaillée des intérêts des visiteurs sur l'espace de plusieurs visites successives.

La troisième entité du modèle est le moteur de navigation qui constitue le cœur de SARIM. Ce moteur décrit la politique d'interaction du visiteur avec la scène sonore. Ce moteur est organisé selon trois couches concentriques. Chaque couche fournit un niveau d'interactivité particulier, et chacune des deux couches supérieures du moteur fournit les contenus d'interactivité de la couche qu'elle contient. La couche de la *Traque* est le noyau du moteur, Celle-ci réalise une spatialisation sonore simple des sons émis par les œuvres suivant la position et l'orientation du visiteur. La couche du *contrôle* permet au visiteur de contrôler volontairement le paysage sonore à travers différents gestes de têtes prédéfinis et activés par le concepteur du scénario de la visite. La dernière couche est l'*adaptation* où le système affine automatiquement le parcours du visiteur en cours de sa visite pour l'adapter au mieux à ses intérêts inférés en étudiant ses choix en termes d'œuvres regardées et de sons écoutés.

Après avoir proposé un modèle de visite augmentée, nous avons réalisé une évaluation empirique contrôlée du système SARIM et de certaines de ses fonctionnalités. Deux expériences ont été menées dans deux salles du musée des arts et métiers. La première expérience vise à étudier l'acceptabilité et l'apport du système SARIM en le comparant avec l'audioguide classique utilisé dans le musée. La deuxième expérience consiste à évaluer l'utilisabilité de SARIM pour sonoriser des œuvres placées les unes à côté des autres dans un espace restreint. Malgré le nombre restreint des participants, les résultats de l'évaluation montrent que SARIM est utilisable et offre un meilleur apport à la fois informatif et ludique à la visite du musée. Il permet aussi de faciliter la localisation des œuvres et accroît l'intérêt que porte le visiteur sur les œuvres visitées. Les résultats de la deuxième expérience montrent la gêne que peut apporter l'activation de plusieurs sources sonores (très rapprochés dans l'espace) lors d'une visite de musée. Elle montre ainsi le besoin de développer une configuration de scène sonore basée sur la notion de clusters pour regrouper les œuvres qui sont très proches géographiquement et similaires fonctionnellement.

6.2 Perspectives

Le travail de la thèse étant à la croisée de plusieurs domaines, elle ouvre le chemin vers différentes questions de recherche intéressantes. Durant notre thèse nous avons entamé la conception de quelques solutions qui répondent à des questions associées à notre sujet de thèse. Nous les évoquons ici dans ce chapitre.

6.2.1 Esquisse d'un logiciel auteur pour SARIM

Nous entendons par logiciel auteur, l'interface graphique permettant au concepteur de composer des scènes sonores virtuelles et de préciser l'interaction entre le visiteur et le monde réel. Notre outil de création est une interface graphique d'utilisateur conçue spécialement pour la réalité augmentée audio appliquée dans le contexte de musées.

La conception d'outils de création est plus fréquente dans le contexte des jeux et de la réalité virtuelle, mais elle est moins fréquente dans le domaine de la réalité augmentée, et encore moins dans le contexte des visites de musées.

Quelques logiciels qui permettent de créer des scènes et des paysages sonores ont été proposés au cours de ces dernières années. Le *Spatialisateur* [Jot99] développé par l'Ircam¹ est un Framework logiciel dédié à la spatialisation des créations musicales en temps réel, post-production et dans le contexte de concerts. Il permet de spécifier les paramètres de spatialisation indépendamment du mode de restitution électro-acoustique, mais aussi d'identifier intuitivement les paramètres d'effets de salle. Toutefois, ce logiciel est conçu pour les situations générales et présente peu de capacités pour la visualisation et l'expérimentation de la scène créée. Il a donc besoin d'être étendu par des modules spécialement conçus pour l'application finale et offrant plus de facilités pour la conception des scènes sonores.

Sur la base de cette observation, *ListenSpace* [Leprado07, Warusfel04, GoBmann04] a été développé comme une couche haut-niveau de Spat. Il a été conçu comme un outil de création spécialement adapté à la réalisation du projet *Listen* [Zimmermann08]. *ListenSpace* permet de décrire les propriétés géométriques de tous les objets qui composent une scène sonore dans *Listen*, et de combiner les paramètres acoustiques physiques et perceptifs pour définir les sources sonores dans la scène. Cependant, la description de la scène dans *ListenSpace* est limitée à une représentation bidimensionnelle (2D). Cette limitation prive le concepteur d'informations cruciales, et en particulier la hauteur (altitude) des sources sonores.

6.2.1.1 Concepts de base

La conception d'un outil auteur pour la création de paysages augmentés par le son et de scénarios de visite de musée, se compose de plusieurs étapes telles que la création des contenus sonores, la création des rendus visuels et la création de scénarios de visite du musée. Nous avons décidé de concentrer nos efforts sur la conception et le développement de fonctionnalités de base nécessaires à la conception des paysages sonores et des scénarios de visite puisque pour les deux autres tâches, des logiciels plus performants existent déjà. Par conséquent, l'outil actuel n'intègre pas d'outil de conception ou de gestion des contenus sonores, ni de fonctionnalités avancées pour la création des rendus graphiques. Cependant, le concepteur a la possibilité d'utiliser des formes géométriques simples pour créer les pièces du musée et les objets sonores également. En outre, l'outil intègre la possibilité d'importer des paysages graphiques 3D créés à travers des outils logiciels de modélisation 3D (ex. Blender) et permettant d'exporter les environnements 3D en langage standards de description d'univers graphiques virtuels (ex. VRML et Collada).

1. www.ircam.fr

Bien que l'outil d'auteur soit principalement conçu pour la création du paysage sonore, cet outil serait incomplet sans une fonctionnalité permettant au concepteur de tester virtuellement le résultat de ses scénarios. Ainsi, deux modes d'utilisation sont disponibles : le mode d'édition qui est utilisé pour créer et manipuler les objets graphiques et sonores, tandis que le mode de démonstration est utilisé pour vérifier et démontrer virtuellement le résultat du paysage sonore.

6.2.1.2 Modes utilisateurs

Mode d'édition Nous décrivons ici le fonctionnement de l'interface homme machine en mode édition. Ce mode est accessible à travers la fenêtre principale organisée de manière à offrir le maximum d'espace de travail au concepteur du paysage graphique et sonore. L'espace du travail est présenté au centre de la fenêtre. Des menus de paramétrage sont disposés dans des barres d'outils latérales et inférieures pouvant être cachées et affichées en un seul clic.

Mode de démonstration Pour tester le paysage sonore qui a été créé, le logiciel offre la possibilité de "jouer" ou de "simuler" une visite grâce à son interface mode de démonstration. Le mode de démonstration est accessible à l'aide du bouton "Play" en bas à droite de l'interface d'édition. Le mode de démonstration offre deux vues, la première est une vue générale dans laquelle la caméra (le point de vue) est située au dessus du visiteur. Ce dernier est représenté par un avatar qui reflète sa position et son orientation. L'utilisateur du logiciel, muni d'un casque stéréo, et d'un dispositif de navigation (souris, joystick) peut simuler une visite dans l'espace du musée et ainsi percevoir les contenus audio que le visiteur devra entendre selon sa position et son orientation.

Outre la vue générale, le mode de démonstration offre également une vue subjective qu'on appellera la vue du visiteur. Cette vue est accessible en activant le bouton de la vue subjective en bas à gauche. Elle permet au simulateur d'apercevoir ce que voit et entend le visiteur durant sa visite au musée. C'est la vue la plus idéale pour se mettre à la place du visiteur et faire concorder les deux rendus sonore et visuel.

6.2.1.3 Naviguer dans la scène sonore

La navigation dans l'environnement virtuel en 3D nécessite des appareils avancés offrant plusieurs degrés de liberté (DDL). La souris classique n'est pas en mesure de détecter les mouvements de rotation et est exclusivement sensible aux mouvements de translation (2 DDL).

Habituellement, les logiciels de CAO et les jeux vidéo permettent la navigation à travers les environnements virtuels en utilisant soit la souris classique combinée au clavier soit en utilisant les boutons et la molette de défilement de la souris [Almeida09]. Pour

des applications professionnelles, l'interaction bi-manuelle est favorisée en complétant, par exemple, la souris par un trackball. D'autres dispositifs fournissent également 6 degrés de liberté comme le SpaceNavigator de la compagnie *3DConnection* [O'Brien08]. Cependant ce dispositif est fixe et ne tourne pas sur 360 degrés. Cela crée un fossé entre la navigation dans le monde réel est la rotation par la souris. Pour parcourir des environnements destinés aux tests d'applications de réalité augmentée, ceci peut s'avérer contraignant. Ce type de dispositif 3D est plus approprié pour commander la vitesse de la rotation plutôt que la position et l'orientation d'un objet en scène [Almeida09].

Dans le contexte de la navigation au sein de paysages sonores virtuels, un dispositif de capture simultanée de la position et de l'orientation de l'utilisateur serait utile. Dans [Almeida07] un tel dispositif a été évalué pour d'autres applications telles que la sélection de la couleur dans une palette graphique et le mixage de sons avec des widgets rotatifs [Jacob94]. Ces expérimentations étaient faites avec la 4D Mouse. La 4D Mouse fait partie du kit de la tablette Wacom Intuos 2. Bien que stable et précis (0,2 degré), ce dispositif est plutôt sous-exploité par le monde des interfaces graphiques. La figure 6.1 montre un exemple de trajectoire intégrale réalisée en utilisant la 4D mouse.

La navigation dans la scène 3D en utilisant la vitesse peut se faire à travers les analogies suivantes :

- La distance (d) séparant la tête de la souris à la ligne rouge horizontale : la vitesse (signée) de déplacement du visiteur
- L'angle entre l'orientation de la souris et l'orientation de la ligne verticale : la vitesse angulaire (signée) de l'orientation du visiteur



FIGURE 6.1 – La trajectoire intégrale utilisant la souris 4D.

6.2.2 Protocole de description de scènes sonores

Étant donné la généralisation de l'utilisation des Smartphones et des dispositifs intelligents permettant la traque et l'accompagnement des utilisateurs en temps réel, pourra-t-on imaginer un langage ou un protocole unifié de description et d'interaction

avec les scènes sonores qui pourra assurer l'interopérabilité entre les différents systèmes de RAS appliquée à la visite de musée.

Dans ce cadre, TourML a été introduit par Robert Stein, et Nancy Proctor [Stein11] comme étant un langage standard pour la définition des parcours mobiles des musées. C'est un schéma XML qui décrit la structure, le contenu et l'expérience prévue d'un parcours mobile de musée en identifiant ses éléments constitutifs. TourML décrit les éléments qui composent un parcours en utilisant trois identifiants essentiels :

1. L'"Asset" qui décrit un composant élémentaire compris dans le parcours, peut être défini donc par un "Asset" de type vidéo, image, web, etc. Une instance TourML doit au minimum présenter l'ensemble d'"Asset" contenus dans ce parcours.
2. Le "Stop" qui constitue une étape du parcours et qui se compose d'un groupe d'"Asset" perceptibles au mêmes temps. Il peut être par exemple un diaporama de photos joué sur un arrière plan audio.
3. La "Connection" qui définit le sens du parcours et qui relie un "Stop" au "Stop" qui le suit dans le parcours

Ceci permet au final de créer un graphe orienté du tour définissant comment le visiteur doit se déplacer entre les "Stop".

