

DÉPARTEMENT D'INFORMATIQUE

PROJET DE FIN D'ÉTUDES

MASTER SCIENCES ET TECHNIQUES
SYSTÈMES INTELLIGENTS & RÉSEAUX

RÉALISATION D'UN PROTOTYPE D'UNE INTERFACE MULTISENSORIELLE POUR LES MALVOYANTS



LIEU DU STAGE : Laboratoire MIS (Modélisation, Information & Systèmes)

Université de Picardie Jules Verne, Amiens, France

Réalisé par :

- Hamza ASBAYO

Encadré par :

- Pr. Younes DHASSI
- Pr. Dominique GROUX
- Pr. Alexis POTELLE

Soutenu le 13.07.2021 devant le jury composé de :

- | | | |
|--------------------------|---|----------------|
| - Pr. Mohammed ABOUNAIMA | Faculté des Sciences et Techniques de Fès | (Président) |
| - Pr. Loubna LAMRINI | Faculté des Sciences et Techniques de Fès | (Examinatrice) |
| - Pr. Dominique GROUX | Université de Picardie Jules Verne d'Amiens | (Encadrante) |
| - Pr. Younes DHASSI | Faculté des Sciences et Techniques de Fès | (Encadrant) |

Année Universitaire 2020 – 2021

Remerciements

Avant tout développement sur cette expérience, il apparaît opportun de commencer ce rapport de stage par des remerciements.

En premier lieu, je tiens à remercier Mme. Dominique GROUX et M. Alexis POTELLE membres de l'équipe PR du laboratoire MIS d'Amiens, de m'avoir accordé l'opportunité de découvrir et participer à la réalisation d'un projet aussi innovant. Je leur suis aussi redevable pour leur encadrement, tous leurs conseils, le soutien dont ils m'ont fait bénéficier, de leurs encouragements et les aides précieuses qu'il n'ont pas cessés de m'apporter tout au long de la période de ce stage.

Je tiens également à remercier et à témoigner toute ma reconnaissance à mon encadrant pédagogique M. Younes DHASSI pour son encadrement, son assistance et surtout pour la confiance qu'il m'a accordée.

Je remercie Pr. Mohammed CHAOUKI ABOUNAIMA et Pr. Loubna LAMRINI membres de jury, pour m'avoir honoré en acceptant d'évaluer et de juger ce travail.

J'adresse également des remerciements particuliers à l'ensemble du corps professoral et administratif de La Faculté des Sciences et Techniques de Fès (FSTF), pour l'incalculable qualité de l'enseignement qu'ils assurent.

Un grand merci est également adressé à mes collègues du Master, pour leur présence, leur participation et leurs interventions qui ont contribué à l'approfondissement de nos connaissances et l'enrichissement de cette expérience.

Que tous ceux et celles qui ont contribué de près ou de loin à l'accomplissement de ce travail trouvent l'expression de mes remerciements les plus chaleureux et les mots du respect les plus distingués.

Résumé

Afin de mettre à disposition des malvoyants le patrimoine architectural, l'équipe PR du laboratoire MIS d'Amiens a initié le projet PHARES (**P**ercevoir grâce à une interface **H**Aptique pour **l**Es malvoyant**S**) qui vise à utiliser les modèles 3D du patrimoine pour apporter aux malvoyants une solution multisensorielle qui diffère des expériences actuelles.

Dans ce cadre vient ce projet qui a comme but, d'initier la conception et la réalisation du système en réalisant un prototype sous Unity, qui permet de détecter l'axe de regard de l'utilisateur ainsi que la synchronisation de la vue à un autre sens (celui de l'audition) et lancer un fichier audio issu d'une base de données selon la partie de l'objet 3D observée.

***Mots clés :** patrimoine architectural, multisensorielle, déficience visuelle, modèle 3D, Unity*

Abstract

In order to make the architectural patrimony available to the visually impaired, the PR team of the MIS laboratory in Amiens has initiated a project intitled PHARES (*Percevoir grâce à une interface HAptique pour lEs malvoyantS*). The project aims to use 3D models of the architectural patrimony to provide the visually impaired with a multisensory experience that differs from current experiences.

Within this context, comes this project which aims to initiate the design and realization of the system by creating a prototype using the game engine Unity. The implemented prototype allows to detect the user's line of sight, as well as the synchronization of the sight with another sense (the hearing one) and to launch an audio file from a database depending on the observed part of the 3D object.

Keywords: architectural patrimony, multisensory, visual impairment, 3D model, Unity

Table des matières

Remerciements.....	i
Résumé.....	ii
Abstract.....	iii
Liste des tableaux	vii
Liste des figures	viii
Liste d'Acronymes et Abréviations.....	x
Introduction générale.....	1
Chapitre 1 : Cadre général du projet	2
Introduction.....	3
1. Présentation du laboratoire d'accueil	3
1.1. Laboratoire MIS	3
1.2. Equipe PR	4
2. Projet PHARES.....	5
3. Problématique.....	5
4. Cadre du projet.....	6
4.1. Missions	6
4.2. Conduite du projet	7
4.2.1. Processus de développement.....	7
4.2.2. Planification.....	9
Conclusion.....	9
Chapitre 2 : Contexte et état de l'art	10
Introduction.....	11
1. Système visuel humain.....	11
1.1. L'œil.....	11
1.1.1. La cornée.....	12
1.1.2. L'iris.....	12
1.1.3. Le cristallin	12
1.1.4. La rétine	12
2. Les déficiences visuelles	13
2.1. Classifications des déficiences visuelles.....	14
2.2. Les causes	14

2.3.	Description des pathologies.....	15
2.3.1.	La cataracte.....	15
2.3.2.	Le glaucome.....	16
2.3.3.	La dégénérescence maculaire liée à l'âge.....	17
3.	Relevé numérique du patrimoine.....	17
3.1.	Outils et méthodes d'acquisition 3D.....	17
3.1.1.	Photogrammétrie.....	18
3.1.2.	Lasergrammétrie.....	20
3.2.	Traitement des données et enrichissement sémantique.....	22
3.2.1.	Segmentation.....	22
3.2.2.	Annotation sémantique.....	24
3.2.3.	Bilan.....	26
	Conclusion.....	26
	Chapitre 3 : Analyse et conception.....	27
	Introduction.....	28
1.	Objectifs du projet.....	28
2.	Modélisation.....	30
2.1.	Diagramme des cas d'utilisation.....	30
2.1.1.	Identification des acteurs.....	30
2.1.2.	Diagramme des Cas d'utilisation de l'assistant.....	30
2.1.3.	Diagramme des Cas d'utilisation du malvoyant.....	32
2.2.	Diagramme de séquence.....	33
2.2.1.	Paramétrage de la scène.....	33
2.2.2.	Déclencher un fichier audio.....	34
2.3.	Fonctionnement réel.....	35
3.	Architecture générale du système.....	37
	Conclusion.....	38
	Chapitre 4 : Mise en œuvre.....	39
	Introduction.....	40
1.	Identification des outils de travail.....	40
1.1.	Unity.....	40
1.2.	Blender.....	41
1.3.	C#.....	42

1.4. Visual studio Community	42
1.5. MySQL	43
2. Segmentation avec Blender	43
3. Présentation de l'interface	45
3.1. Vue d'entrée	45
3.2. Vue de choix de pathologie.....	46
3.3. Vue de choix de la scène.....	46
3.4. Vues de la scène 3D	46
3.4.1. Vues de l'intérieure de la scène	47
3.4.2. Modèle de l'ange.....	47
Conclusion.....	49
Conclusion & perspectives.....	50
Bibliographie et Webographie.....	51

Liste des tableaux

Tableau 1 : Liste d'Acronymes et Abréviations.....	x
Tableau 2: Plan de communication entre les acteurs du stage.....	8
Tableau 3: Les différents types et niveaux de déficience visuelle	14
Tableau 4: Description des cas d'utilisation de l'assistant.....	31
Tableau 5: Description des cas d'utilisation du malvoyant.....	32

Liste des figures

Figure 1: Organigramme du laboratoire MIS	4
Figure 2: Image originale.....	6
Figure 3: Image de sortie après application du filtre "Glare"	6
Figure 4: Cycle de développement XP	8
Figure 5: Diagramme de GANTT.....	9
Figure 6: Schéma en coupe de l'œil humain [1].....	11
Figure 7: Schéma en coupe de la rétine humaine [3]	13
Figure 8: les limites normales du champ visuel [4, p. 7].....	13
Figure 9: Simulation d'une vision normale; et d'une vision dans le cas de la cataracte [7]	15
Figure 10: Schéma d'un œil atteint de glaucome [8].....	16
Figure 11: Simulation d'une vision normale; et d'une vision dans le cas du glaucome [7]	16
Figure 12: Simulation d'une vision normale (à gauche) et d'une vision dans le cas de la DMLA (à droite) [7].....	17
Figure 13: Classification des principales techniques d'acquisition 3D [10].....	18
Figure 14: Les différentes techniques de restitution photogrammétrique. De gauche à droite : mono, stéréo, multi-images [11, Fig. 27].....	19
Figure 15: Reconstruction 3D à partir d'un ensemble d'images terrestres d'une partie de la cathédrale d'Amiens (E-Cathedrale (MIS, ENSG, IGN), 2011)	20
Figure 16: Système de balayage du scanner laser [10, Fig. 27].....	21
Figure 17: Nuage de points représentant un modèle photogrammétrique d'un ange [14]	22
Figure 18: Gauche : Image originale. Droite : Image segmentée [10].....	23
Figure 19: Modèle 3D d'un ange segmenté.....	23
Figure 20: Les modèles du dispositif LISA	28
Figure 21: Modèle d'un ange segmenté-annoté (Source : cahier de charge confidentiel)	29
Figure 22: Les acteurs.....	30
Figure 23: Diagramme des cas d'utilisation de l'assistant	31
Figure 24: Diagramme des cas d'utilisation du malvoyant	32
Figure 25: Diagramme de séquence -Paramétrage de la scène-.....	34
Figure 26 : Diagramme de séquence -Déclencher fichier audio-	35
Figure 27: Schéma représentant l'architecture générale du système.....	37
Figure 28: Modèle 3D de l'ange avant segmentation.....	44
Figure 29: Modèle de l'ange 3D segmenté dans Blender	44
Figure 30: La fenêtre Hierarchy de Unity	45
Figure 31: Vue d'accueil de l'application.....	45
Figure 32: Vue de choix de pathologie	46
Figure 33: Vue de choix de scène ou objet 3D	46
Figure 34: Illustration de quelques angles de vue de la scène 3D	47
Figure 35: Modèle 3D de l'ange	47

Figure 36: Illustration représentant la ligne de regard de l'utilisateur 48

Figure 37: Illustration du bon fonctionnement de la synchronisation de la vue et l'audition. 48

Figure 38: La table de la base de données contenant le fichier audio récupéré..... 48

Liste d'Acronymes et Abréviations

DMLA	<i>Dégénérescence Maculaire Liée à l'Age</i>
MIS	<i>Modélisation, Information & Systèmes</i>
PHARES	<i>Percevoir grâce à une interface HAptique pouR IEs malvooyantS</i>
PR	<i>Perception Robotique</i>
COVE	<i>Commande et Véhicules</i>
GOC	<i>Graphes, Optimisation et Contraintes</i>
SDMA	<i>Systèmes Distribués, Mots et Applications</i>
CMN	<i>Centre des Monuments Nationaux</i>
DRAC	<i>Direction Régional des Affaires Culturelles</i>
LISA	<i>muLti-sensoriel pour leS MalvoyAnts</i>
XP	<i>eXtreme Programming</i>
OMS	<i>Organisation Mondiale de la Santé</i>
UML	<i>Unified Modeling Language</i>
VR	<i>Virtual Reality</i>
API	<i>Application Programming Interface</i>
PSD	<i>Photoshop Data file</i>
JPEG	<i>Joint Photographic Experts Group</i>
PNG	<i>Portable Network Graphics</i>
SVG	<i>Scalable Vector Graphics</i>
OGG	<i>Operation Good Guys</i>
WAV	<i>Waveform Audio File Format</i>
B3D	<i>Bodypaint 3D Format</i>
FBX	<i>Filmbox</i>
GPU	<i>Graphics Processing Unit</i>
CPU	<i>Central Processing Unit</i>
IA	<i>Intelligence Artificielle</i>
IDE	<i>Integrated Development Environment</i>

Tableau 1 : Liste d'Acronymes et Abréviations

Introduction générale

Aujourd'hui, nous sommes face à une augmentation du nombre de malvoyants dans la population de par son vieillissement. Les méthodes de numérisation et de reconstruction 3D ont notamment connu des évolutions importantes avec le développement des scanners laser 3D et de logiciels de reconstruction 3D. Il devient important de repenser le patrimoine architectural, pour qu'il soit accessible par tous, aussi bien spatialement que pour les publics empêchés.