Ce standard de métadonnées fournit un certain nombre d'avantages pour les musées. En effet, il permet de faire migrer le contenu et la structure d'un parcours parmi les divers outils mobiles disponibles et futurs, sans nécessité de réadaptation manuelle par le personnel du musée. Il permet également aux musées de choisir les outils auteur qui répondent le mieux à leurs besoins tout en continuant à utiliser les appareils mobiles qui s'adaptent le mieux à leur contenu, et finalement il contribue à la survie du contenu des parcours des musée au avancés technologique de présentation du contenu multimédia.

TourML peut être utilisé, après une personnalisation importante, pour représenter notre modèle de la scène sonore. Ceci nécessite bien le développement de plusieurs `assetType`, `stopType` et `propertyType` associés. Nous pouvons imaginer dans notre cas la définition d'asset "zone d'audibilité" ceci nécessite la définition du `assetType` "ZoneDAudibiliteAsset", ainsi que plusieurs `propertyType` qui lui sont liées.

Mais pour définir le modèle de navigation et le modèle visiteur, une extension du standard s'impose. En effet un élément "visitor" pourra être rajouté pour décrire le ou les visiteurs évoluant dans le graphe du parcours. Deux éléments "Action" et "State" pourront être introduits pour définir une politique d'interaction. Un "State" représente un état d'un "Visitor" par rapport à un "Asset". Une "Action" représente un acte fait par ce visiteur, soit un geste soit un arrêt prolongé. Cette "Action" peut être définie en utilisant des éléments de temps, de position et d'orientation. Nous pouvons définir également un élément "interaction" qui se compose de deux "State", et une "Action" qui définit l'acte permettant le passage d'un "State" à un autre.

Par contre, le contenu audio qui nous intéresse dans cette thèse, ne peut être représenté dans TourML qu'entant que simple contenu média, sans prendre en considération la fonctionnalité de la spatialisation sonore. En outre, le standard ne permet pas de spécifier :

- à quel objet exposé un son est associé ?
- où devrait être joué le son géographiquement, pourrait-il y avoir des superpositions entre les "Stop" ?
- comment va être joué le son : en boucle ou une seule fois, émanant d'une source sonore ou réparti régulièrement dans une ZA ?

Ainsi d'autres travaux d'extension de TourML doivent être faits pour une représentation complète de notre modèle de scène sonore.

6.2.3 Autres directions de recherche

Cette thèse se trouvant au carrefour de différents domaines, elle nous ouvre plusieurs directions de recherche dont l'étude s'avère nécessaire pour une meilleure utilisation de la spatialisation sonore comme moyen d'autoguidage.

L'un des sujets importants à étudier serait une analyse approfondie du comportement naturel des visiteurs afin de déduire leurs intérêts et préférences. Outre l'utilisation du profil du visiteur, l'idée la plus basique appliquée dans ce travail est de se référer au temps passé en regardant un objet et en écoutant le son qui lui est associé pour estimer l'intérêt porté à ce dernier. La question est certainement beaucoup plus complexe. L'intérêt peut varier dans le temps. Il peut dépendre de l'expérience du visiteur mais aussi des événements avant et après la visite. Le profil tout seul ne peut pas refléter l'intérêt puisque ce même profil est variable. Le long regard porté vers un objet pourra même dans certaines conditions être interprété comme un désintérêt, une contestation ou un refus. Comment pourra-t-on déduire si l'intérêt concerne la totalité de l'objet ou qu'une seule composante de ce dernier : sa forme, sa couleur, sa fonction, le son associé. Concernent la gestuelle ou le "body language", il convient de noter que les codes ne sont pas similaires pour tout type de public, nous pouvons penser aux indiens par exemple qui balancent leurs tête de droite à gauche pour dire oui, contrairement aux européens par exemple. Le public malvoyant nécessite également une attention particulière pour étudier son propre langage du corps.

Une autre direction de recherche serait l'étude de la traque continue de la position et de l'orientation du visiteur sur un plan sociétal. La question de la vie privée a suscité et suscite encore beaucoup d'inquiétude vis à vis des nouvelles technologies qui imposent une traque et un suivi continu de l'activité de l'utilisateur. L'objectif étant d'enrichir la visite, comment évaluer le besoin réel de l'utilisateur et le niveau d'implication du système dans son comportement ? Serait-il plus judicieux de lui laisser un espace de liberté en n'interprétant que certains de ses comportements et pas tous ? Quel serait le seuil au niveau du temps de parcours, du nombre d'objets sonorisé et en termes de contenu

sonore à partir duquel l'augmentation sonore gênera la visite ? Dépassant ce seuil cette application de la RAS dans la visite au musée mènera-t-elle vers une surcharge ou une diminution de la réalité ? Empêchera-t-elle le visiteur d'apercevoir quelques détails liés aux objets exposés ?

Par ailleurs, étant donné la généralisation de l'utilisation des Smartphones et des dispositifs intelligents permettant la traque des utilisateurs en temps réel, jusqu'où peut-on étendre le modèle SARIM pour qu'il se généralise à d'autres situations de la vie quotidienne ? Pourra-t-on imaginer dans le futur un système basé sur la RAS qui nous assiste dans nos tâches quotidiennes ? La question se pose également pour le public malvoyant pour lequel une assistance généralisée pourra être intéressante. Pourra-t-on imaginer un langage ou un protocole unifié de description et d'interaction avec les scènes sonores qui pourront assurer l'interopérabilité entre les différents systèmes de RAS appliquée à la vie quotidienne. Une analogie peut être faite avec le système GSM qui fait passer le téléphone portable d'un relais à l'autre sans interruption et sans nécessité d'adaptation. Quelles sont les défis techniques posés pour une telle approche (architecture client/serveur, p2p,...)

La question de la mise en échelle du système et de la granularité des détails de l'environnement que l'on souhaite sonoriser se pose également à ce stade. La vitesse du déplacement mais aussi le domaine d'application joueront nécessairement dans la définition de cette dernière.

La question des canaux d'interaction est également intéressante à étudier. Étant donné que l'objectif est de réduire au maximum l'attention requise pour interagir avec l'environnement sonore, comment pourra-t-on enrichir le langage utilisé pour communiquer avec cet environnement ? Devra/pourra-t-on introduire de nouveaux moyens pour interpréter les gestes corporels et surtout manuels ou vocaux des utilisateurs. Nous penserons éventuellement à des systèmes de détection de mouvements à travers l'analyse du son, de l'image, etc. Le modèle pourra être également adapté pour qu'il soit appliqué à une visite d'un groupe de visiteurs. Comment interpréter le comportement du groupe et guider le groupe comme une visite d'un musée dans le cadre d'un voyage organisé.

Enfin, plutôt que les visiteurs soient seulement récepteurs du contenu sonore, la visite d'un musée pourrait même devenir participative. Des contenus de la scène sonore pourraient être créés ou commentés en temps réels par les visiteurs eux même tel le Web2.0. D'autres visiteurs pourraient ainsi réagir et interagir avec ces contenus et éventuellement en proposer d'autres. Ceci nécessiterait bien évidemment un procédé de modération et de sélection des contenus et commentaires pertinents.

A

Annexe

Nous présentons dans ce chapitre, dans un premier temps, le calcul effectué pour récupérer les valeurs des angles d'Euler données par le *SparkfunIMU 6DOF V4*. Dans un second temps, nous présentons le matériel utilisé lors des expériences menées dans le musée des arts et métiers de Paris.

A.1 Calcul de l'orientation avec Sparkfun IMU-6DOF-V4

Nous avons utilisé ce capteur dans le but d'obtenir des mesures précises des angles de l'orientation de la tête du visiteur, surtout pour le *Lacet* aussi appelé *Heading*, le *Tangage* ou *Elevation* et le *Roulis* ou *Roll*. Le *Lacet* est défini comme l'angle dans le plan horizontal local mesuré dans le sens des aiguilles de la montre à partir de la direction du vrai Nord (axe polaire de la Terre). Le *Tangage* est l'angle formé entre l'axe X et l'horizon/sol, et le *Roulis* est l'angle formé entre l'axe Y et l'horizon/sol (Figure A.1).

A.1.0.1 Le calcul du tangage et du roulis

Nous pouvons obtenir des mesures assez fiables du *Tangage* et du *Roulis* en utilisant les données fournies par l'accéléromètre embarqué dans le *IMU-6DOF-V4*. En pratique, au repos, le capteur renvoie les coordonnées du vecteur force de gravité dans un repère local au circuit. Celui-ci ne correspond pas à la verticale quand le capteur est posé à plat. Il faut donc effectuer une calibration. Pour cette fin, la procédure que nous avons suivie est celle décrite pour le calibrage de la *Wiimote* [Kionix07]. Elle consiste à poser

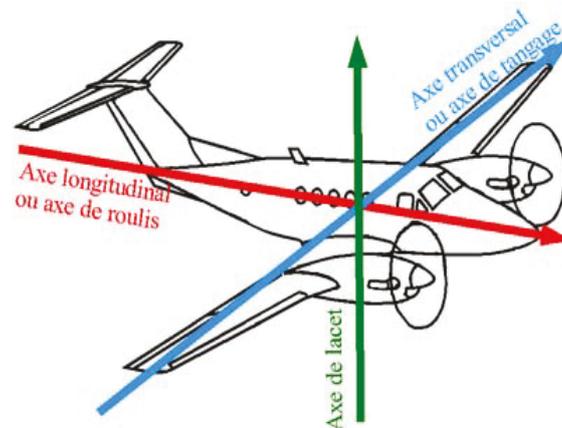


FIGURE A.1 – Dessin des axes de roulis, tangage et lacet sur un avion.

le capteur de manière à avoir successivement les trois axes X , Y et Z à la verticale et à collecter les valeurs renvoyées : x_1, y_1, z_1 lorsque l'axe X est vertical, x_2, y_2, z_2 lorsque l'axe Y est vertical et x_3, y_3, z_3 lorsque l'axe Z est vertical.

On calcule alors les coordonnées du point origine :

$$x_0 = (x_1 + x_2) / 2 \quad y_0 = (y_1 + y_2) / 2 \\ z_0 = (z_2 + z_3) / 2$$

On obtient ainsi les coordonnées du vecteur force (exprimées en g) :

$$a_x = (x_{raw} - x_0) / (x_3 - x_0) \\ a_y = (y_{raw} - y_0) / (y_2 - y_0) \\ a_z = (z_{raw} - z_0) / (z_1 - z_0)$$

où $x_{raw}, y_{raw}, z_{raw}$ sont les mesures de l'accéléromètre.

Si l'accéléromètre n'est pas en mouvement accéléré, la mesure des coordonnées du vecteur gravité permet d'obtenir l'orientation dans l'espace du dispositif, les valeurs du *Tangage* et du *Roulis* sont donc calculées à partir des formules :

$$\text{Tangage} = \arctan(a_x / \sqrt{a_y^2 + a_z^2}) \\ \text{Roulis} = \arctan(a_y / \sqrt{a_x^2 + a_z^2})$$

A.1.0.2 Le calcul du lacet

Une approche naïve pour calculer le *Lacet* consiste à une intégration de la vitesse angulaire donnée par le gyroscope. Il peut également être calculé à partir des données du magnétomètre. Toutefois, ces deux méthodes ne donnent pas des résultats exacts. En effet, les capteurs magnétiques sont soumis à des perturbations imprévi-

sibles. Par ailleurs, les capteurs gyroscopiques sont eux aussi sujets d'une dérive (*Drift*) de leurs mesures d'origines (*Bias*) qui augmente de manière non régulière au fil du temps. Afin d'éviter la dérive du gyroscope et les perturbations magnétiques, notre approche consiste à coupler les données du magnétomètre, de l'accéléromètre et du gyroscope embarquées dans l'IMU afin d'avoir le suivi le plus optimal de l'orientation. Notre approche de couplage consiste à calculer le *Lacet* fourni par le magnétomètre et le *Lacet* fourni par le gyroscope, puis à les combiner selon les algorithmes décrits dans ce qui suit.

Lacet du magnétomètre :

Nous commençons par calculer le *Lacet* en utilisant uniquement les données du magnétomètre [Caruso00]. La première étape consiste à calibrer la boussole. Pour cela, nous calculons les valeurs minimales et maximales des mesures du magnétomètre selon les 3 axes (x_{min} , y_{min} , z_{min} , x_{max} , y_{max} , z_{max}). L'objectif est d'inclure les facteurs de corrections ferreux. La proximité des matériaux ferreux peut créer des distorsions, les nouvelles valeurs de la boussole sont :

$$\begin{aligned} y_c &= 2 \quad (y_i - y_{min}) / (y_{max} - y_{min}) - 1 \\ x_c &= 2 \quad (x_i - x_{min}) / (x_{max} - x_{min}) - 1 \\ z_c &= 2 \quad (z_i - z_{min}) / (z_{max} - z_{min}) - 1 \end{aligned}$$

Etant donné que la sensibilité du magnétomètre diminue lorsque le *Tangage* et le *Roulis* augmentent, nous les incluons dans le calcul du *Lacet* :

$$\begin{aligned} X_h &= x_c \cos(\text{Tangage}) \\ &\quad + y_c \sin(\text{Roulis}) \sin(\text{Tangage}) \\ &\quad + z_c \cos(\text{Roulis}) \sin(\text{Tangage}); \\ Y_h &= y_c \cos(\text{Roulis}) + z_c \sin(\text{Roulis}); \end{aligned}$$

Ainsi, la formule pour calculer le *Lacet* à partir du magnétomètre et de l'accéléromètre est :

$$Lacet_m = \arctan(Y_h / X_h)$$

Lacet du gyroscope :

Les données fournies par le gyroscope correspondent aux vitesses auxquelles les trois angles (*Lacet*, *Tangage* et *Roulis*) varient. Elles ne nous permettent pas ainsi directement de mesurer l'angle du *Lacet*. Nous ne pourrions que calculer les variations d'angles sur des périodes successives puis calculer le *Lacet* en se basant sur la formule suivante :

$$Lacet_g(i) = Lacet_g(i-1) + \Delta Lacet_g$$

Théoriquement, la vitesse angulaire est la dérivée de l'angle par rapport au temps. Soit $\omega_{Lacet}(r)$ la mesure correspondante à la valeur réelle de la vitesse angulaire, et $\omega_{Lacet}(m)$ correspond à la valeur mesurée de la vitesse angulaire par le gyroscope. $Lacet_g$ est la valeur du lacet mesurée par le gyroscope.

$$\omega_{Lacet}(r) = dLacet_g / dt$$

$$dLacet_g = \omega_{Lacet}(r) dt$$

pour une fonction discrète, cette formule s'écrira $\omega_{Lacet}(r) = \Delta Lacet_g / \Delta t$

$$\Delta Lacet_g = \omega_{Lacet}(r) \Delta t$$

Les mesures récupérées par le gyroscope ne sont pas prises dans une échelle réelle, mais bien multipliées par un facteur d'échelle correspondant aux spécifications du constructeur. On doit multiplier le résultat par ce *facteur d'échelle* f_e pour pouvoir rendre les mesures en degré.

$$\omega_{Lacet}(r) = \omega_{Lacet}(m) f_e$$

Ce facteur peut être calculé expérimentalement, dans notre cas il correspond à la valeur $f_e = 4/3$. En plus de ce *facteur d'échelle* qui reste fixe dans le temps, Les vitesses mesurées par le gyroscope sont retournées avec un *Biais*. Ceci signifie que :

$$f_e = 4/3$$

Les vitesses mesurées par le gyroscope sont retournées avec un *Biais*. Soit $\omega_{Lacet}(mb)$ correspond à la mesure brute de la vitesse angulaire retournée par le gyroscope. Ceci signifie que :

$$\omega_{Lacet}(m) = \omega_{Lacet}(mb) + Biais$$

$$\Delta Lacet_g = (\omega_{Lacet}(mb) + Biais) f_e \Delta t$$

La valeur du *Biais* est égale à la mesure du gyroscope lorsque celui-ci est immobile ($\omega_{Lacet}(r) = 0$). Cependant, les gyroscopes souffrent d'une dérive. Plus précisément, le *Biais* change dans le temps indépendamment des mesures. Cela rend la vitesse angulaire inexacte après quelques secondes. Afin de remédier à cet inconvénient, nous proposons d'utiliser la variance de la vitesse angulaire du *Lacet* en vue de déterminer si le capteur est en mouvement ou non. Si sa valeur reste proche du zéro pendant une période $T_0 = 0,5$ secondes (période nécessaire pour faire 100 mesures de gyroscope, sachant que la fréquence des mesures du capteur est de 195 Hz), nous supposons qu'il n'y a pas de mouvement, et nous mettons à jour le *Biais* avec la mesure du gyroscope. En fait, la variance de la vitesse angulaire du *Lacet* peut s'approcher du zéro également en cas d'un mouvement avec une vitesse constante, mais cette situation se produit rarement puisque les mouvements effectués pour tourner la tête sont en général courts et avec une vitesse souvent variable.

if($\sigma^2(\omega_{Lacet}(m)) < \epsilon$) pendant une durée spécifiée

Biais = $E(\omega_{Lacet}(m))$ avec E : l'espérance mesurée sur la période T

Les algorithmes que nous avons cités ci-dessus permettent de réduire l'effet des perturbations magnétiques, la dérive du gyroscope, et d'obtenir une estimation fiable de l'angle du *Lacet*. Nous avons cité ces algorithmes comme étant un premier pas vers une API "open source" qui traite les fonctionnalités de captation de la position et de l'orientation en utilisant des capteurs IMU.

A.1.1 Couplage du magnétomètre et du gyroscope

Après avoir calculé séparément les deux lacets (du magnétomètre $Lacet_m$ et du gyroscope $Lacet_g$), nous combinons les deux valeurs pour obtenir une meilleure mesure du $Lacet_c$ et surmonter les problèmes de chaque capteur [Roetenberg05]. Nous supposons qu'il existe des perturbations magnétiques lorsqu'il y a une différence importante entre la variance de la vitesse angulaire obtenue par le gyroscope et la variance de la vitesse angulaire obtenue par le magnétomètre. Par conséquent, quand il n'y a pas de perturbation nous mettons $Lacet_c = Lacet_m$, si une perturbation est détectée, la différence entre $Lacet_g$ et le dernier $Lacet_c$ non perturbé est : $Lacet_c = Lacet_c + \Delta(Lacet_g)$ où $\Delta(Lacet_g)$ est la différence entre le $Lacet_g$ et le dernier $Lacet_g$ non perturbé.

A.2 Consignes

Expérience dans la salle des communications

« Vous allez visiter cette salle comme vous le feriez en situation réelle. Nous vous demandons simplement de vous concentrer uniquement sur les objets présentant une description audio, indiqués par une vignette de couleur. Allez à votre rythme ; ce n'est pas un jeu de piste, mais une situation de visite "normale" du musée. Au bout de quelques minutes, nous interrompons la visite pour vous poser des questions sur votre expérience. Je vous suivrai tout au long de la visite pour en observer le déroulement, mais je n'aurai pas le droit d'interagir avec vous. Rappelez vous seulement que ce n'est pas vous que nous évaluons, mais le dispositif que vous utiliserez. »

Expérience dans la salle des automates

Condition 1 : « En observant cette salle, vous allez entendre des musiques provenant de certains des automates exposés. Si vous repérez un objet qui vous intéresse, il vous suffira d'orienter votre tête dans sa direction et de le fixer quelques secondes pour écouter le sons qu'il émet. Après avoir écouté le morceau de musique d'un automate, vous pourrez écouter les commentaires audio qui lui sont relatifs si vous continuez de regarder dans sa direction. Cependant, si vous en avez envie, vous pouvez à tout moment passer à l'observation d'un autre objet en tournant la tête vers celui-ci. Nous vous laissons libre d'explorer la scène comme vous le voudrez. Au bout de quelques minutes nous arrêterons l'expérience pour vous poser des questions sur ce que vous viendrez de vivre. Rappelez vous seulement que ce n'est pas vous que nous évaluons, mais le dispositif que vous utiliserez. »

Condition 2 : « En observant cette salle, vous allez entendre des musiques provenant de certains des automates exposés. Si vous repérez un objet qui vous intéresse, il vous suffira d'orienter votre tête dans sa direction pour écouter le sons qu'il émet. Après avoir écouté le morceau de musique d'un automate, vous pourrez écouter les commentaires audio qui lui sont relatifs si vous continuez de regarder dans sa direction. Cepen-

de recrutement.png

Participez à une expérience de réalité augmentée sonore dans le cadre d'une visite du Musée des Arts et Métiers

Ce formulaire va nous permettre de prendre vos coordonnées pour vous proposer un rendez-vous en fonction de vos disponibilités (que pouvez nous transmettre via le doodle envoyé par mail) :

Rappel :
Pour participer, vous devez impérativement ne jamais avoir utilisé de dispositif d'audioguide tel qu'on en trouve au musée.

Merci de votre participation !

NOM Prénom *

Âge *

Mail *

Téléphone

Avez vous déjà visité le Musée des Arts et Métiers ? *

Oui

Non

Environ combien de fois avez vous visité le musée au cours des 3 dernières années ? *

Aucune

une fois

2 ou 3 fois

4 ou 5 fois

Plus de 5 fois

Merci !