C'est à partir de ces constats que les membres de l'équipe PR du laboratoire MIS d'Amiens ont initié le projet de recherche PHARES (**P**ercevoir grâce à une interface **H**Aptique pour **L**es malvoyant**S**). Le projet PHARES vise à répondre à la problématique suivante : « Comment mettre à disposition des malvoyants le patrimoine 3D par une expérience multisensorielle qui diffère des expériences actuelles ? ». C'est dans ce cadre s'inscrivent les travaux réalisés lors de ce projet de fin d'études, qui a pour but de contribuer à une partie de cette problématique, en implémentant le modèle usager de l'interface que nous souhaitons réalisée.

Le présent rapport décrit l'essentiel du travail réalisé lors de ce projet, il est organisé en quatre chapitres :

Dans le premier chapitre, nous présentons l'organisme d'accueil, après nous présentons le sujet de recherche auquel notre stage appartient ainsi que la problématique traitée. Après, nous exposons les missions confiées et la démarche suivie pour la gestion du projet.

Dans le deuxième chapitre, nous présentons le système visuel humain ainsi que des descriptions de quelques pathologies. Ensuite, nous exposons un état de l'art des méthodes existantes d'acquisition 3D et quelques techniques de traitement de données.

Dans l'avant-dernier chapitre, analyse et conception, nous exposons les acteurs du système, les cas d'utilisation, les diagrammes de séquence et l'infrastructure matérielle générale du système.

Le dernier chapitre est consacré pour présenter la mise en œuvre du projet. Il décrit les outils et les technologies utilisés, ainsi que les interfaces graphiques de l'application.

Chapitre 1

Cadre général du projet

Introduction

Le premier chapitre de ce rapport a pour objectif de situer notre travail dans son cadre général. Nous commençons par la présentation de l'organisme d'accueil, son organisation ainsi que l'équipe au sein de laquelle le stage a été effectué. Ensuite, nous présentons également le projet de recherche auquel appartient le sujet du stage. Enfin, nous passons à la problématique, les objectifs et les missions confiées.

1. Présentation du laboratoire d'accueil

1.1. Laboratoire MIS

Créé en 2008, le laboratoire Modélisation, Information & Systèmes (MIS) d'Amiens, dépendant de l'Université de Picardie Jules Verne (UPJV), son directeur est actuellement M. Gilles DEQUEN et dirigé par un conseil de laboratoire. Il rassemble des enseignants-chercheurs de l'UPJV en informatique, automatique, robotique et vision par ordinateur. Les travaux du MIS portent sur des thématiques appartenant aux Sciences et Technologies de l'Information et de la Communication (STIC).

Le MIS est composé de 80 personnes, dont 40 enseignants-chercheurs répartis en quatre équipes de recherche :

- Commande et Véhicules (COVE),
- Graphes, Optimisation et Contraintes (GOC),
- Systèmes Distribués, Mots et Applications (SDMA),
- Perception Robotique (PR), et c'est au sein de laquelle nous avons effectué ce stage.



Figure 1: Organigramme du laboratoire MIS

1.2. Equipe PR

Ce projet de fin d'études a été réalisé au sein de l'équipe PR (Perception Robotique). Son responsable est M. Pascal VASSEUR. Cette équipe travaille sur différents projets dans les thématiques suivantes :

- Localisation et navigation de robots,
- E-santé,
- Contrôle qualité,
- Cartographie du patrimoine monumental.

L'équipe a initié le projet E-Cathédra@le qui appartient à l'axe e-Patrimoine qui est un programme de recherche et de cartographie du patrimoine monumental soutenu par la région Haut de France, Amiens métropole, le Centre des Monuments Nationaux (CMN) et de la Direction Régional des Affaires Culturelles (DRAC) de Picardie. Ce programme vise entre autres à réaliser un modèle numérique 3D dense de la cathédrale d'Amiens.

En outre, l'équipe traite des problématiques autour de la vision non conventionnelle.

2. Projet PHARES

Le sujet de ce stage s'inscrit dans le cadre du projet de thèse de doctorat de Mme. Audrey AMBLES intitulé **PHARES** (Percevoir grâce à une interface **HA**ptique pour **IEs** malvoyant**S**) et dans le programme de recherche E-Cathédra@le. Ce projet vise à faciliter l'accès au patrimoine numérique pour les malvoyants et leur permettre d'appréhender les modèles numériques 3D qui sont actuellement restitués sur des ordinateurs ou dans des systèmes d'immersion. Pour cela il s'agit, en partant des besoins des malvoyants, de proposer dans une démarche de conception, les critères d'une nouvelle interface multi-sensorielle dédiée aux malvoyants. Cette proposition aboutira à la conception d'un prototype haptique. Ainsi, il tente d'apporter une réponse à la problématique suivante : « Comment mettre à disposition des malvoyants le patrimoine 3D par une expérience multi-sensorielle qui diffère des expériences actuelles ? »

3. Problématique

La déficience visuelle touche aujourd'hui plus de 315 millions de personnes à travers le monde, un chiffre qui pourrait doubler d'ici 2030 du fait du vieillissement de la population. Pour compenser la perte ou l'absence de vision, il existe deux approches holistiques : les systèmes de substitution sensorielle visant à restituer l'information visuelle par d'autres modalités sensorielles, principalement l'audition ou le toucher et les neuro-prothèses visuelles. Face à cette population grandissante, de plus en plus d'institutions culturelles, musées et monuments, proposent des services spécifiques pour ce public empêché.

Ces dernières années, divers projets ont été consacrés à la création d'expériences multisensorielles associant principalement des explorations tactiles à des informations auditives. Ces diverses expériences multisensorielles utilisent le plus souvent des objets 3D fabriqués et du contenu multimédia (par des auto-description vocales ou braille) pour faciliter l'interprétation des malvoyants et sont généralement localisés dans des salles tactiles. Ils constituent des reproductions de sculptures, souvent en nombre limité. Il existe également des expériences où des techniques de mise en relief de supports culturels 2D sont utilisées, qui utilisent diverses méthodes pour extraire les contours. Cependant, ces méthodes génèrent trop d'informations qui peuvent dérouter un utilisateur malvoyant. Enfin, notons également des projets de prise en compte du contexte, qui permettent d'ajuster les recommandations pour guider les malvoyants.

D'un autre côté, les projets de numérisation 3D du patrimoine se multiplient et aboutissent à l'élaboration de corpus importants et de très grande qualité. Ces derniers ouvrent des perspectives de recherche et de médiation prometteuses, offrant notamment des possibilités

de visite virtuelle et de découverte intéressante. Malheureusement, les malvoyants n'ont actuellement pas les moyens d'en bénéficier.

Pour relever les défis scientifiques de demain, l'objectif de cette recherche vise à concevoir un dispositif permettant aux malvoyants d'accéder à ces modèles 3D par une expérience multisensorielle qui diffère des expériences actuelles : le projet PHARES (Percevoir grâce à une interface HAptique pour lEs malvoyantS). Mais alors, Comment mettre à disposition des malvoyants ce patrimoine 3D par une expérience multisensorielle qui diffère des expériences actuelles ?

En proposant le dispositif *LISA (mulTi-sensoriel pour leS MalvoyAnts)* qui se place en rupture des dispositifs déjà existants.

4. Cadre du projet

Le stage s'inscrit dans une démarche exploratoire de recherche et développement qui a pour mission de proposer un dispositif aux malvoyants afin de leur permettre d'accéder aux numérisations 3D du patrimoine.

4.1. Missions

Les tâches demandées dans ce stage se divisent en deux axes principaux :

Premièrement, il est demandé de réaliser une interface graphique sous Unity permettant de simuler visuellement les pathologies des malvoyants possédant une déficience visuelle en pouvant agir sur des objets 3D. Cette interface doit être différente des simulateurs courants qui utilisent en entrée une image (Figure 2) de la caméra et génèrent une image en sortie avec un filtre appliqué (Figure 3).

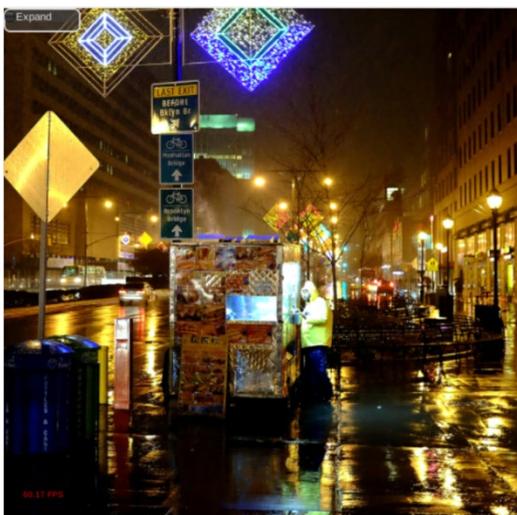


Figure 2: Image originale

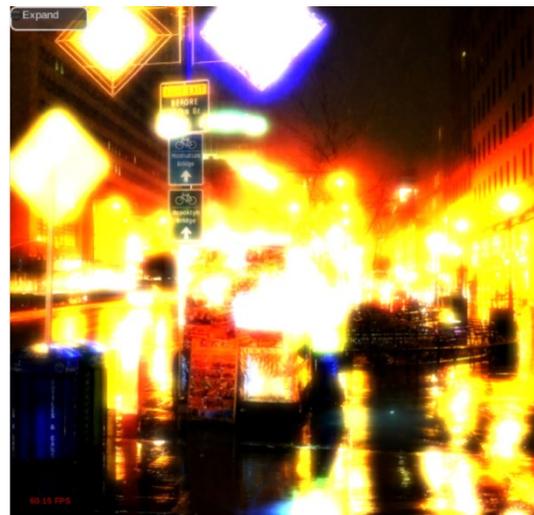


Figure 3: Image de sortie après application du filtre "Glare"

Deuxièmement, contribuer à une partie de la problématique du sujet PHARES, de ce fait, proposer un modèle conceptuel d'une interface multisensorielle destinée aux malvoyants afin de leur permettre d'accéder aux numérisations 3D du patrimoine.

En effet, notre travail est principalement focalisé sur cette deuxième partie, plus précisément, essayer d'apporter des réponses et proposer des solutions aux problématiques suivantes :

- Comment segmenter et annoter les numérisations 3D réalisées par le laboratoire qui sont présentées sous forme des nuages de points 3D ?
- Comment manipuler les modèles 3D afin de pouvoir les utiliser sous Unity ?
- Proposer une conception d'une interface multisensorielle générique (scalable) qui sera basée sur d'autres sens (le toucher, l'ouïe, l'odorat) en plus de la vue, dans le but de maximiser la vue des malvoyants pour éprouver une expérience multisensorielle optimale.

4.2. Conduite du projet

Cette section est consacrée à la présentation du processus de développement adopté pour la planification et la réalisation du projet.