FIGURE A.2 – Formulaire de recrutement.

dant, si vous en avez envie, vous pouvez à tout moment passer à l'observation d'un autre objet en tournant la tête vers celui-ci. Nous vous laissons libre d'explorer la scène comme vous le voudrez. Au bout de quelques minutes nous arrêterons l'expérience pour vous poser des questions sur ce que vous viendrez de vivre. Rappelez vous seulement que ce n'est pas vous que nous évaluons, mais le dispositif que vous utiliserez. »

A.3 Formulaire de recrutement

Afin de recruter les sujets, nous avons mis en place un formulaire (figure A.2) auquel doit répondre les sujets qui souhaitent participer à cette expérimentation. A la base de leurs réponses, nous avons sélectionné ceux qui vont participer à cette expérience.

A.4 Questionnaires

A.4.1 Questionnaire : salle des communications

Condition1 : SARIM

Le questionnaire proposé aux participants à la condition de SARIM dans la salle des communications est représenté dans les figures A.3, A.4, A.5, A.6.

Condition2 : audioguide

Pour la condition de l'audioguide, les participants à l'expérimentation dans la salle des communications, ont répondu au questionnaire suivant (figures A.7, A.8, A.9, A.10)

A.4.2 Questionnaire : salle des automates

Dans la salle des automates, les sujets qui ont participé à cette expérimentation dans les deux conditions (zones croisées ou indépendantes), ont répondu au questionnaire suivant (A.11, A.12, A.13)

14/09/2012

Questionnaire - Salle des com

Questionnaire - Salle des com

Sujet :

Prototype

Êtes-vous habituellement intéressé(e) par ce genre d'exposition ?

- Pas du tout intéressé(e)
- Peu intéressé(e)
- Plutôt intéressé(e)
- Fortement intéressé(e)

Dans l'ensemble, la visite vous a semblé...

- Ennuyeuse
- Moyennement intéressante
- Intéressante
- Passionnante

D'après vous, combien de temps a-t-elle duré ?

La visite vous a-t-elle paru amusante ?

- Pas du tout amusante
- Pas très amusante
- Plutôt amusante
- Très amusante

Dans l'ensemble, les objets que vous avez observés vous ont semblé...

- Pas du tout intéressants
- Peu intéressants
- Plutôt intéressants
- Très intéressants

D'après vous, quel a été l'impact de la presse rotative située au centre de la salle dans l'histoire de l'imprimerie ?

- Très faible
- Plutôt faible
- Plutôt fort

docs.google.com/spreadsheet/viewfor...

1/4

FIGURE A.3 – Questionnaire de la salle des communications-condition SARIM (page1).

14/09/2012

Questionnaire - Salle des com

 Très fort

Vous êtes vous senti(e) projeté(e) dans les époques correspondant aux objets que vous avez observés ?

- Pas du tout
- Partiellement
- Tout à fait

Vous êtes-vous senti(e) à l'aise avec l'utilisation du dispositif ?

- Très peu à l'aise
- Plutôt peu à l'aise
- Plutôt à l'aise
- Très à l'aise

Comment avez-vous trouvé la réactivité du système ?

- Beaucoup trop lente
- Plutôt lente
- Suffisamment rapide

A quelle vitesse pensez-vous que la plupart des gens apprendrait à utiliser ce système ?

- Très lentement
- Plutôt lentement
- Plutôt rapidement
- Très rapidement

Personnellement, avez-vous trouvé le dispositif facile à utilisé ?

- Très difficile
- Plutôt difficile
- Plutôt facile
- Très facile

Avez-vous compris comment écouter les commentaires audio liées aux objets observés ?

- Pas du tout
- Je pense que oui
- Absolument

Si oui, comment ?

docs.google.com/spreadsheet/viewfor...

2/4

FIGURE A.4 – Questionnaire de la salle des communications-condition SARIM (page2).

14/09/2012

Questionnaire - Salle des com

Vous avez compris comment écouter le son d'un objet en particulier ?

- Pas du tout
- Je pense que oui
- Absolument

Si oui, comment ?**Vous est-il arrivé de déclencher des sons que vous ne vouliez pas écouter ?**

- Oui
- Non

Si oui, vous avez pu corriger vos actions..

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Ces erreurs de manipulation vous ont-elles aidé(e) à mieux comprendre le fonctionnement du dispositif ?

- Pas du tout aidé(e)
- Pas trop aidé(e)
- Plutôt aidé(e)
- Tout à fait aidé(e)

Dans l'ensemble, vous avez vous compris comment utiliser le dispositif..

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Vous avez réussi à localiser les objets à observer...

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Les sons du casque vous-ont-ils aidé(e) à trouver les objets à observer ?

- Pas du tout aidé(e)
- Pas tellement aidé(e)

docs.google.com/spreadsheet/viewfor...

3/4

FIGURE A.5 – Questionnaire de la salle des communications-condition SARIM (page3).

14/09/2012

Questionnaire - Salle des com

- Plutôt aidé(e)
- Tout à fait aidé(e)

Dans l'ensemble, Vous avez trouvé le dispositif...

- Très encombrant
- Plutôt encombrant
- Plutôt pratique
- Très pratique

Durant la visite, avez-vous ressenti le besoin de pouvoir contrôler le dispositif autrement que par vos déplacements ? Pourquoi et comment ?

D'une manière générale, êtes vous satisfait(e) du dispositif ?

- Très insatisfait(e)
- Plutôt insatisfait(e)
- Plutôt satisfait(e)
- Très satisfait(e)

Si un tel système se généralisait, aimeriez-vous pouvoir l'utiliser à nouveau lors de vos futures visites de musées ?

- Je n'aimerais pas du tout
- Pourquoi pas, ça m'est égal
- J'aimerais assez
- J'aimerais beaucoup

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

FIGURE A.6 – Questionnaire de la salle des communications-condition SARIM (page4).

14/09/2012

Questionnaire - Salle des com

Questionnaire - Salle des com

Sujet :

Audioguide

Êtes-vous habituellement intéressé(e) par ce genre d'exposition ?

- Pas du tout intéressé(e)
- Peu intéressé(e)
- Plutôt intéressé(e)
- Fortement intéressé(e)

Dans l'ensemble, la visite vous a semblé...

- Ennuyeuse
- Moyennement intéressante
- Intéressante
- Passionnante

D'après vous, combien de temps a-t-elle duré ?

La visite vous a-t-elle paru amusante ?

- Pas du tout amusante
- Pas très amusante
- Plutôt amusante
- Très amusante

Dans l'ensemble, les objets que vous avez observés vous ont semblé...

- Très intéressants
- Plutôt intéressants
- Peu intéressants
- Pas du tout intéressants

D'après vous, quel a été l'impact de la presse rotative située au centre de la salle dans l'histoire de l'imprimerie ?

- Très faible
- Plutôt faible
- Plutôt fort

docs.google.com/spreadsheet/viewfor...

1/4

FIGURE A.7 – Questionnaire de la salle des communications-condition de l'audioguide (page1).

14/09/2012 Questionnaire - Salle des com

Très fort

Vous êtes vous senti(e) projeté(e) dans les époques correspondant aux objets que vous avez observés ?

Pas du tout
 Partiellement
 Tout à fait

Vous êtes-vous senti(e) à l'aise avec l'utilisation du dispositif ?

Très peu à l'aise
 Plutôt peu à l'aise
 Plutôt à l'aise
 Très à l'aise

Comment avez-vous trouvé la réactivité du système ?

Beaucoup trop lente
 Plutôt lente
 Suffisamment rapide

A quelle vitesse pensez-vous que la plupart des gens apprennent à utiliser ce système ?

Très lentement
 Plutôt lentement
 Plutôt rapidement
 Très rapidement

Personnellement, avez-vous trouvé le dispositif facile à utilisé ?

Très difficile
 Plutôt difficile
 Plutôt facile
 Très facile

Avez-vous compris comment écouter les commentaires audio liées aux objets observés ?

Pas du tout
 Je pense que oui
 Absolument

Si oui, comment ?

FIGURE A.8 – Questionnaire de la salle des communications-condition de l'audioguide (page2).

14/09/2012

Questionnaire - Salle des com

Vous est-il arrivé de faire des erreurs de manipulation du dispositif durant la visite ?

- Oui
- Non

Si oui, vous avez pu corriger vos actions..

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Ces erreurs de manipulation vous ont-elles aidé(e) à mieux comprendre le fonctionnement du dispositif ?

- Pas du tout aidé(e)
- Pas trop aidé(e)
- Plutôt aidé(e)
- Tout à fait aidé(e)

Dans l'ensemble, vous avez vous compris comment utiliser le dispositif..

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Vous avez réussi à localiser les objets à observer...

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Dans l'ensemble, Vous avez trouvé le dispositif..

- Très encombrant
- Plutôt encombrant
- Plutôt pratique
- Très pratique

D'une manière générale, êtes vous satisfait(e) du dispositif ?

- Très insatisfait(e)
- Plutôt insatisfait(e)
- Plutôt satisfait(e)
- Très Très satisfait(e)

docs.google.com/spreadsheet/viewfor...

3/4

FIGURE A.9 – Questionnaire de la salle des communications-condition de l'audioguide (page3).

14/09/2012 Questionnaire - Salle des com

Aimeriez vous utiliser à nouveau ce dispositif lors de vos futures visites de musées ?

- Je n'aimerais pas du tout
- Pourquoi pas, ça m'est égal
- J'aimerais assez
- J'aimerais beaucoup

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

docs.google.com/spreadsheet/viewfor...

4/4

FIGURE A.10 – Questionnaire de la salle des communications-condition de l'audioguide (page4).

14/09/2012

Questionnaire - Salle des Automates

Questionnaire - Salle des Automates

Sujet :

Êtes-vous habituellement intéressé(e) par ce genre d'exposition ?

- Pas du tout intéressé(e)
- Peu intéressé(e)
- Plutôt intéressé(e)
- Fortement intéressé(e)

Dans l'ensemble, l'expérience vous a semblé(e)...

- Ennuyeuse
- Moyennement intéressante
- Intéressante
- Passionnante

D'après vous, combien de temps a-t-elle duré ?

L'expérience vous a-t-elle paru amusante ?

- Pas du tout amusante
- Pas très amusante
- Plutôt amusante
- Très amusante

Dans l'ensemble, les objets que vous avez observés vous ont semblé...

- Très intéressants
- Plutôt intéressants
- Peu intéressants
- Pas du tout intéressants

Vous êtes vous senti(e) projeté(e) dans les époques correspondant aux objets que vous avez observés ?

- Pas du tout
- Partiellement
- Tout à fait

Vous êtes-vous senti(e) à l'aise avec l'utilisation du dispositif ?

- Très peu à l'aise
- Plutôt peu à l'aise

docs.google.com/spreadsheet/viewfor...

1/4

FIGURE A.11 – Questionnaire de la salle des automates (page1).

14/09/2012

Questionnaire - Salle des Automates

- Plutôt à l'aise
- Très à l'aise

Pourquoi ?**Comment avez-vous trouvé la réactivité du système ?**

- Beaucoup trop lente
- Plutôt lente
- Suffisamment rapide

A quelle vitesse pensez-vous que la plupart des gens apprendrait à utiliser ce système ?

- Très lentement
- Plutôt lentement
- Plutôt rapidement
- Très rapidement

Personnellement, avez-vous trouvé le dispositif facile à utiliser ?

- Très difficile
- Plutôt difficile
- Plutôt facile
- Très facile

Vous est-il arrivé de déclencher des sons que vous ne vouliez pas écouter ?

- Oui
- Non

Si oui, vous avez pu corriger vos actions...

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Combien d'objets avez-vous localisés ?docs.google.com/spreadsheet/viewfor...

2/4

FIGURE A.12 – Questionnaire de la salle des automates (page2).

14/09/2012

Questionnaire - Salle des Automates

Vous avez réussi à localiser les objets à observer...

- Très difficilement
- Plutôt difficilement
- Plutôt facilement
- Très facilement

Dans l'ensemble, Vous avez trouvé le dispositif...

- Très encombrant
- Plutôt encombrant
- Plutôt pratique
- Très pratique

Durant l'expérience, avez-vous ressenti le besoin de pouvoir contrôler le dispositif autrement que par l'orientation de votre tête ? Pourquoi et comment ?

D'une manière générale, êtes vous satisfait(e) du dispositif ?

- Très insatisfait(e)
- Plutôt insatisfait(e)
- Plutôt satisfait(e)
- Très satisfait(e)

Si un tel système était disponible dans certains musées, aimeriez-vous pouvoir l'utiliser à nouveau lors de vos visites ?

- Je n'aimerais pas du tout
- Pourquoi pas, ça m'est égal
- J'aimerais assez
- J'aimerais beaucoup

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

docs.google.com/spreadsheet/viewfor...

3/4

FIGURE A.13 – Questionnaire de la salle des automates (page3).

Bibliographie

- [Abowd97] Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper & Mike Pinkerton. *Cyberguide : a mobile context-aware tour guide*. *Wirel. Netw.*, vol. 3, no. 5, pages 421–433, Springer-Verlag New York, Inc., Secaucus, NJ, USA, October 1997.
- [Alais04] David Alais & David Burr. *The Ventriloquist Effect Results from Near-Optimal Bimodal Integration*. *Current Biology*, vol. 14, no. 3, pages 257–262, Cell Press, feb 2004.
- [Albaum97] Gerald Albaum. *The Likert scale revisited : an alternate version.(product preference testing)*. *Journal of the Market Research Society*, vol. 39, no. 2, pages 331–343, apr 1997.
- [Almeida07] Rodrigo Almeida & Pierre Cubaud. *Nearly-Integral Manipulation of Rotary Widgets*. In *INTERACT*, pages 489–492, 2007.
- [Almeida09] Rodrigo Almeida. *Thèse de doctorat : Contribution aux techniques pour enrichir l'espace moteur et l'espace visuel des dispositifs d'interaction bureautique*. Préparée au : Centre d'étude et de recherche en Informatique du CNAM, Paris, France, 2009.
- [AMNH] AMNH. *AMNH Explorer from the American Museum of Natural History*. <http://www.amnh.org/apps/explorer.php>, consulté le 11/09/2013.
- [Anastassova07] M. Anastassova, J. M. Burkhardt & C. Mégard. *User-Centred Design and Evaluation of Augmented Reality Systems for Industrial Applications : Some Deadlocks and Breakthroughs*. *Virtual Reality International Conference (VRIC) 2007, April 18-20, Laval, France (Best Paper Award)*, *Virtual Reality International Conference (VRIC) 2007, April 18-20, Laval, France (Best Paper Award)*, 2007.
- [Apeloig08] Yves Apeloig. *La révolution des audioguides : Eshkar SARL Faldard Industrie*. 2008. http://www.eurosfaire.prd.fr/7pc/documents/1329753872_7_eshkarfrv10_2_0.pdf, consulté le 10/09/2013.
- [Azuma97] Ronald T. Azuma. *A survey of augmented reality*. *Presence : Teleoperators and Virtual Environments*, vol. 6, no. 4, pages 355–385, aug 1997.

- [Azuma01] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier & Blair MacIntyre. *Recent Advances in Augmented Reality*. IEEE Comput. Graph. Appl., vol. 21, pages 34–47, IEEE Computer Society Press, Los Alamitos, CA, USA, November 2001.
- [Baudel95] Thomas Baudel. *Thèse de doctorat : Réalité Augmentée : pour une définition et une nouvelle dénomination du domaine*. préparée à l'Université Paris Sud, 1995.
- [Bederson95] Benjamin B. Bederson. *Audio augmented reality : a prototype automated tour guide*. In CHI '95 : Conf.companion on Human factors in computing systems, pages 210–211, New York, NY, USA, 1995. ACM.
- [Bellotti02] Francesco Bellotti, Riccardo Berta, Alessandro De Gloria & Massimiliano Margarone. *User Testing a Hypermedia Tour Guide*. IEEE Pervasive Computing, vol. 1, no. 2, pages 33–41, IEEE Educational Activities Department, Piscataway, NJ, USA, april 2002.
- [Benelli99] Giuliano Benelli, Alberto Bianchi, Patrizia Marti, David Sennati & E. Not. *HIPS : Hyper-Interaction within Physical Space*. Multimedia Computing and Systems, International Conference on, vol. 2, page 1075, IEEE Computer Society, Los Alamitos, CA, USA, 1999.
- [Bimber03] Oliver Bimber, L. Miguel Encarnação & Dieter Schmalstieg. *The virtual showcase as a new platform for augmented reality digital storytelling*. In Proceedings of the workshop on Virtual environments 2003, EGVE '03, pages 87–95, New York, NY, USA, 2003. ACM.
- [Blauert83] Jens Blauert. *Spatial Hearing : The Psychophysics of Human Sound Localization*. vol. 508, The MIT Press ; revised edition (October 2, 1996), Cambridge, MA, USA, 1983.
- [Bosseur96] Jean-Yves Bosseur. *Vocabulaire de la musique contemporaine*. Editions Minerve, 1996.
- [Boulin10] Jean Luc Boulin. *Réalité augmentée et M-tourisme en ville : Rotterdam et Amsterdam en Urban Augmented Reality*. mobil'it studio, November 2010.
- [Bourdot06] Patrick Bourdot, Damien Touraine & Jean-Marc Vézien. *Les techniques d'intégration multimodale en environnement virtuel*. In *Traité de la Réalité Virtuelle, Volume 1 : L'interfaçage, l'immersion et l'interaction en environnement virtuel*, volume 2 of *Collection Sciences mathématiques et informatiques*, pages 381–387. Les Presses de l'École des Mines de Paris, 2006.
- [Bouvier09] Patrice Bouvier. *Thèse de doctorat : La présence en réalité virtuelle, une approche centrée sur l'utilisateur*. Préparée à : L'Université Paris-Est, France, 2009.

- [Brungart99] Douglas S. Brungart. *Auditory localization of nearby sources. III. Stimulus effects*. Journal of the Acoustical Society of America, vol. 106, no. 6, pages 3589–3602, Altera Corporation, 1999.
- [Burr06] David Burr & David Alais. *Combining visual and auditory information*. In Visual Perception Fundamentals of Awareness : Multi-Sensory Integration and High-Order Perception, volume 155, Part B of *Progress in Brain Research*, pages 243 – 258. Elsevier, 2006.
- [B.V.09] Xsens Technologies B.V. *MTi and MTx User Manual and Technical Documentation*. The Netherlands, 2009. <http://edge.rit.edu/content/P10010/public/PDF/XsensMtx.pdf>, consulté le : 11/09/2013.
- [Carro06] Manuel Carro, José F. Morales, Henk L. Muller, G. Puebla & M. Hermenegildo. *High-level languages for small devices : a case study*. In CASES '06 : Proc. of the 2006 int. Conf. on Compilers, architecture and synthesis for embedded systems, pages 271–281, New York, NY, USA, 2006. ACM.
- [Carroll87] John M. Carroll & Mary Beth Rosson. *The paradox of the active user*. In *Interfacing thought : cognitive aspects of human-computer interaction*, pages 80–111. MIT Press, Cambridge, MA, USA, 1987.
- [Caruso00] Michael J Caruso. *Applications of magnetic sensors for low cost compass systems*. Position Location and Navigation Symposium, IEEE 2000, pages 177–184, IEEE, 2000.
- [Chalon04] René Chalon. *Thèse de doctorat : Réalité Mixte et Travail Collaboratif : IRVO, un modèle de l Interaction Homme- Machine*. Préparée à : L'Ecole Centrale de Lyon, France, 2004.
- [Cheverst00] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday & Christos Efstratiou. *Developing a context-aware electronic tourist guide : some issues and experiences*. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI '00, pages 17–24, New York, NY, USA, 2000. ACM.
- [Cheverst02] Keith Cheverst, Keith Mitchell & Nigel Davies. *The role of adaptive hypermedia in a context-aware tourist GUIDE*. Communications of the ACM Journal, vol. 45, no. 5, pages 47–51, ACM, New York, NY, USA, may 2002.
- [Childs03] Edward Childs & Ville Pulkki. *Using multi-channel spatialization in sonification : A case study with meteorological data*. Proceedings of the 9th International Conference on Auditory Display (ICAD2003), vol. 2003, Boston University Publications Production Department., Boston, USA, July 2003.

- [Chueng02] Priscilla Chueng. *Designing sound canvas : the role of expectation and discrimination*. In CHI '02 Extended Abstracts on Human Factors in Computing Systems, CHI EA '02, pages 848–849, New York, NY, USA, 2002. ACM.
- [Coleman68] Paul D. Coleman. *Dual Rôle of Frequency Spectrum in Determination of Auditory Distance*. Journal of the Acoustical Society of America, pages 631–632, Altera Corporation, 1968.
- [Crawford03] Chris Crawford. Chris crawford on game design. New Riders Games, jun 2003.
- [Damala09] Areti Damala. *Thèse de doctorat : Interaction Design and Evaluation of Mobile Guides for the Museum Visit : A Case Study in Multimedia and Mobile Augmented Reality*. Préparée au : Centre d'étude et de recherche en Informatique du CNAM, Paris, France, 2009.
- [Damala10] Areti Damala, Isabelle Astic & Aunis Coline. *PLUG, Université Paris Nuit : A Design Reiteration of a Mobile Museum Edutainment Application*. In VAST'10, Virtual Reality, Archaeology and Cultural Heritage, The 8th EUROGRAPHICS Workshop on Graphics and Cultural Heritage, Short and Project Papers, Paris, France, September 21-24, 2010, pages 91–94, September 2010.
- [Damala12] Areti Damala & Nenad Stojanovic. *Tailoring the Adaptive Augmented Reality (A2R) Museum Visit : Identifying Cultural Heritage Professionals ? Motivations and Needs*. Atlanta, Georgia, USA, 2012.
- [Damala13] Areti Damala, Tobias Schuchert, I. Rodriguez, J. Moragues, K. Gil-leade & Ljiljana Stojanovic. *Exploring the Affective Museum Visiting Experience : Adaptive Augmented Reality (A2R) and Cultural Heritage*. International Journal of Heritage in the Digital Era, vol. 2, no. 1, pages 117–141, Multi Science Publishing, 2013.
- [Davies09] Russell Davies. *Blog : La réalité augmentée sonore*, 2009. <http://www.la-realite-augmentee.fr/videos/la-realite-augmentee-sonore>, consulté le 03/09/2013.
- [Dean94] David Dean. *Museum exhibition : theory and practice. The heritage : care, preservation, management programme*. Routledge, Taylor and Francis Group, London, 1994.
- [Delerue02] Olivier Delerue & Olivier Warusfel. *Authoring of virtual sound scenes in the context of the LISTEN project*. In Proceedings of the AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio, Helsinki, 2002.
- [Deshays02] Sophie Deshays. *Audioguides et musées*. pages 24–31, Office de Coopération et d'Information Muséales (OCIM), 2002.