4.2.1. Processus de développement

Nous avons choisi pour la gestion de notre projet la méthodologie eXtreme Programming (XP). Elle est une des méthodologies agiles les plus connues et elle regroupe de bonnes pratiques de développement découlant des principes de l'agilité et les poussent à l'extrême. Elle repose sur les quatre principes du développement agile : la communication, le feedback, la simplicité et le courage.

L'utilisation de cette méthode est dû à la nature des tâches qui ont été réalisés, notamment les tests et les essais de l'interface. Ainsi que cette méthode convient parfaitement les équipes de petite taille (5 personnes dans notre cas).

Le cycle de développement de la méthodologie XP se base sur une approche des itérations fréquentes, notamment, nous commençons par sélectionner le scénario à réaliser à partir des besoins, ensuite nous définissons et répartissons les tâches, après, nous planifions le développement et les tests et finalement, nous fournissons une application exécutable. Après l'avoir évaluée, nous passons à l'incrément suivant et nous refaisons le processus comme décrit le schéma ci-dessous (Figure 4).

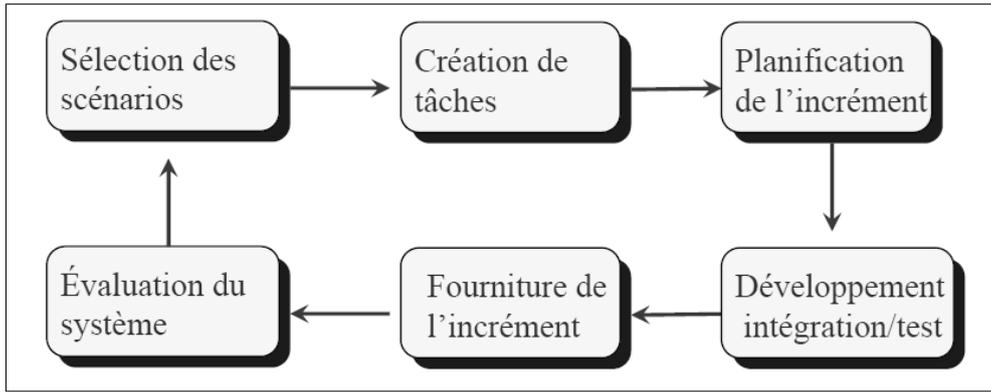


Figure 4: Cycle de développement XP

En dehors des outils agiles, nous nous sommes appuyés sur un outil de gestion de projet plus traditionnel, comme avec un plan de communication :

	Entre le stagiaire et l'équipe du laboratoire
Qui	Hamza ASBAYO (stagiaire) Dominique GROUX (encadrante) Alexis Potelle (encadrant) Audrey AMBLES (doctorante)
Quoi	Discussions sur les travaux effectués, les travaux restants, les travaux futurs, questions sur le sujet ou une notion particulière, préconisations, réunions de travail
Pourquoi	Pour informer de l'avancement des travaux, présenter les productions, faire des préconisations, échanger sur les résultats obtenus et définir les axes à suivre pour la suite des travaux
Fréquence	Réunions hebdomadaires : fin de semaine pour faire un point sur les travaux effectués lors de la semaine Sans fréquences particulières pour les autres communications
Canal	Mails, visio via zoom

Tableau 2: Plan de communication entre les acteurs du stage

4.2.2. Planification

Le planning est une activité qui va nous aider à déterminer et ordonnancer les tâches de notre projet, à estimer leurs charges et à déterminer les profils nécessaires à leur réalisation, il va nous permettre également de suivre et de communiquer l'avancement du projet en déterminant si les objectifs sont réalisés ou dépassés.

Diagramme de Gantt

Le diagramme de Gantt, couramment utilisé en gestion de projet, est l'un des outils les plus efficaces pour représenter visuellement l'état d'avancement des différentes activités (tâches) qui constituent un projet.

La figure suivante résume la planification réelle du projet :



Figure 5: Diagramme de GANTT

Conclusion

Dans ce chapitre, nous avons présenté le cadre général du projet, en mettant en valeur les problématiques à résoudre et les questions à répondre, ainsi que la démarche adoptée pour la gestion du projet.

Chapitre 2

Contexte et état de l'art

Introduction

Ce chapitre est divisé en deux parties, la première a pour objet de présenter l'organisation du système visuel humain afin de mieux comprendre les pathologies des malvoyants possédant une déficience visuelle. La deuxième partie dresse un état de l'art des méthodes et outils existants d'acquisition 3D du patrimoine ainsi que les types et les techniques de la segmentation et l'annotation sémantique du relevé numérique.

1. Système visuel humain

1.1. L'œil

L'œil est l'organe sensoriel de la vision. Premier élément de la chaîne de traitement de l'information visuelle, sa fonction principale est de focaliser la lumière incidente pour former une image sur la rétine, puis de convertir ces ondes électromagnétiques en influx nerveux, qui seront transmis par les nerfs optiques au cortex visuel, qui nous renvoie l'image traitée et permet ainsi l'interprétation de notre environnement.

La vision humaine est la principale source d'informations du cerveau provenant du monde extérieur : environ 70 % des informations sont transmises par la vision. Le champ de vision humain est d'environ 220° horizontalement et 140° verticalement [1].

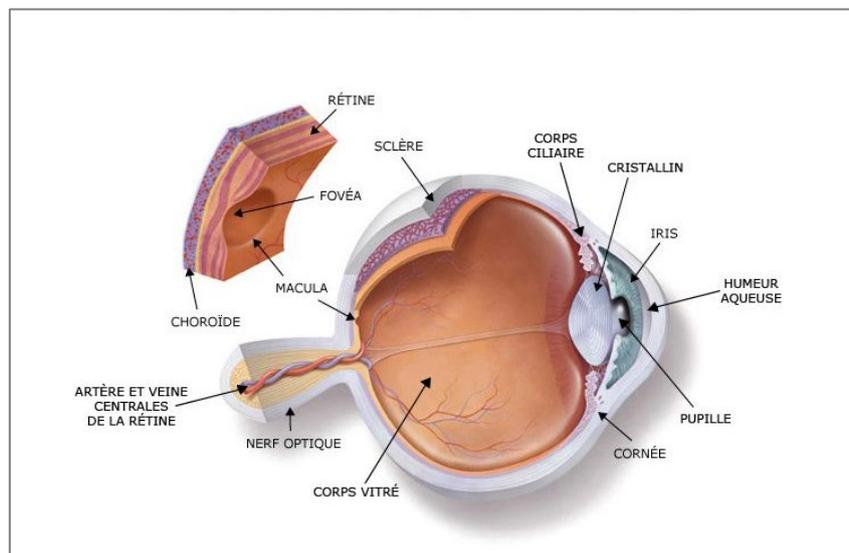


Figure 6: Schéma en coupe de l'œil humain [1]

L'œil peut se définir comme une caméra, la plus complexe et la plus aboutie qui existe. Dans cet organe ultra-perfectionné, chaque élément a son rôle et son importance (Figure 6).

1.1.1. La cornée

La cornée est la partie antérieure de l'œil. Elle constitue le premier élément optique qui permet de focaliser les signaux lumineux sur la rétine. Elle agit comme une fenêtre à travers laquelle la lumière pénètre dans l'œil. En effet, la cornée est une membrane fine et fragile ainsi elle doit toujours être parfaitement propre et transparente. Sa puissance optique (vergence) ne varie pas.

1.1.2. L'iris

L'iris est situé entre la cornée et le cristallin. Il est une membrane contractile ayant la capacité de modifier le diamètre de l'ouverture en son centre, appelée pupille, afin de réguler la quantité de lumière pénétrant dans l'œil en fonction de la luminosité ambiante et de réduire les aberrations optiques pour la formation d'une image rétinienne nette. Il est quasiment collé à la face avant du cristallin.

1.1.3. Le cristallin

Le cristallin est une lentille représentant le deuxième élément optique de l'œil. Son rôle est de focaliser l'image des signaux optiques issus de la cornée sur la rétine pour la rendre nette quels que soient leurs distances de provenance. Il a une puissance optique (vergence) variable. L'adaptation de la vergence du cristallin est réalisée par un phénomène appelé l'accommodation. C'est la contraction des muscles fixant celui-ci dans l'œil. Cela a pour effet de modifier les courbures de sa face avant de sa face arrière. En vieillissant, le cristallin perd de sa plasticité, devient un peu plus dur et les muscles qui le contrôlent deviennent moins actifs. Il peut aussi s'opacifier et sa capacité d'accommodation baisse, dans ce cas, on parle de cataracte, la pathologie représentant la première cause de cécité à l'échelle mondiale, que nous allons voir en détail par la suite.

1.1.4. La rétine

La rétine est la couche photosensible de l'œil d'un quart millimètre d'épaisseur, placée au fond de l'œil, sur sa paroi interne. C'est la membrane qui permet la phototransduction : la conversion de l'information lumineuse en signal électrique par un ensemble d'évènements biochimiques dans les photorécepteurs. La rétine est composée de centaines de millions de cellules nerveuses photoréceptrices : 6 à 7 millions de cônes et 130 millions de bâtonnets. Ces deux types de cellules ont des structures très proches (Figure 7). Les cônes permettent la vision des couleurs tandis que les bâtonnets permettent d'interpréter les signaux de faible luminosité. Ces cellules transforment l'information lumineuse en courants ioniques qui se transmettent aux cellules bipolaires et horizontales. Les cellules bipolaires propagent ces potentiels jusqu'aux cellules ganglionnaires. L'influx nerveux est généré dans les cellules ganglionnaires et se propage via leurs axones dont le rassemblement forme le nerf optique [2, p. 16].

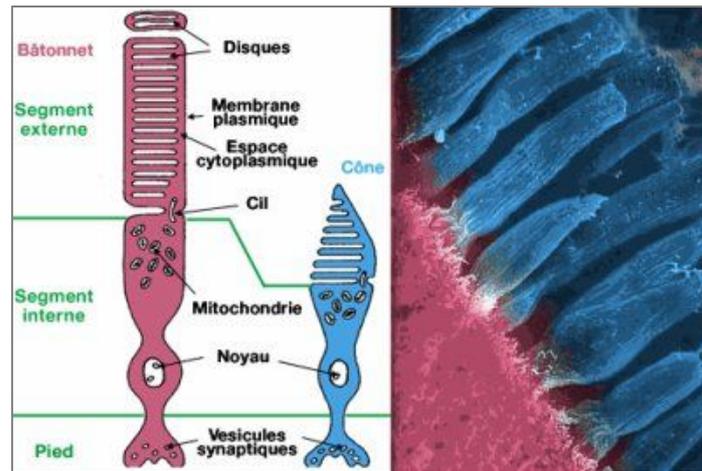


Figure 7: Schéma en coupe de la rétine humaine [3]

2. Les déficiences visuelles

Avant de présenter un certain nombre de troubles de la vision, il est important de présenter d'abord quelques critères utilisés couramment par les opticiens et les ophtalmologues, permettant de qualifier la vision d'une personne. Ils sont généralement utilisés dans l'évaluation de la déficience visuelle : l'acuité et le champ visuel.

L'acuité visuelle mesure le sens morphoscopique, c'est-à-dire la capacité de l'œil à voir distinctement les détails d'un objet dans le but de le reconnaître (à l'échelle décimale, une acuité de 10/10^{ème} représente une vision correcte), alors que l'examen du champ visuel évalue la portion de l'espace vue par l'œil fixe regardant face à soi (pour un œil sain, le champ visuel normal, par rapport à l'axe optique de l'œil, est en moyenne de 60° vers le nez, 100° vers la tempe, 60° vers le haut et 75° vers le bas (Figure 8)).