- [Deshays04] Sophie Deshays. *L usage des supports mobiles au musée, des audioguides classiques au multimédia nomade*. In Digital Culture and Heritage. Patrimoine and Culture Numérique (ICHIM 04), Haus der Kulturen der Welt, Berlin, Août 2004.
- [Dubois01a] Emmanuel Dubois. *Thèse de doctorat : Chirurgie Augmentée, un cas de Réalité Augmentée. Conception et réalisation centrées sur l'utilisateur*. ANRT, Grenoble, 2001. Préparée au Laboratoire de Communication Langagière et Interaction Personne-Système (IMAG), Université Joseph Fourier.
- [Dubois01b] Emmanuel Dubois, Laurence Nigay & Jocelyne Troccaz. *Consistency in Augmented Reality Systems*. In Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction, EHCI '01, pages 111–122, London, UK, 2001. Springer-Verlag.
- [Dünser07a] Andreas Dünser, Raphael Grasset, Hartmut Seichter & Mark Billinghurst. *Applying HCI principles to AR systems design*. Piscataway, NJ, USA, 2007. University of Canterbury. Human Interface Technology Laboratory.
- [Dünser07b] Andreas Dünser & Eva Hornecker. *An observational study of children interacting with an augmented story book*. In Proceedings of the 2nd international conference on Technologies for e-learning and digital entertainment, Edutainment'07, pages 305–315, Berlin, Heidelberg, 2007. Springer-Verlag.
- [Dünser08] Andreas Dünser, Raphaël Grasset & Mark Billinghurst. *A survey of evaluation techniques used in augmented reality studies*. In ACM SIGGRAPH ASIA 2008 courses, SIGGRAPH Asia '08, pages 1–27, New York, NY, USA, 2008. ACM.
- [Durand11] Peggy Durand. *Blog : La réalité augmentée, un nouvel outil au service du patrimoine*. 2011. <http://www.la-realite-augmentee.fr/videos/la-realite-augmentee-sonore>, consulté le 03/09/2013.
- [Eckel01a] Gerhard Eckel. *Immersive Audio-Augmented Environments - The LISTEN Project*. In 5th Int. Conf. on Information Visualization (IV2001), Los Alamitos, CA, USA, 2001.
- [Eckel01b] Gerhard Eckel. *The Vision of the LISTEN Project*. In Proceedings of the Seventh International Conference on Virtual Systems and Multimedia (VSMM'01), VSMM '01, pages 393–, Washington, DC, USA, 2001. IEEE Computer Society.
- [Faller04] Christof Faller & Juha Merimaa. *Source localization in complex listening situations : Selection of binaural cues based on interaural coherence*. Journal of the Acoustical Society of America, vol. 116, no. 5, pages 3075–3089, ASA, 2004.

- [Farman12] Jason Farman. *Mobile interface theory : Embodied space and locative media*. Media studies. Taylor & Francis Group, 2012.
- [Fauche08] Anne Fauche. *La médiation scientifique et culturelle de musée*. Laboratoire de Didactique et d'Épistémologie des Sciences, 2008.
- [Filippini03] Fantoni Silvia Filippini. *Personalization through it in museums. Does it really work? the case of the marble museum web site*. Proceedings of ICHIM03 on Cultural Institutions and Digital Technology, Paris, France, 2003.
- [Fink02] Josef Fink & Alfred Kobsa. *User Modeling for Personalized City Tours*. *Artif. Intell. Rev.*, vol. 18, no. 1, pages 33–74, Kluwer Academic Publishers, Norwell, MA, USA, September 2002.
- [Fodróczy08] Zoltán Fodróczy & András Radványi. *Localization of directional sound sources supported by a priori information of the acoustic environment*. *EURASIP J. Adv. Signal Process*, vol. 2008, Hindawi Publishing Corp., New York, NY, United States, jan 2008.
- [Gabbard08] Joseph L. Gabbard & J. Edward Swan II. *Usability Engineering for Augmented Reality : Employing User-Based Studies to Inform Design*. *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 3, pages 513–525, IEEE Educational Activities Department, Piscataway, NJ, USA, may 2008.
- [Gaudy08] Thomas Gaudy. *Thèse de doctorat : Etude et développements de jeux vidéo sonores accessibles aux personnes aveugles*. Préparée au : Centre d'étude et de recherche en Informatique du CNAM, Paris, France, 2008.
- [Gee04] James Paul Gee. *What video games have to teach us about learning and literacy*. Palgrave Macmillan, New York, 2004.
- [Gilkey95] Robert H. Gilkey & Janet M. Weisenberger. *The Sense of Presence for the Suddenly-Deafened Adult : Implications for Virtual Environments*. *Presence*, vol. 4, no. 4, pages 357–363, 1995.
- [Gleue01] Tim Gleue & Patrick Dähne. *Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project*. In Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage, VAST '01, pages 161–168, New York, NY, USA, 2001. ACM.
- [GoBmann02] Joachim GoBmann & Marcus Specht. *Location Models for Augmented Environments*. *Personal Ubiquitous Computing.*, vol. 6, pages 334–340, Springer-Verlag, London, UK, January 2002.
- [GoBmann04] Joachim GoBmann. *Connecting virtual sound and physical space in audioaugmented environments*. Proceedings of ICAD 04. Meeting of the

- International Conference on Auditory Display, Sydney, Australia, 2004.
- [Golledge91] Reginald G. Golledge, Jack M. Loomis, Roberta L. Klatzky, Andreas Flury & Xiao-Li Yang. *Designing a personal guidance system to aid navigation without sight : progress on the GIS component*. International Journal of Geographical Information Science, vol. 5, no. 4, pages 373–395, 1991.
- [Golledge98] Reginald G. Golledge, Roberta L. Klatzky, Jack M. Loomis, Jon Speigle & Jerome Tietz. *A Geographical Information System for a GPS Based Personal Guidance System*. International Journal of Geographical Information Science, vol. 12, no. 7, pages 727–749, 1998.
- [Gonot08] Antoine Gonot. *Thèse de doctorat : Conception et évaluation d interfaces de navigation dans les environnements sonores 3D*. Préparée au : Centre d'étude et de recherche en Informatique du CNAM, Paris, France, 2008.
- [Hatala04] Marek Hatala, Leila Kalantari, Ron Wakkary & Kenneth Newby. *Ontology and rule based retrieval of sound objects in augmented audio reality system for museum visitors*. In ACM symposium on applied computing, SAC '04, pages 1045–1050, New York, NY, USA, 2004. ACM.
- [Hatala05] Marek Hatala & Ron Wakkary. *Ontology-Based User Modeling in an Augmented Audio Reality System for Museums*. User Modeling and User-Adapted Interaction, vol. 15, no. 3-4, pages 339–380, Kluwer Academic Publishers, Hingham, MA, USA, 2005.
- [Heller09] Florian Heller, Thomas Knott, Malte Weiss & Jan Borchers. *Multi-user interaction in virtual audio spaces*. pages 4489–4494, ACM, New York, NY, USA, 2009.
- [Heller11] Florian Heller & Jan Borchers. *Corona : Audio Augmented Reality in Historic Sites*. MobileHCI 2011 Workshop on Mobile Augmented Reality : Design Issues and Opportunities, pages 51–54, Stockholm, Sweden, August 2011.
- [Holland02] Simon Holland, David R. Morse & Henrik Gedenryd. *AudioGPS : Spatial Audio Navigation with a Minimal Attention Interface*. Personal Ubiquitous Comput., vol. 6, no. 4, pages 253–259, Springer-Verlag, London, UK, 2002.
- [Inc.10] InterSense Inc. *INERTIACUBE BT : Dynamic wireless inertial orientation sensor*. MA 01821, USA, 2010. http://www.est-kl.com/fileadmin/media/pdf/InterSense/InertiaCube_BT_Datasheet.pdf, consulté le 11/09/2013.

- [Jacob94] Robert J. K. Jacob, Linda E. Sibert, Daniel C. Mcfarlane & M. Preston Mullen. *Integrality and Separability of Input Devices*. ACM Transactions on Computer-Human Interaction, 1994.
- [Jacomy02] Bruno Jacomy. *L'âge du plip, chroniques de l'innovation technique*. Editions du Seuil, Février 2002.
- [Jones08] Matt Jones, Steve Jones, Gareth Bradley, Nigel Warren & David Bainbridge. *ONTRACK : Dynamically adapting music playback to support navigation*. Personal and Ubiquitous Computing, pages 513–525, 2008.
- [Jot99] Jean-Marc Jot. *Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces*. Multimedia Syst., vol. 7, no. 1, pages 55–69, Springer-Verlag New York, Inc., Secaucus, NJ, USA, January 1999.
- [Kang96] Seong-Hoon Kang & Sung-Han Kim. *Realistic Audio Teleconferencing using Binaural and Auralization Techniques*. Etri Journal, vol. 18, pages 41–51, 1996.
- [Katz10] Brian F.G. Katz, Philippe Truillet, Simon Thorpe & Christophe Jouffrais. *NAVIG : Navigation Assisted by Artificial Vision and GNSS*. Workshop on Multimodal Location Based Techniques for Extreme Navigation - Pervasive 2010 conference, ACM, Helsinki, Finland, 2010.
- [Kionix07] Kionix. *Tilt-Sensing with Kionix MEMS Accelerometers*. Application Note AN005, 2007.
- [Kobsa01] Alfred Kobsa, Jürgen Koenemann & Wolfgang Pohl. *Personalised hypermedia presentation techniques for improving online customer relationships*. Knowl. Eng. Rev., vol. 16, no. 2, pages 111–155, Cambridge University Press, New York, NY, USA, march 2001.
- [Koizumi92] Nobuo Koizumi, Michael Cohen & Shigeaki Aoki. *Design of Virtual Conferencing Environment in Audio Telecommunication*. Audio Engineering Society Convention 92, March 1992.
- [Lefebvre91] Bernard Lefebvre & Hélène Lefebvre. *Le visiteur, le guide et l'éducation*. Revue canadienne de l'éducation, pages 331–337, 1991.
- [Lemordant11] Jacques Lemordant. *Indoor Outdoor AR Navigation - Autonomy Projet (Global Competitive Cluster Minalogic)*. Free & Open Source Software for GeoSpatial (FOSS4G), Osaka, Japan, 2011.
- [Leprado07] Cécile Leprado. *LISTEN LISBOA : scripting languages for interactive musical installations*. 4th Sound and Music Computing Conference (SMC), Lefkada, Greece, July 2007.
- [Légal07] Jean-Baptiste Légal. *Cours universitaire : Méthodologie Expérimentale*. Université Paris X-Nanterre, 2007. <http://j.b.legal.free>.