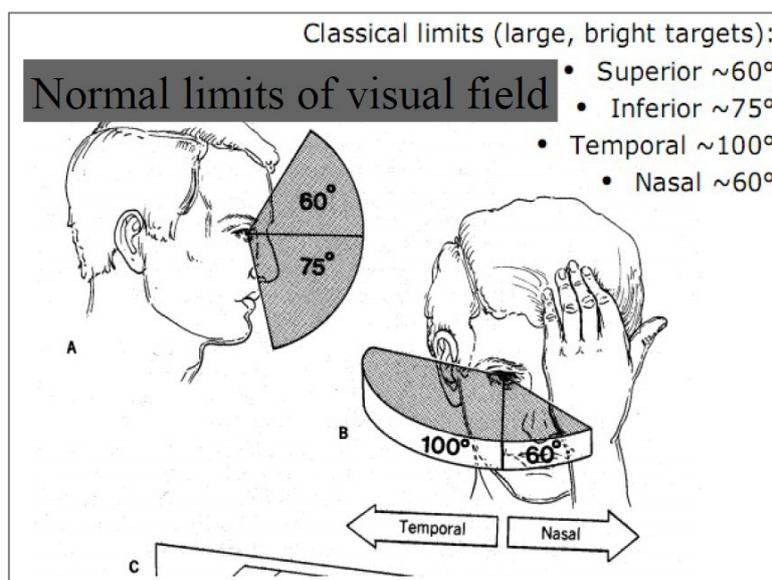


Figure 8: les limites normales du champ visuel [4, p. 7]

2.1. Classifications des déficiences visuelles

Selon la dernière classification internationale des maladies (2018), l'Organisation Mondiale de la Santé (OMS) distingue deux types des déficiences visuelles, selon que la vision de loin ou la vision de près est affectée.

Les différents types et niveaux des déficiences visuelles sont basés sur la mesure de l'acuité visuelle et du champ visuel comme indiqué ci-dessous :

Déficiência de l'acuité visuelle de loin	
Niveau	Acuité visuelle
Légère	6/12 mais plus grande ou égale à 6/18
Modérée	Inférieure à 6/18
Sévère	Inférieure à 6/60
Cécité	Inférieure à 3/60, 1/60 ; ou aucune perception de la lumière
Déficiência de l'acuité visuelle de près	
Acuité visuelle	
Inférieure à N6 ¹ à 40 cm	

Tableau 3: Les différents types et niveaux de déficiences visuelles

L'expérience personnelle de la déficiences visuelles dépend de nombreux facteurs différents, tels que la disponibilité d'interventions préventives et thérapeutiques, la disponibilité de rééducation visuelle (y compris des produits auxiliaires tels que des lunettes ou des bâtons de marche blancs), et tout problème rencontré en raison de l'inaccessibilité des transports et de l'information [5].

2.2. Les causes

À l'échelle mondiale, au moins 2,2 milliards de personnes ont une déficiences visuelles touchant la vision de près ou la vision de loin. Pour au moins 1 milliard de ces personnes, soit près de la moitié d'entre elles, la déficiences visuelles aurait pu être évitée ou n'a pas encore été prise en charge. Ce milliard de personnes comprend les personnes atteintes d'une déficiences modérée ou sévère ou d'une cécité due à un défaut de réfraction non traité (88,4 millions), à la cataracte (94 millions), au glaucome (7,7 millions), aux opacités de la cornée (4,2 millions), à la rétinopathie diabétique (3,9 millions) et au trachome (2 millions), ainsi que celles atteintes d'une déficiences visuelles de près causée par la presbytie non prise en charge (826 millions) [6].

¹ La vision de près "normale" est appelée N6, le "N" faisant référence à la vision de près « Near » et le "6" à la taille des lettres ; N5 est meilleur que N8.

Au niveau mondial, les principales causes de déficience visuelle sont les suivantes :

- Les troubles de la réfraction non corrigés ;
- La cataracte ;
- La dégénérescence maculaire liée à l'âge (DMLA) ;
- Le glaucome ;
- La rétinopathie diabétique ;
- L'opacité cornéenne ;
- Le trachome.

2.3. Description des pathologies

Nous pouvons constater qu'il existe plusieurs pathologies pouvant affecter la vision. Dans cette section, nous nous limiterons à présenter une description succincte des principales pathologies responsables de malvoyance. La liste présentée ici, n'est pas exhaustive. On peut citer :

- La cataracte
- Le glaucome
- La dégénérescence maculaire liée à l'âge (DMLA)

2.3.1. La cataracte

La cataracte est la pathologie provoquant le plus grand nombre de cécités dans le monde. Elle correspond à une opacification du cristallin qui entraîne une baisse progressive de l'acuité visuelle, jusqu'à la cécité complète si elle n'est pas traitée. Elle se développe généralement avec le vieillissement (elle touche une personne sur cinq à partir de 65 ans).

Le cristallin permet de fixer l'image sur la rétine lorsqu'il est normal. Sinon, s'il est « embué » à cause d'une cataracte, l'image devient « embrouillée » (Figure 9).

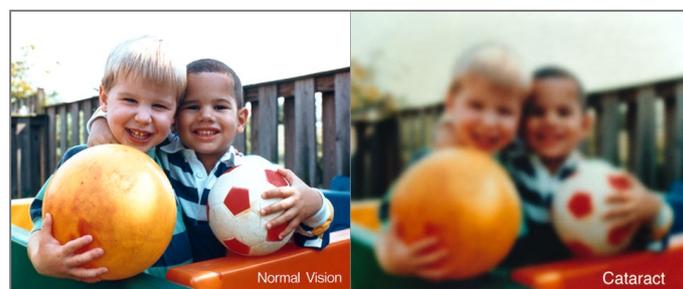


Figure 9: Simulation d'une vision normale; et d'une vision dans le cas de la cataracte [7]

La cataracte peut aussi avoir comme symptômes :

- Une vision floue, voilée ou brouillée ;
- Des petites taches ou des points dans le champ de vision ;
- Une diminution de l'acuité visuelle ;
- Une mauvaise perception des contrastes ;
- Des difficultés à distinguer les reliefs ;
- Des difficultés à différencier certaines couleurs.

2.3.2. Le glaucome

Le glaucome est une maladie dégénérative du nerf optique qui entraîne une perte progressive et irréversible du champ visuel commençant tout d'abord en périphérie et progressant graduellement vers le centre. Cela va occasionner des troubles et des pressions (Figure 10) réduisant la quantité d'informations transmises au niveau de la papille optique, zone de la rétine dépourvue de photorécepteurs qui permet la connexion du nerf optique. Le glaucome engendre une baisse de l'acuité visuelle s'il n'est pas traité rapidement. Il peut ainsi causer les symptômes suivants :

- Un brouillard visuel intermittent ;
- Une vision réduite, surtout sur les côtés. La vision centrale est normale, mais l'environnement qui l'entoure est perturbé (Figure 11).



Figure 10: Schéma d'un œil atteint de glaucome [8]

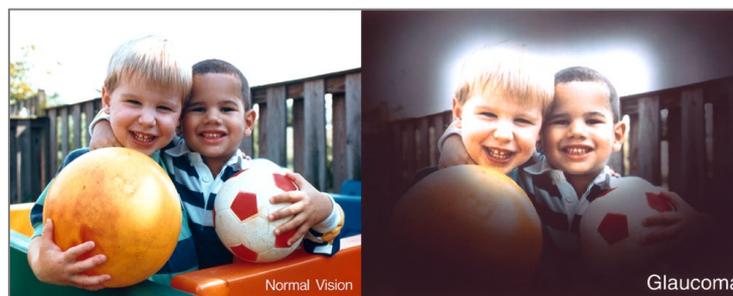


Figure 11: Simulation d'une vision normale; et d'une vision dans le cas du glaucome [7]

2.3.3. La dégénérescence maculaire liée à l'âge

La dégénérescence maculaire liée à l'âge (DMLA) représente la première cause de malvoyance chez les personnes âgées, elle touche près de 10 % des personnes de 65-75 ans et 25 à 30 % des plus de 75 ans. Elle résulte de la dégradation de la macula, zone centrale de la rétine de l'œil où l'acuité visuelle est maximale. Cette zone permet la vision précise des détails et des couleurs. La DMLA entraîne d'abord des gênes à la lecture, la reconnaissance des visages ou à la conduite, et peut évoluer vers des déficiences visuelles plus graves allant jusqu'à la cécité. Cette pathologie affecte le champ visuel avec un scotome, zone peu ou pas fonctionnelle dans le champ visuel (Figure 12).



Figure 12: Simulation d'une vision normale (à gauche) et d'une vision dans le cas de la DMLA (à droite) [7]

3. Relevé numérique du patrimoine

Dans la deuxième partie de ce chapitre, nous allons présenter l'état de l'art des méthodes et outils existants d'acquisition 3D du patrimoine. Ensuite, nous présenterons les types et les techniques utilisés dans la segmentation et l'annotation du relevé numérique qui est présenté généralement sous la forme d'un nuage de points.

3.1. Outils et méthodes d'acquisition 3D

Les dispositifs existants d'acquisition des données 3D sont généralement divisées en deux catégories : avec et sans contact d'après Jacquot [9, p. 30]. Elles sont majoritairement basées sur l'exploitation du spectre électromagnétique et classées en deux sous-catégories : les techniques actives et passives (Figure 13).

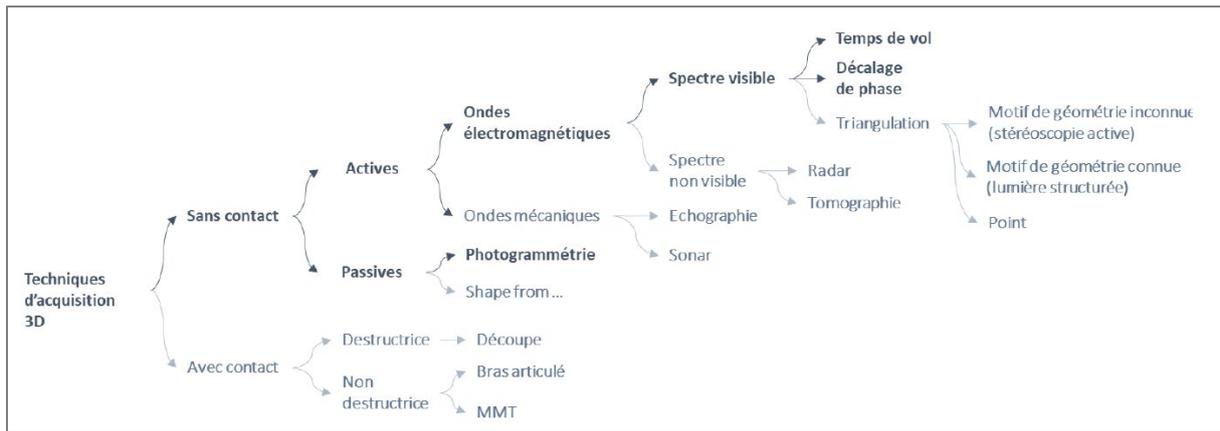


Figure 13: Classification des principales techniques d'acquisition 3D [10]

Les technologies actives se basent généralement sur l'utilisation d'ondes électromagnétiques du spectre visible (lasergrammétrie) ou non visible (tomographie, radar, etc.), mais aussi sur l'utilisation d'ondes mécaniques (sonar, échographie, etc.). La numérisation active tridimensionnelle appliquée au relevé du patrimoine architectural est principalement basée sur la lasergrammétrie.

Contrairement aux techniques actives qui émettent un signal spécifique, les techniques passives se basent sur l'étude du rayonnement ambiant naturel [9]. Parmi les techniques passives, la photogrammétrie est de loin la plus répandue, nous ne détaillerons donc pas les autres dans ce document. Par ailleurs, dans le cas du patrimoine architectural, il est évident que les techniques d'acquisition destructives sont absolument inenvisageables, ainsi que les techniques d'acquisition avec contact sont peu adaptées du fait de la fragilité et de la valeur historique des monuments que nous étudions, nous ne nous attarderons donc pas sur ces systèmes.