- fr/Blog/share/Dynamiques/Methodo.pdf, consulté le : 11/09/2013.
- [Loomis98] Jack M. Loomis, Roberta L. Klatzky & Reginald G. Golledge. *Auditory distance perception in real, virtual, and mixed environments*. In *Mixed reality : Merging real and virtual worlds*, pages 201–214, Tokyo, Japon, 1998.
- [Lyons00] Kent Lyons, Maribeth Gandy & Thad Starner. *Guided by Voices : An Audio Augmented Reality System*. In *PROCEEDINGS OF INTL. CONF. ON AUDITORY DISPLAY (ICAD) 2000*, 2000.
- [Mackay93] Wendy Mackay, Gilles Velay, Kathy Carter, Chaoying Ma & Daniele Pagani. *Augmenting reality : adding computational dimensions to paper*. *Commun. ACM*, vol. 36, pages 96–97, ACM, New York, USA, July 1993.
- [Mackay96] Wendy Mackay. *Réalité augmentée : le meilleur des deux mondes. La recherche*, pages 32–37, 1996. <http://www.larecherche.fr/savoirs/dossier/realite-augmentee-meilleur-deux-mondes-01-03-\1996-87977>, consulté le : 11/09/2013.
- [Mackay98] Wendy E. Mackay. *Augmented reality : linking real and virtual worlds : a new paradigm for interacting with computers*. In *Proceedings of the working conference on Advanced visual interfaces, AVI '98*, pages 13–21, New York, USA, 1998. ACM.
- [Mallem08] Malik Mallem & David Roussel. *Realite augmentee : Principes, technologies et applications*. Ed. Techniques Ingénieur, Référence TE5920, 2008.
- [Mariette09] Nicholas Mariette & Brian F.G. Katz. *Sounddelta - Large scale, Multi-user audio augmented reality*. *Proc. of the EAA Symposium on Auralization*, Espoo, Finland, 2009.
- [Mariette10] Nicholas Mariette. *Navigation performance effects of render method and head-turn latency in mobile audio augmented reality*. In *Proceedings of the 6th international conference on Auditory Display, CMMR/I-CAD'09*, pages 239–265, Berlin, Heidelberg, 2010. Springer-Verlag.
- [Middlebrooks91] John C. Middlebrooks & David M. Green. *Sound Localization by Human Listeners*. *Annual Review of Psychology*, vol. 42, no. 1, pages 135–159, Department of Neuroscience, University of Florida, Gainesville 32610., 1991.
- [Milgram94a] Paul Milgram & Fumio Kishino. *A taxonomy of mixed reality visual displays*. *IEICE Transactions on Information Systems*, 1994.
- [Milgram94b] Paul Milgram, Haruo Takemura, Akira Utsumi & Fumio Kishino. *Augmented Reality : A Class of Displays on the Reality-Virtuality Conti-*

- nuum*. pages 282–292, ATR Communication Systems Research Laboratories. Kyoto 619-02, Japan, 1994.
- [Milgram99] Paul Milgram & Herman Colquhoun Jr. *A Taxonomy of Real and Virtual World Display Integration*. International Symposium on Mixed Reality, 1999.
- [Mina10] An Xiao Mina. *Louvre-DNP Museum Lab : Augmented Reality Dish*. 2010. <http://www.frequency.com/video/louvre-dnp-museum-lab-augmented-reality/2260257>, consulté le 10/09/2013.
- [Mladenic96] Dunja Mladenic. *Personal WebWatcher : Implementation and Design*. Rapport technique IJS-DP-7472, Department for Intelligent Systems, J. Stefan Institute, Slovenia, 1996.
- [Murat08] Aytakin Murat, Moss Cynthia F. & Simon Jonathan Z. *A Sensorimotor Approach to Sound Localization*. ACM Trans. Appl. Percept., vol. 20, no. 3, pages 603–635, MIT Press, Cambridge, MA, Etats-unis, 2008.
- [Nees08] Michael A. Nees & Bruce N. Walker. *Data density and trend reversals in auditory graphs : Effects on point-estimation and trend-identification tasks*. ACM Trans. Appl. Percept., vol. 5, no. 3, pages 13 :1–13 :24, ACM, New York, NY, USA, sep 2008.
- [Oberlander97] Jon Oberlander, Chris Mellish, Mick O'Donnell & Alistair Knott. *Exploring a Gallery with Intelligent Labels*. In International Cultural Heritage Informatics Meetings (ICHIM), pages 80–87, 1997.
- [O'Brien08] Trevor M. O'Brien, Daniel F. Keefe & David H. Laidlaw. *A case study in using gestures and bimanual interaction to extend a high-DOF input device*. In Proceedings of the 2008 symposium on Interactive 3D graphics and games, I3D '08, pages 9 :1–9 :1, New York, NY, USA, 2008. ACM.
- [O'donnell01] Michael O'donnell, Chris Mellish, J Jon Oberlander & Alistair Knott. *ILEX : an architecture for a dynamic hypertext generation system*. Nat. Lang. Eng., vol. 7, no. 3, pages 225–250, Cambridge University Press, New York, NY, USA, September 2001.
- [Oppermann99] Reinhard Oppermann, Marcus Specht & Igor Jaceniak. *Hippie : a nomadic information system*. In 1st international symposium on Handheld and Ubiquitous Computing, HUC '99, pages 330–333, London, UK, 1999. Springer-Verlag.
- [Owen05] Ruth Owen, Dimitrios Buhalis & Daniël Pletinckx. *Visitors evaluations of ICTs used in cultural heritage*. In Proceedings of the 6th International conference on Virtual Reality, Archaeology and Intelligent Cultural Heritage, VAST'05, pages 129–136, Aire-la-Ville, Switzerland, Switzerland, 2005. Eurographics Association.

- [Papagiannakis02] George Papagiannakis, Michal Ponder, Tom Molet, Sumedha Kshirsagar, Frederic Cordier, Nadia Magnenat-thalmann & Daniel Thalmann. *LIFEPLUS : Revival of life in ancient Pompeii*. In In Proceedings of the 8th International Conference on Virtual Systems and Multimedia (VSMM 02), pages 25–27, Gyeongju, Korea, 2002.
- [Paterson10] Natasa Paterson, Katsiaryna Naliuka, Soren Kristian Jensen, Tara Carrigy, Mads Haahr & Fionnuala Conway. *Design, implementation and evaluation of audio for a location aware augmented reality game*. In Fun and Games '10 : Proceedings of the 3rd International Conference on Fun and Games, pages 149–156, New York, NY, USA, 2010. ACM.
- [Pazzani97] Michael Pazzani & Daniel Billsus. *Learning and Revising User Profiles : The Identification of Interesting Web Sites*. Machine Learning, vol. 27, pages 313–331, Kluwer Academic Publishers, Hingham, MA, USA, June 1997.
- [Peters10] Nils Peters. *Thèse de doctorat : Sweet reproduction : Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception*. Préparée à : Music Technology Area, Department of Music Research Schulich School of Music, McGill University Montreal, QC, Canada, 2010.
- [Petrelli99] Daniela Petrelli, Elena Not, Marcello Sarini, Oliviero Stock, Carlo Strapparava & Massimo Zancanaro. *HyperAudio : location-awareness + adaptivity*. In extended abstracts on Human factors in computing systems, CHI EA '99, pages 21–22, New York, NY, USA, 1999. ACM.
- [Petrelli01] Daniela Petrelli, Elena Not, Massimo Zancanaro, Carlo Strapparava & Oliviero Stock. *Modelling and Adapting to Context*. Personal Ubiquitous Comput., vol. 5, pages 20–24, Springer-Verlag, London, UK, January 2001.
- [Petrie96] Helen Petrie, Valerie Johnson, Thomas Strothotte, Andreas Raab, Steffi Fritz & Rainer Michel. *MOBIC : Designing a Travel Aid for Blind and Elderly People*. The Journal of Navigation, vol. 49, no. 01, pages 45–52, 1996.
- [Pontus07] Larsson Pontus & Våljamäe Aleksander. *Auditory-visual perception of room size in virtual environments*. In Proceedings of the 19th International Congress on Acoustics, Madrid, Spain, September 2-7, 2007.
- [Proctor03] Nancy Proctor & Jane. Burton. *Tate modern multimedia tour pilot*. In MLEARN 2003 conference, Learning with Mobile Devices, pages 127–130., The King's College London & Tate Britain Colloquium, London, May 2003.

- [Proctor05] Nancy Proctor. *Off Base or On Target? Pros and cons of wireless and location-aware applications in the museum*. In Digital Culture and Heritage Patrimoine and Culture Numérique (ICHIM 05), Bibliothèque nationale de France, Paris, Septembre 2005.
- [Ran04] Lisa Ran, Sumi Helal & Steve Moore. *Drishti : An Integrated Indoor/Outdoor Blind Navigation System and Service*. In Proceedings of the Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04), PERCOM '04, page 23, Washington, DC, USA, 2004. IEEE Computer Society.
- [Randell04] Cliff Randell, Sara Price, Yvonne Rogers, Eric Harris & Geraldine Fitzpatrick. *The Ambient Horn : designing a novel audio-based learning experience*. Personal Ubiquitous Comput., vol. 8, no. 3-4, pages 177–183, Springer-Verlag, London, UK, 2004.
- [Regenbrecht04] H Holger Regenbrecht, T. Lum, P. Kohler, Claudia Ott, Michael Wagner, Wilhelm Wilke & E. Mueller. *Using augmented virtuality for remote collaboration*. Presence : Teleoper. Virtual Environ., vol. 13, no. 3, pages 338–354, MIT Press, Cambridge, MA, USA, 2004.
- [Roads07] Curtis Roads. *L'audionumérique : Musique et informatique*, volume 672. Dunod édition : 2e édition révisée et actualisée (31 octobre 2007), 2007.
- [Robson94] Colin Robson. *Experiment, design and statistics in psychology*. Penguin Psychology Books, 1994.
- [Roentgen08] Uta R Roentgen, Gert Jan Gelderblom, Mathijs Soede & Luc P De Witte. *Inventory of Electronic Mobility Aids for Persons with Visual Impairments : A Literature Review*. Journal of Visual Impairment and Blindness, vol. 102, no. 11, pages 702–724, American Foundation for the Blind, 2008.
- [Roetenberg05] Daniel Roetenberg, Henk J. Luinge, Chris T. M. Baten & Peter H. Veltink. *Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 13, pages 395–405, 2005.
- [Rolim12] Cledja Rolim & Veronica Teichrieb. *A Viewpoint about Diminished Reality : Is it Possible Remove Objects in Real Time from Scenes?* pages 141–146, IEEE Computer Society, Washington, DC, USA, 2012.
- [Scapin97] Dominique L. Scapin & J. M. Christian Bastien. *Ergonomic criteria for evaluating the ergonomic quality of interactive systems*. Behaviour and Information Technology, pages 220–231, Taylor and Francis Ltd, July 1997.
- [Schafer10] Raymond Murray Schafer. *Le paysage sonore : Le monde comme musique*. Wildproject Editions, 2010.

- [Serrell96] Beverly Serrell. *The Question of Visitor Styles*. Jacksonville, Alabama, 1996. Center for Social Design.
- [Sodnik05] Jaka Sodnik, Rudolf Susnik, Mitja Stular & Saso Tomazic. *Spatial sound resolution of an interpolated HRIR library*. *Applied Acoustics*, Elsevier, Kidlington, UK, vol. 66, no. 4, pages 1219–1234, 2005.
- [Sodnik06] Jaka Sodnik, Saso Tomazic, Raphael Grasset, Andreas Duenser & Mark Billinghurst. *Spatial sound localization in an augmented reality environment*. In *OZCHI '06 : Proc. of the 18th Australia Conf. on Computer-Human Interaction*, pages 111–118, Sydney, Australia, 2006. ACM.
- [Sparacino02] Flavia Sparacino. *The Museum Wearable : Real-Time Sensor-Driven Understanding of Visitors Interests for Personalized Visually-Augmented Museum Experiences*. In *Museums and the Web*, Boston, Massachusetts, U.S.A, 2002.
- [Sparkfun07] Electronics. Sparkfun. *IMU 6 Degrees of Freedom v4 Data Sheet*. Colorado, USA, 2007. <https://www.sparkfun.com/datasheets/Sensors/IMU/6DOF-v4-Rev1.pdf>, consulté le : 10/09/2013.
- [Specht99] Marcus Specht, Reinhard Oppermann & Marcus Specht. *User modeling and adaptivity in nomadic information systems*. In Mimo Caenepeel, David Benyon and Duncan Smith (eds.), *Proceedings of the i3 Annual Conference : Community of the Future*, pages 65–68, 1999.
- [Stahl07] Christoph Stahl. *The roaring navigator : a group guide for the zoo with shared auditory landmark display*. pages 383–386, ACM, New York, NY, USA, 2007.
- [Stein11] Robert Stein & Nancy Proctor. *TourML : An Emerging Specification for Museum Mobile Experiences*. In *Museums and the Web*, Philadelphia, PA, U.S.A, April 2011.
- [Stock07] Oliviero Stock & Massimo Zancanaro. *Peach : Intelligent interfaces for museum visits*. Springer, Berlin, 2007.
- [Stojanovic12] Nenad Stojanovic, Areti Damala, Tobias Schuchert, Ljiljana Stojanovic, Stephen Fairclough & John Moores. *Tutorial 1 : Adaptive augmented reality (A2R) : Where AR meets user s interest*. pages 1–1, IEEE Computer Society, Washington, DC, USA, 2012.
- [Strachan05] Steven Strachan, Parisa Eslambolchilar, Roderick Murray-Smith, Stephen Hughes & Sile O'Modhrain. *GpsTunes : controlling navigation via audio feedback*. In *MobileHCI '05 : Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 275–278, New York, NY, USA, 2005. ACM.

- [Strachan09] Steven Strachan & Roderick Murray-Smith. *Bearing-based selection in mobile spatial interaction*. Personal Ubiquitous Comput., vol. 13, no. 4, pages 265–280, Springer-Verlag, London, UK, 2009.
- [Sutherland68] Ivan E. Sutherland. *A head-mounted three dimensional display*. In Proceedings of the December 9-11, 1968, fall joint computer conference, part I, pages 757–764, New York, NY, USA, 1968. ACM.
- [Swan05] J. Edward II Swan & Gabbard Joseph L. *Survey of User-Based Experimentation in Augmented Reality*. In Proceedings of 1st International Conference on Virtual Reality, HCI International 2005, July 22-27 2005.
- [Sylaiou05] Stella Sylaiou, Fotis Liarokapis, Lazaros Sechidis, Petros Patias & Olga Georgoula. *Virtual museums : First results of a survey on methods and tools*. In the XXth International Symposium (the ICOMOS ISPRS Committee on Documentation of Cultural Heritage), pages 1138–1143, the CIPA International Archives for Documentation of Cultural Heritage, 2005.
- [Sylaiou08] Styliani Sylaiou, M. Economou, Athanasis Karoulis & Martin White. *The evaluation of ARCO : a lesson in curatorial competence and intuition with new technology*. Computers in Entertainment, vol. 6, no. 2, pages 23 :1–23 :18, ACM, New York, NY, USA, July 2008.
- [Taylor07] Robyn Taylor, Maryia Kazakevich, Pierre Boulanger, Manuel Garcia & Walter F. Bischof. *Multi-modal Interface for Fluid Dynamics Simulations Using 3-D Localized Sound*. In Proceedings of the 8th international symposium on Smart Graphics, SG '07, pages 182–187, Berlin, Heidelberg, 2007. Springer-Verlag.
- [Thian12] Cherry Thian. *Augmented Reality ?What Reality Can We Learn From It ?* In Museums and the Web, San Diego, CA, USA, 2012.
- [Vazquez09] Yolanda Vazquez Alvarez & Stephen Brewster. *Investigating background and foreground interactions using spatial audio cues*. In Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, CHI EA '09, pages 3823–3828, New York, NY, USA, 2009. ACM.
- [Véron96] Eliséo Véron & Martine Levasseur. *Ethnographie de l'exposition, l'espace le corps et le sens*. page 179, Bibliothèque publique d'information Du Centre Georges Pompidou, Paris, Mars 1996.
- [Vilatte07] Jean-Christophe Vilatte. *Audioguides et musées*. Formation « Evaluation », 2007.
- [Virolainen01] Jussi Virolainen. *Thèse de Master : Design and Implementation of a Stereo Audio Conferencing System*. préparée à : Helsinki University of Technology, 2001.

- [Wakkary07] Leprado Wakkary & Marek Hatala. *Situated play in a tangible interface and adaptive audio museum guide*. Personal Ubiquitous Computing, vol. 11, pages 171–191, Springer-Verlag, London, UK, February 2007.
- [Warusfel04] Olivier Warusfel & Gerhard Eckel. *LISTEN - Augmenting everyday environments through interactive soundscapes*. 2004. Virtual Reality for Public Consumption, IEEE Virtual Reality 2004 Workshop, Chicago IL 27.
- [Wellner93] Pierre Wellner. *Interacting with paper on the DigitalDesk*. Commun. ACM, vol. 36, no. 7, pages 87–96, ACM, New York, NY, USA, jul 1993.
- [Wilson07] Jeff Wilson, Bruce N. Walker, Jeffrey Lindsay, Craig Cambias & Frank Dellaert. *SWAN : System for Wearable Audio Navigation*. In Proceedings of the 11th International Symposium on Wearable Computers (ISWC), pages 91–98. IEEE, 2007.
- [Woodruff01] Allison Woodruff, Paul M. Aoki, Amy Hurst & Margaret H. Szymanski. *Electronic Guidebooks and Visitor Attention*. In Proceedings of the International Conference on Cultural Heritage and Technologies in the Third Millennium, pages 437–454, 2001.
- [Wuestefeld92] A. Paige Wuestefeld. *Auditory looming perception : Influences on anticipatory judgments*. Journal of The Acoustical Society of America, vol. 92, 1992.
- [Zahorik05] Pavel Zahorik, Douglas S Brungart & Adelbert W Bronkhorst. *Auditory distance perception in humans : a summary of past and present research*. Acta Acoustica united with Acustica, vol. 91, pages 409 – 420, S. Hirzel Verlag, 2005.
- [Zimmermann08] Andreas Zimmermann & Andreas Lorenz. *LISTEN : a user-adaptive audio-augmented museum guide*. User Modeling and User-Adapted Interaction, vol. 18, no. 5, pages 389–416, Kluwer Academic Publishers, Hingham, MA, USA, 2008.

Modèle et expériences pour la visite de musées en réalité augmentée sonore

Résumé

Cette thèse consiste à explorer l'usage de la réalité augmentée sonore pour la visite de musées. Notre objectif est de proposer un audioguide permettant d'immerger le visiteur dans une scène sonore constituée de sons ambiants et de commentaires associés aux objets exposés, et de minimiser ses efforts pour découvrir ces objets et interagir avec cet environnement sonore. La première contribution de cette thèse concerne la mise en place d'une preuve de concept de SARIM (Sound Augmented Reality Interface for visiting Museum). Cette preuve de concept a été développée en utilisant des capteurs de position et d'orientation filaire et non filaire. La deuxième contribution concerne la modélisation de la visite augmentée par la dimension sonore. Après une étude des modèles existants, l'objectif est de concevoir un modèle comprenant une représentation du visiteur, du paysage sonore et de la navigation et offrant une grande flexibilité pour créer l'environnement sonore. Ce modèle a comme finalité de faciliter la conception de différents types de scénarios de visite sur la base de la notion de zone d'audibilité. La troisième contribution de cette thèse est le travail d'évaluation mené dans un environnement réel qu'est le musée des arts et métiers de Paris, et qui a permis de confirmer l'utilisabilité, ainsi que l'apport didactique et ludique que procure la réalité augmentée sonore en général et le système SARIM en particulier pour prolonger et enrichir la visite de musées.

Mots clés

Réalité augmentée, visite musée, captation de mouvement, personnalisation, modélisation, spatialisation sonore, audioguide

Abstract

The goal of this thesis is to explore the use of sound to enhance the museum visit. We aim to provide an audioguide to immerse the visitor in a soundstage consisting of ambient sounds and comments associated with the exhibits, and minimize its efforts to discover these objects and interact with the sound environment. The first contribution of this thesis is the implementation of a proof of concept of SARIM (Sound Augmented Reality Interface for visiting Museum). This proof of concept was developed using position sensors and guidance wired and wirelessly. The second contribution concerns the modeling of the augmented visit by the sound dimension. After a review of existing models, the objective is to design a model that includes a representation of the visitor, the soundscape and navigation, offering the flexibility to create the sound environment. This model has the purpose of facilitating the design of different types of scenarios based on the concept of audibility area. The Third contribution of this thesis is the evaluation conducted in a real environment what the Museum of Arts and Crafts in Paris, which confirmed the usability, as well as providing educational and ludic impact of the audio augmented reality in general and the SARIM in particular to extend and enrich the museum visits.

Keywords

Augmented reality, museum visit, motion tracking, context aware, adaptive system, user modeling, sound spatialization, audioguide