Dans cette section, seules les techniques sans contact sont abordées et plus particulièrement la photogrammétrie et la lasergrammétrie, qui sont utilisées dans le cadre du programme e-cathédrale du MIS pour élaborer des modèles 3D. Sachant que le laser scanner est beaucoup plus utilisé car il nous permet d'obtenir des modèles 3D denses.

3.1.1. Photogrammétrie

La photogrammétrie est une technique qui permet d'obtenir des informations fiables sur la forme et les dimensions d'un objet physique par l'enregistrement, la mesure et l'interprétation d'images photographiques. Elle est un moyen dont les fondements remontent à plusieurs siècles.

D'une façon simple, la méthode s'inspire directement de la vision humaine : nos deux yeux voient le même sujet sous un « angle différent », c'est à dire avec un décalage latéral de

quelques centimètres. La parallaxe introduite par les deux images permet au cerveau de reconstituer la troisième dimension, ce que nous voyons de la réalité.

La réalisation d'un relevé photogrammétrique s'effectue généralement en deux étapes :

- tout d'abord la prise de vue, à partir d'un avion pour la cartographie ou terrestre pour les relevés d'architecture.
- puis vient la restitution qui va extraire un ensemble de coordonnées dans l'espace à partir des paramètres de la position (coordonnées suivant les trois axes : X, Y, Z) et de l'orientation (angles autour des trois axes : θ, ψ, ω). Le nombre d'images utilisées va déterminer la méthode : mono, stéréo ou multi-images (Figure 14).

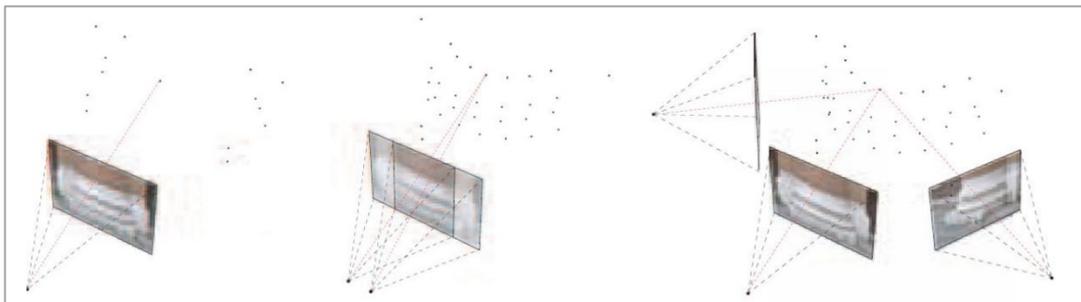


Figure 14: Les différentes techniques de restitution photogrammétrique. De gauche à droite : mono, stéréo, multi-images [11, Fig. 27]

Les progrès réalisés ces dernières décennies dans le domaine de la vision par ordinateur permettent aujourd'hui l'automatisation de l'ensemble de la chaîne de traitement. Le processus général est détaillé par Pierrot-Deseilligny et al. [12].

Les données générées sont stockées sous la forme de cartes de profondeurs qui sont généralement exploitées par la suite pour produire des nuages de points denses dont la densité est liée à la nature de la scène (nombre de points d'intérêts pouvant être appariés entre les images). Le fichier de points est récupérable de façon brut par exportation dans un quelconque logiciel 3D (Dessin assisté par ordinateur) ou dans un système d'information géographique (Figure 15).

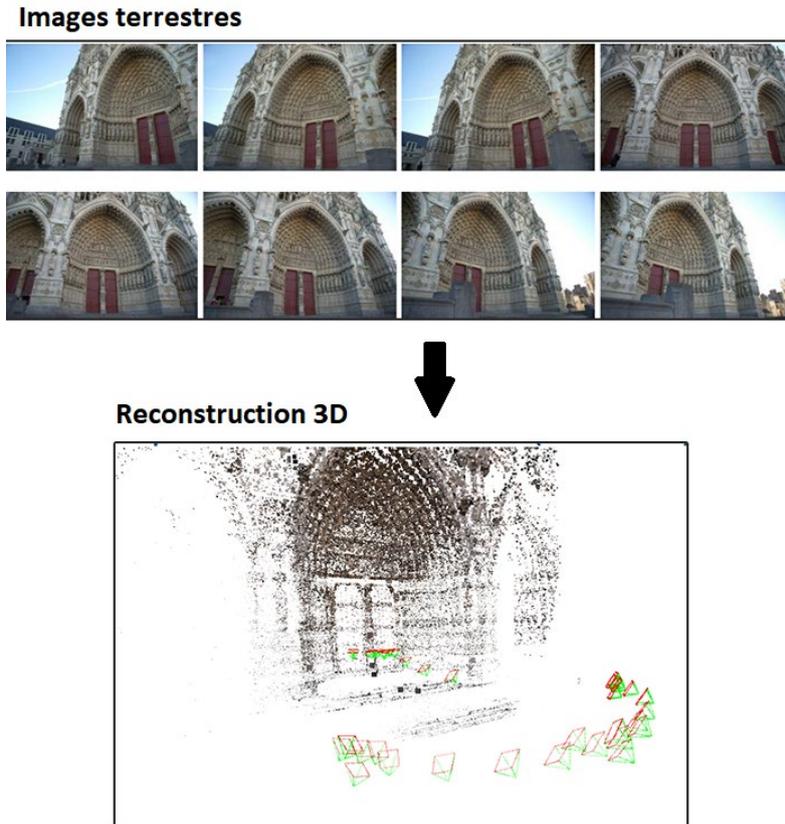


Figure 15: Reconstruction 3D à partir d'un ensemble d'images terrestres d'une partie de la cathédrale d'Amiens (E-Cathedrale (MIS, ENSG, IGN), 2011)

Aujourd'hui, la photogrammétrie peut proposer des résultats dont la qualité est proche de ceux obtenus par lasergrammétrie (Remondino et al. [13]). À la différence de cette dernière, la photogrammétrie offre cependant une grande flexibilité quant aux niveaux de détails, ce qui permet de produire des numérisations 3D multirésolutions.

3.1.2. Lasergrammétrie

La lasergrammétrie ou le relevé par capteur laser est une technique de numérisation 3D sans contact permettant d'obtenir par balayage un ensemble de coordonnées X Y Z des points de la surface de l'objet étudié. Elle est la technique d'acquisition appropriée pour confectionner un modèle numérique de terrain dans les différents domaines de relevés terrestres. Ces dispositifs se sont considérablement améliorés et constituent à nos jours des outils incontournables pour les relevés d'objets patrimoniaux complexes. Ces dernières années ont ainsi vu se multiplier les campagnes de numérisation par scanner laser, contribuant très largement à la production massive de données 3D. Un des avantages immédiats de ce mode de relevé est surtout la précision du nuage de points produit.

Les scanners laser terrestres sont principalement constitués de quatre éléments : une tête laser, un système de balayage permettant de faire varier l'angle du rayon laser, un système récepteur enregistrant le signal retour après réflexion sur l'objet, et un logiciel de pilotage.

Ils se basent sur l'émission d'un signal laser et sa réception par le système récepteur pour déduire la distance séparant l'émetteur du point d'impact sur l'objet étudié. Le système de balayage fait ensuite varier l'angle horizontal du rayon laser pour réaliser la mesure suivante. Lorsque l'appareil a réalisé un tour complet, une variation de l'angle vertical du rayon laser permet de commencer une nouvelle série de mesures à une hauteur différente (Figure 16).

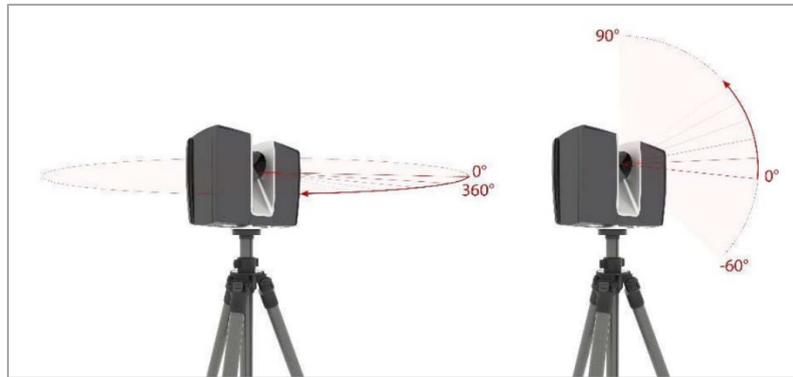


Figure 16: Système de balayage du scanner laser [10, Fig. 27]

Pour conclure, la photogrammétrie et la lasergrammétrie sont aujourd'hui les techniques les plus utilisées pour le relevé permettant d'obtenir les informations nécessaires pour procéder des analyses d'objets patrimoniaux. Plusieurs études ont déjà cherché à comparer les deux techniques. Afin de choisir la méthode à utiliser, les critères cités ci-dessous sont à envisager :

- Facilité de mise en œuvre
- Qualité du nuage de points
- Quantité d'informations
- Variété d'informations

Enfin, ce qui nous intéresse plus particulièrement dans notre travail, ce sont les fichiers résultants de ces processus de numérisation, qui sont représentés sous la forme d'un nuage de points (Figure 17).



Figure 17: Nuage de points représentant un modèle photogrammétrique d'un ange [14]

3.2. Traitement des données et enrichissement sémantique

Quelles que soient les technologies de numérisation que nous avons utilisées, les données résultantes nécessitent dans la plupart du temps des opérations de traitement ou certaines manipulations à l'aide d'outils spécifiques pour permettre aux différents acteurs de les exploiter et d'extraire les informations qu'elles contiennent, avant d'y associer un sens.

Il existe plusieurs méthodes et techniques utilisées soit pour l'extraction d'informations géométriques et visuelles, soit pour l'enrichissement sémantique des données numériques. Dans le cas des images numériques, il peut s'agir d'opérations de retouche, par exemple pour réaliser des transformations géométriques, des modifications de couleurs et de valeurs, des recadrages ou sélections partielles, ou encore l'utilisation de méthodes de traitement d'images plus spécifiques comme l'application de filtres ou d'effets variés. Dans le cas des données 3D, la volonté de les exploiter dans des logiciels tiers peut impliquer notamment des étapes de maillage, de texturage, ou encore d'optimisation.

Dans cette section seules les techniques que nous avons étudiées sont abordées et plus particulièrement, la segmentation rapprochée pour les techniques de traitement de données et l'annotation sémantique pour les méthodes permettant l'enrichissement sémantique des données traitées lors de relevés. En ce qui concerne les autres méthodes et techniques, Abergel V. les présente en détail dans [10, Ch. 2].

3.2.1. Segmentation

L'œil sain d'un humain peut naturellement dissocier des éléments dans une scène, la mise en place des algorithmes permettant d'informatiser ce processus reste encore aujourd'hui l'un des thèmes les plus courants en traitement d'images. La segmentation a pour but de simplifier

et de transformer une image en une représentation significative plus facile à analyser. Pour se faire, la scène est divisée en groupe de pixels (dans le cas d'une image 2D (Figure 18)), ou en maillage de polygones (dans le cas d'un modèle 3D (Figure 19)) homogènes selon certains critères, en formant des régions. Cette méthode peut être utilisée pour atteindre différents objectifs mais dans le cas du relevé numérique du patrimoine architectural, il s'agit principalement de regrouper certains points 3D de la scène pour proposer un découpage susceptible d'être utilisé pour l'annotation sémantique (Section 3.2.2).



Figure 18: Gauche : Image originale. Droite : Image segmentée [10]

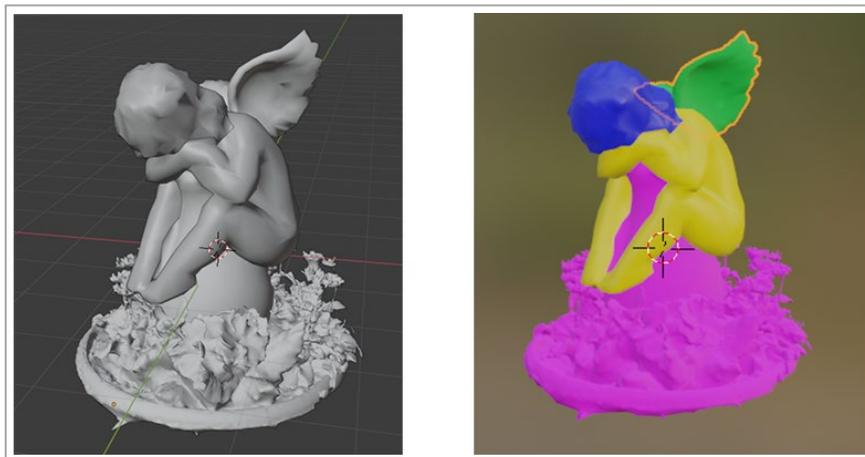


Figure 19: Modèle 3D d'un ange segmenté

Il existe de nombreuses méthodes de segmentation d'images, permettant de faciliter le processus d'extraction d'informations. Elles sont généralement regroupées en quatre catégories :

- Segmentation par région (*region-based*)
- Segmentation par contours (*edge-based*)
- Segmentation par seuillage (*threshold*)

- Segmentations combinant ces trois méthodes.

Cependant, dans le cas d'objets patrimoniaux, l'étape de la segmentation pose beaucoup de problème du fait de l'unicité des objets et de leur complexité. Ainsi, chaque cas d'étude suppose le développement de stratégies adaptées à ses spécificités. Quelques algorithmes existants sont présentés par Manuel A. [11, Chap. 7] qui ne sont généralement applicables qu'à des études de cas spécifiques.

3.2.2. Annotation sémantique

Après avoir appliqué les méthodes et les techniques de traitement des données, il est très important de mettre en relation ces données géométriques et visuelles avec les connaissances propres à l'observateur afin de donner des formalisations sémantiques permettant de mettre en relation les observations avec leurs supports.

Les experts disposent de plusieurs types de représentations qui peuvent être classées selon deux grandes catégories : les représentations 2D et les représentations 3D. Chacune de ces représentations peut être enrichie sémantiquement par un processus d'annotation. Les méthodes d'annotation s'adaptent donc en fonction de la représentation choisie.

Le principe de l'annotation sémantique repose sur la mise en relation entre un document et des informations complémentaires, afin d'établir des relations entre les informations liées aux données plutôt qu'entre les données elles-mêmes. Autrement dit, il s'agit de passer d'une structure de données à une structure de connaissances, afin de permettre efficacement la recherche, le partage, et le croisement d'informations.

Dans le domaine de la conservation du patrimoine, l'annotation sémantique d'image se révèle comme un moyen important pour documenter leur contenu. Ce procédé d'annotation passe par la définition de zones d'intérêt sur l'image auxquelles sont alors associées à une description à l'aide de mots-clés. Elle comprend deux aspects : l'annotation (définition ou sélection spatiale des aspects formels concernés par la description) et la description sémantique (assignation des éléments sélectionnés).

Les annotations peuvent être liées aux images de deux façons :

- **Annotation globale** : l'annotation est associée à l'ensemble du document, elle décrit le contenu du document.
- **Annotation locale** : l'annotation est associée à des zones du document (points, courbes, polygones, etc.), chaque zone est annotée séparément des autres.

Concernant la description sémantique, Andrews et al. [15] distinguent quatre modèles, qui correspondent à différentes méthodes pour associer une description aux éléments annotés :

- **Tags** : Un tag est un mot-clé (un mot ou une séquence de caractères sans espaces) ou un terme assigné à une ressource qui décrit implicitement une propriété particulière de la ressource. L'ordinateur ne connaît pas, dans ce cas, la signification sous-entendue par le commentateur.
- **Attributs** : Un attribut est une paire d'éléments : le nom de l'attribut (définissant la propriété de la ressource) et sa valeur.
- **Relations** : Une relation est une paire de composants : le nom de la relation et la ressource associée. La ressource annotée est donc liée par rapport au nom de la relation. Ainsi les relations sont une extension des attributs pour les ressources.
- **Ontologies** : Une ontologie est un modèle qui consiste à aligner une ressource ou une partie de celle-ci avec une description de certaines de ses propriétés et caractéristiques selon un modèle conceptuel formel (ou ontologie).

Les techniques d'annotations d'images sont classifiées selon trois catégories :

- **Annotations manuelles** : les régions sont tracées manuellement sur chaque image devant faire l'objet d'annotations, et les descriptions sont choisies par l'utilisateur (définies librement ou sélectionnées parmi des classes préétablies).
- **Annotations automatiques** : ces méthodes incluent la sélection automatique des régions d'intérêt sur les images, et leur description. Globalement, il s'agit d'apprendre automatiquement les concepts sémantiques à partir d'un grand nombre d'échantillons d'images, et de les utiliser ensuite pour décrire de nouvelles images.
- **Annotations semi-automatiques** : Ces méthodes sont des combinaisons de méthodes manuelles et de méthodes automatiques. Elles consistent à faire intervenir les utilisateurs pour valider les décisions de la machine, afin de profiter des performances des méthodes automatiques tout en garantissant la pertinence des descriptions réalisées.

Comme pour les approches 2D, des données 3D enrichies sémantiquement permettent d'envisager des recherches exploitant les propriétés des données mais aussi les connaissances qu'elles représentent. Conceptuellement ces approches sont assez semblables. Ainsi, les annotations sémantiques de données 3D peuvent être globales ou locales, et les modèles de descriptions (par tags, attributs, relations, ou ontologies) sont les mêmes. À l'instar des approches 2D, les approches 3D soulèvent donc des questions semblables concernant les points critiques de l'annotation, c'est-à-dire l'efficacité et la précision des outils permettant la sélection de régions, et les modèles de description à privilégier pour garantir une richesse sémantique tout en assurant l'interopérabilité des descripteurs.

3.2.3. Bilan

Vu que notre application est destinée aux personnes qui ont une déficience visuelle et ils n'ont pas la capacité d'acquérir facilement de grandes quantités de données, alors les scènes du patrimoine présentées seront en principe des scènes simples contenant quelques objets. C'est pour cela que nous avons choisi de faire simple et segmenter les scènes à la main. En effet, la segmentation manuelle est celle qui reste la plus précise en termes d'exactitude de dissociation des scènes 3D.

En ce qui concerne l'annotation sémantique, l'utilisation du logiciel de modélisation 3D Blender (Section **Erreur ! Source du renvoi introuvable.**) pour la segmentation nous a facilité la tâche, car il nous a permis d'avoir une scène divisée en objets annotés (*GameObjects*²) que nous pouvons manipuler par la suite sous Unity (Section **Erreur ! Source du renvoi introuvable.**).

Conclusion

Dans ce chapitre, nous avons présenté brièvement l'organisation du système visuel humain ainsi que les pathologies pouvant affecter la vision. Nous avons cité les principaux outils et méthodes existants d'acquisition 3D du patrimoine architectural. En outre, nous avons présenté les opérations de traitement nécessaires à appliquer à ces données résultantes afin de mieux les exploiter.

² Les *GameObjects* sont les objets fondamentaux d'Unity qui représentent les personnages, les accessoires et les décors.

Chapitre 3

Analyse et conception

Introduction

À travers ce chapitre, nous entamons l'analyse et la conception qui présente une étape fondamentale qui précède la réalisation. Nous commençons par un rappel du cadre général ainsi que les missions principales du projet. Nous procédons ensuite à l'identification des acteurs et des cas d'utilisation en les détaillant par des diagrammes de séquence. Et nous terminons par présenter l'infrastructure matérielle et logicielle du système ainsi que l'interaction entre ces composants.

1. Objectifs du projet

Rappelons que notre travail s'inscrit dans le cadre du sujet PHARES, qui a comme objectif de mettre à disposition des malvoyants le patrimoine 3D par une expérience multisensorielle.

La figure 20 ci-dessous représente la modélisation générale du dispositif LISA proposée par Mme. Groux et l'équipe du laboratoire dans le cahier de charge qui m'a été fournis, et à travers cette modélisation, nous pouvons remarquer la complexité du système. Notre travail est principalement focalisé sur les deux modèles : **modèle 3D-Annoté** qui est décrit par des **modalités-sensorielles**.

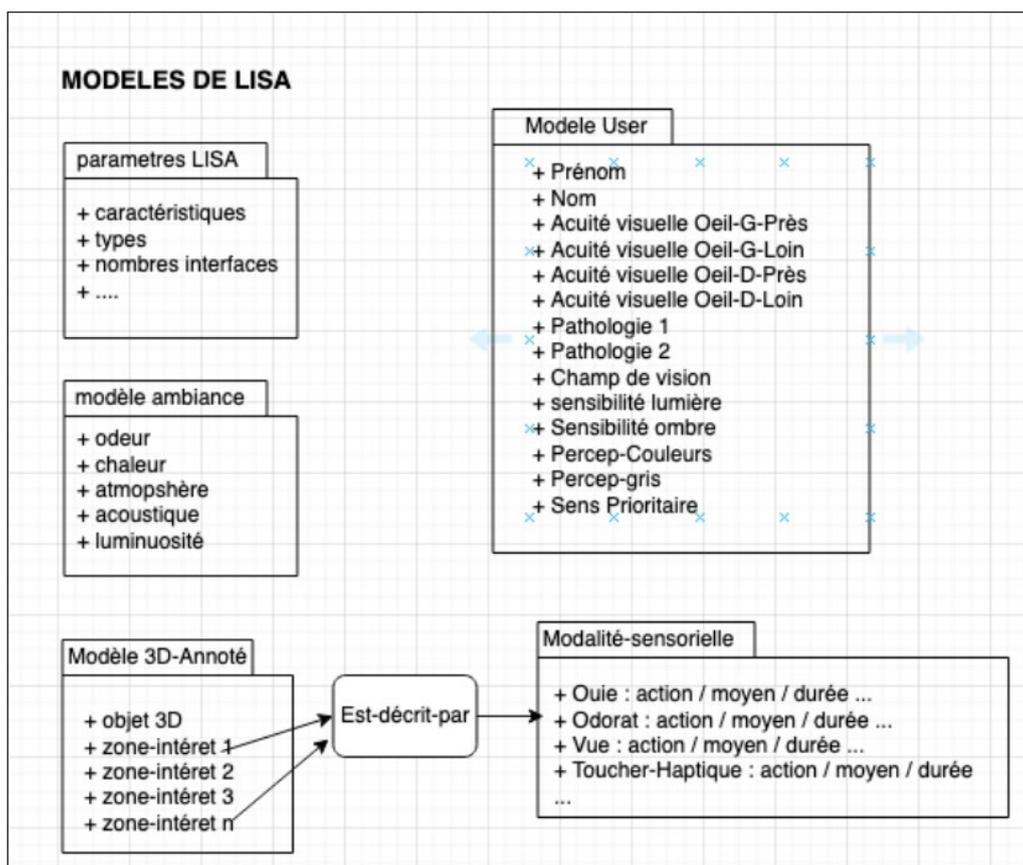


Figure 20: Les modèles du dispositif LISA

Le schéma ci-dessous (Figure 21) représente un exemple d'une instance d'un modèle 3D segmenté-annoté qui est extrait d'un fichier confidentiel fourni par l'équipe du laboratoire. Chaque zone d'intérêt du modèle est décrite par des modalités sensorielles (la vue, l'ouïe, le toucher, l'odorat) que nous souhaitons mettre en œuvre à travers notre système. Dans un premier temps, nous visons à travers ce travail à expliciter le modèle multisensoriel de LISA en se basant sur deux sens : la vue et l'ouïe.

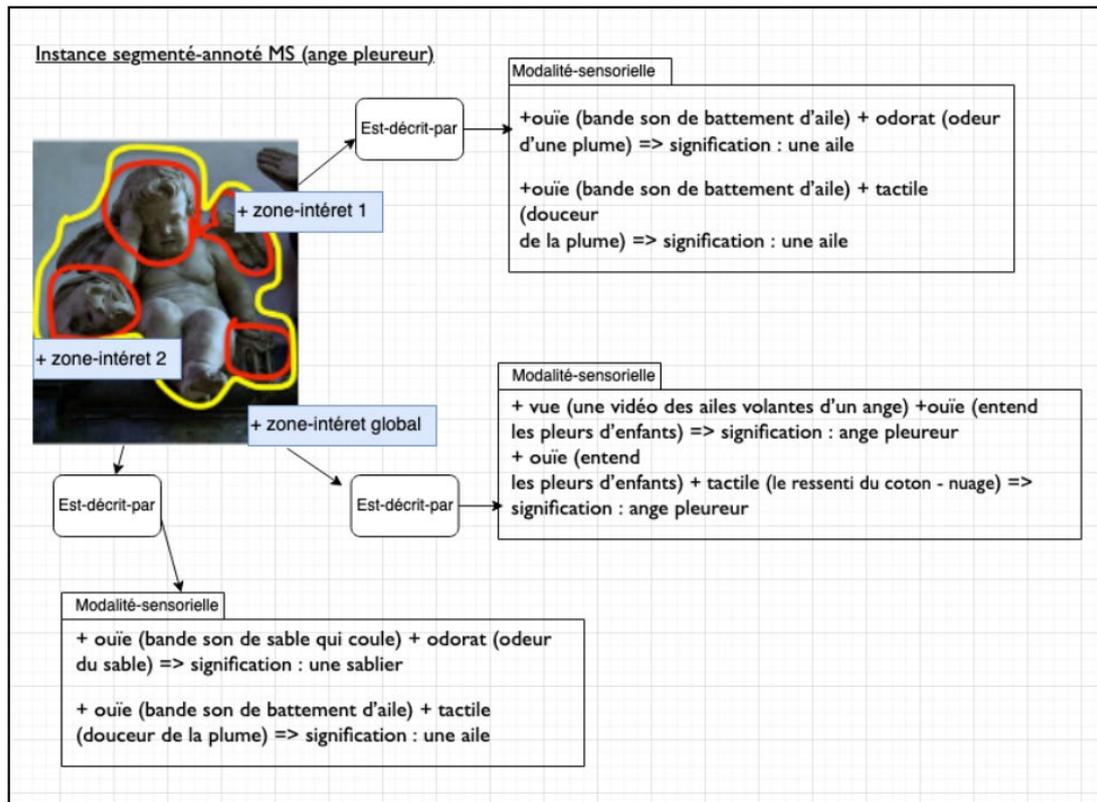


Figure 21: Modèle d'un ange segmenté-annoté (Source : cahier de charge confidentiel)

Dans cette section, nous décrivons notre contribution dans le projet en présentant les missions principales confiées :

- Segmenter et annoter les objets 3D issus des numérisations faites par le laboratoire afin de pouvoir les manipuler sous Unity.
- Créer une interface sous Unity en prenant en considération deux sens (la vue et l'ouïe) permettant de :
 - Détecter la ligne de regard de l'utilisateur, autrement dit, savoir où l'utilisateur regarde exactement dans la scène 3D.
 - Déclencher une action sensorielle autre que la vue en fonction de l'objet entrant en collision avec l'axe de regard de l'utilisateur.

2. Modélisation

Afin de représenter le comportement fonctionnel de notre système, nous utilisons les diagrammes de cas d'utilisation UML.

2.1. Diagramme des cas d'utilisation

Le diagramme de cas d'utilisation décrit le comportement du système du point de vue utilisateur sous forme d'actions et de réactions. Il existe deux concepts fondamentaux dans la modélisation par les cas d'utilisation :

- Les acteurs qui agissent sur le système,
- Les cas d'utilisation qui représentent la façon dont le système est manipulé par les acteurs.

Chaque cas d'utilisation indique une fonctionnalité du système déclenché par un acteur externe au système. Ce diagramme permet de mettre en place et de comprendre les besoins des utilisateurs.

2.1.1. Identification des acteurs

Au niveau de cette partie, nous présentons les acteurs susceptibles d'interagir avec le système, mais tout d'abord, nous donnons une définition du concept acteur.

Acteur : le rôle joué par des entités externes qui interagissent directement avec le système étudié, il peut être un utilisateur, un matériel externe ou un autre système.

Notre système comprend deux acteurs principaux « le malvoyant » et « l'assistant » dans le cas où le malvoyant a une déficience visuelle sévère.

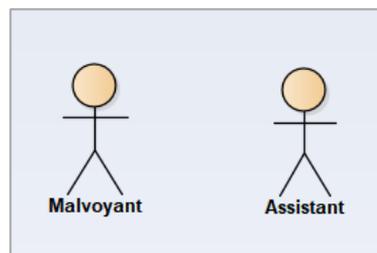


Figure 22: Les acteurs

2.1.2. Diagramme des Cas d'utilisation de l'assistant

Le cas d'utilisation ci-dessous (Figure 23) englobe l'ensemble des fonctionnalités majeures fournies par le système au premier acteur « l'assistant ».

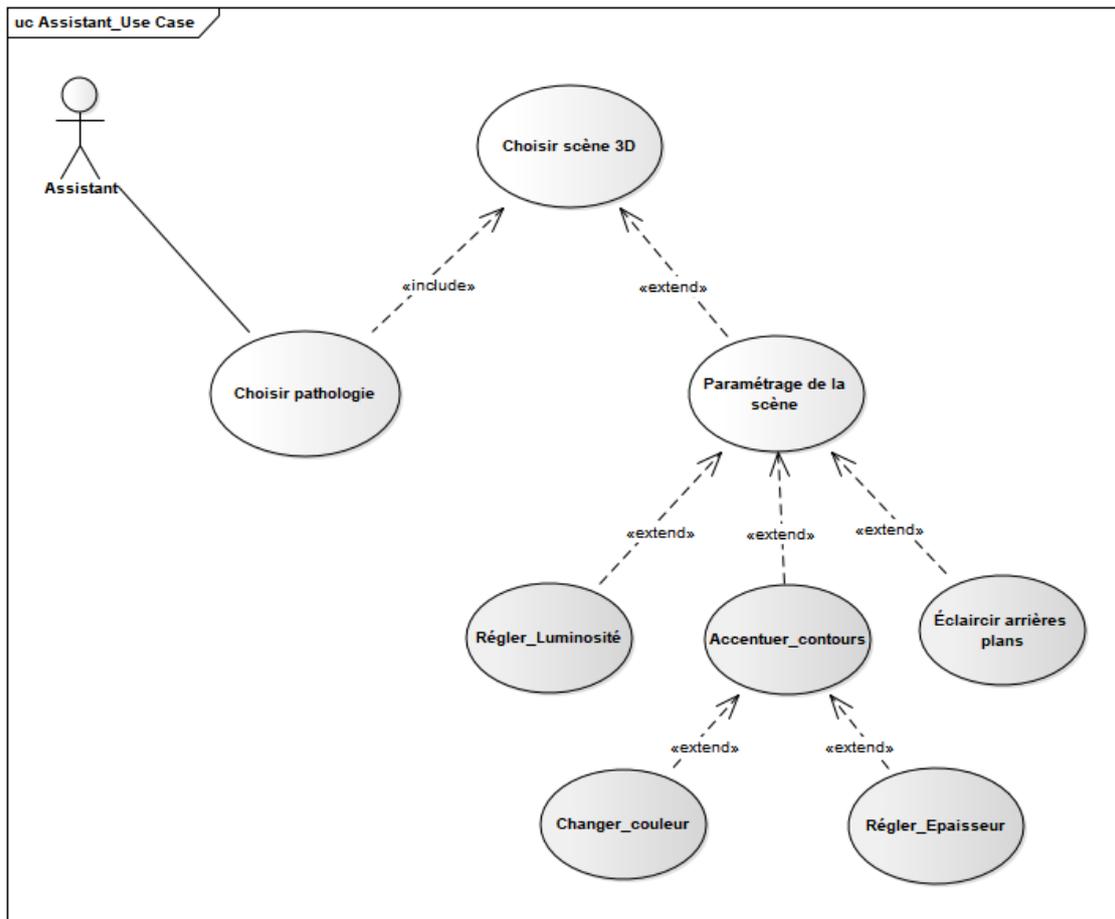


Figure 23: Diagramme des cas d'utilisation de l'assistant

Le tableau 4 décrit les fonctions fournies par notre système à l'acteur « Assistant » :

Cas d'utilisation	Description
Choisir pathologie	L'assistant choisit la ou les pathologies dont le malvoyant souffre pour que le système lui retourne une scène adaptée à son cas.
Choisir scène 3D	L'assistant peut choisir la scène ou l'objet 3D que le malvoyant souhaite explorer.
Paramétrer la scène	L'assistant peut adapter la scène à travers plusieurs paramètres, comme la luminosité et l'éclairage. Il peut aussi accentuer les contours des objets 3D, en modifiant la couleur et l'épaisseur.

Tableau 4: Description des cas d'utilisation de l'assistant

2.1.3. Diagramme des Cas d'utilisation du malvoyant

Le cas d'utilisation ci-dessous (Figure 24) donne une vision globale des fonctionnalités fournies par le système à l'acteur principal « le malvoyant ».

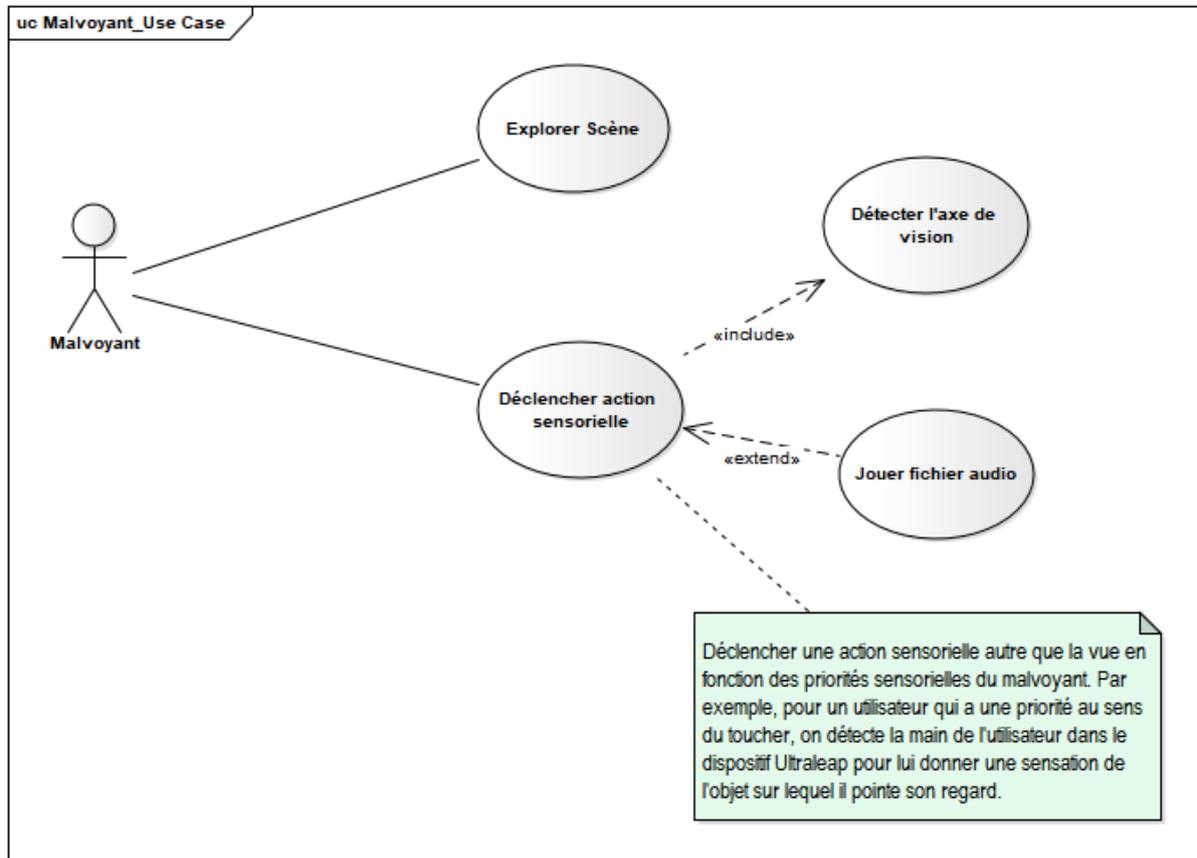


Figure 24: Diagramme des cas d'utilisation du malvoyant

Le tableau suivant décrit les fonctions fournies par notre système à l'acteur « malvoyant » :

Cas d'utilisation	Description
Explorer la scène	Le malvoyant peut naviguer dans la scène 3D virtuelle avec une souris, un clavier et un écran
Déclencher action sensorielle	En temps réel, en détectant l'axe de regard du malvoyant, on déclenche une action sensorielle en fonction des priorités sensorielles de l'utilisateur. Dans notre cas d'étude, nous avons pris comme exemple l'ouïe, c'est-à-dire que nous allons déclencher un fichier audio issu d'une base de données selon l'objet sur lequel le regard est pointé.

Tableau 5: Description des cas d'utilisation du malvoyant

2.2. Diagramme de séquence

Parmi les diagrammes intéressants d'UML, on trouve le diagramme de séquence qui illustre une représentation graphique des interactions entre l'acteur et le système selon un ordre chronologique.

L'utilité du diagramme de séquence est de montrer les interactions d'objet dans le cadre d'un scénario de cas d'utilisation.

La dimension verticale du diagramme représente le temps, permettant de visualiser l'enchaînement des actions dans le temps, et de spécifier la naissance et la mort d'objets. Les périodes d'activité des objets sont symbolisées par des rectangles, et ces objets dialoguent à l'aide de messages.

Cette partie vise à présenter les cas d'utilisation les plus importants dans le projet, ce qui permet de mieux voir et tracer l'enchaînement de l'application.

2.2.1. Paramétrage de la scène

Le graphe ci-dessous schématise l'ensemble des interactions qui permettent à l'assistant de régler la scène pour répondre aux besoins du malvoyant (Figure 25).

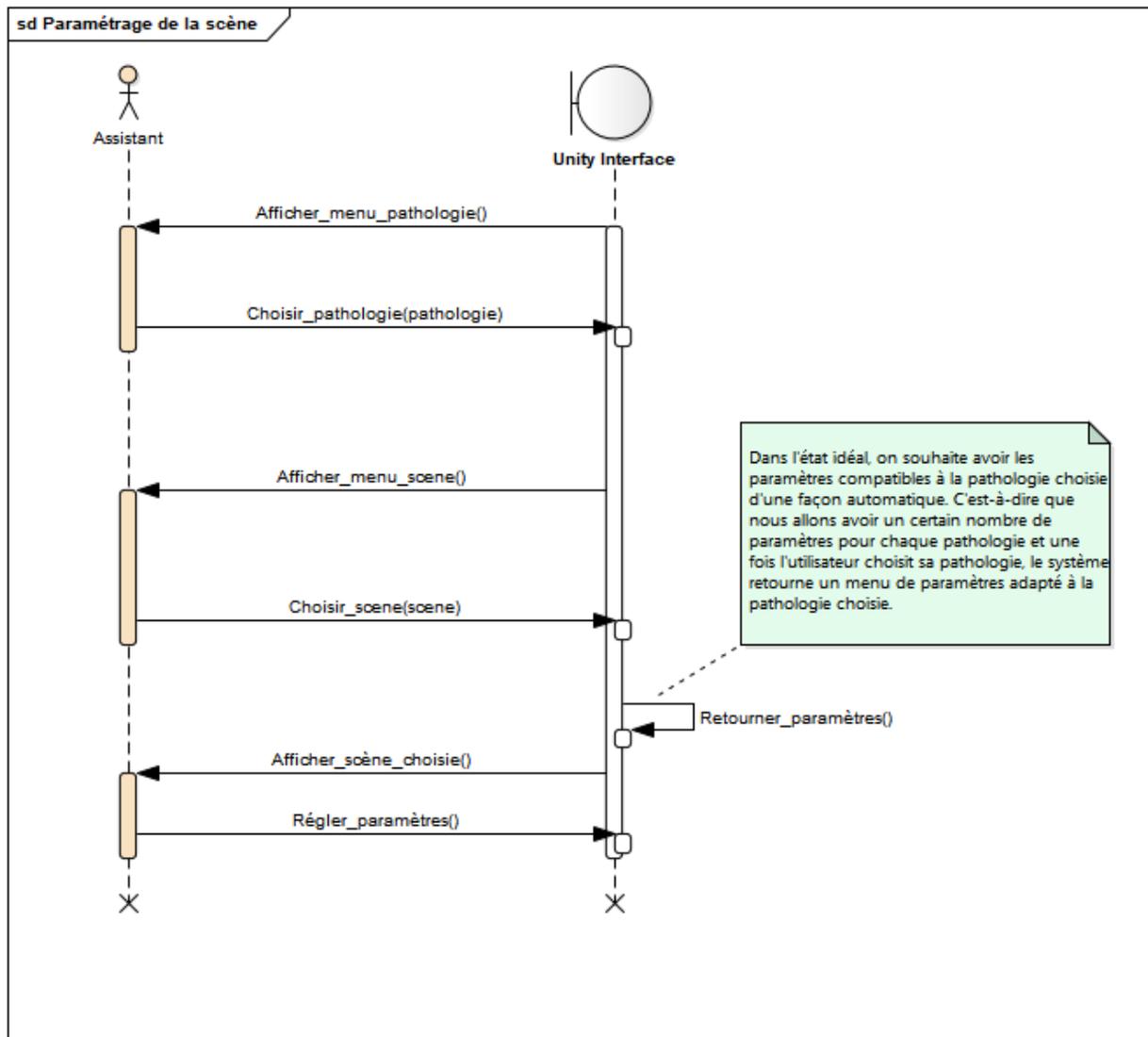


Figure 25: Diagramme de séquence -Paramétrage de la scène-

Pour un malvoyant ayant une déficience visuelle sévère, un assistant choisit à sa place la pathologie et la scène ou l'objet 3D qu'il souhaite explorer. Le système affiche un menu pour choisir la pathologie du malvoyant et ensuite il affiche un autre menu pour choisir la scène. Finalement, l'assistant règle les paramètres de la scène pour qu'elle soit compatible avec la pathologie du malvoyant.

Dans l'état idéal, nous souhaitons avoir la scène paramétrée automatiquement en fonction des pathologies. Autrement dit, quand un utilisateur choisit sa pathologie, il reçoit automatiquement les réglages de la scène compatibles avec son cas.

2.2.2. Déclencher un fichier audio

Le diagramme ci-dessous présente la séquence de messages transmis entre les entités, et modélise le déroulement logique entre la détection de l'axe de regard de l'utilisateur et le déclenchement d'un fichier audio (Figure 26).

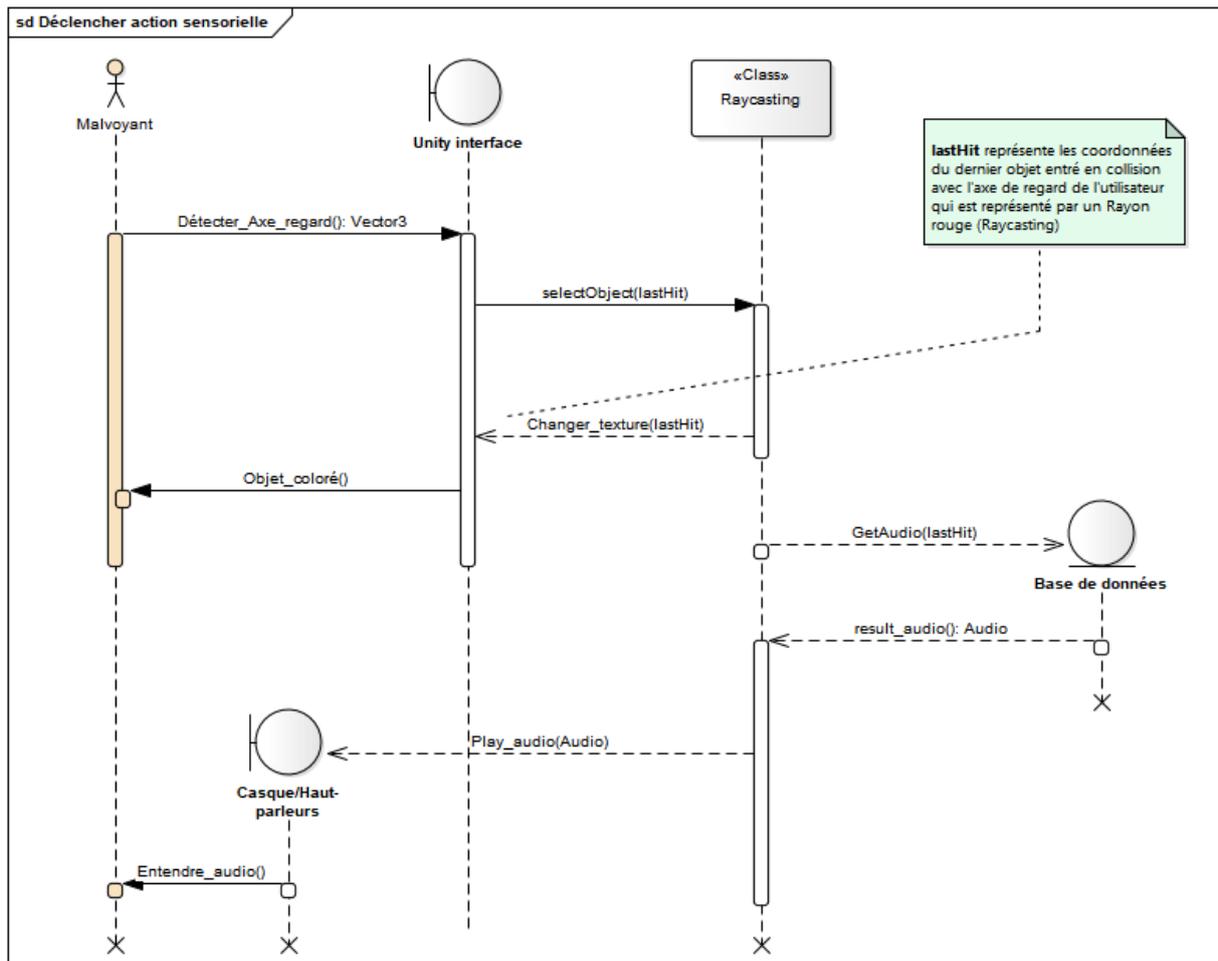


Figure 26 : Diagramme de séquence -Déclencher fichier audio-

En temps réel, le système détecte l’axe de regard de l’utilisateur grâce à la classe « *Raycasting* », en appelant la méthode *selectObject()* qui retourne à son rôle l’objet « *lastHit* » sur lequel l’utilisateur pointe son regard. Ainsi, deux actions se réalisent simultanément, d’une part, le système change la texture de l’objet en modifiant sa couleur, d’une autre part, il interroge la base de données afin de récupérer le fichier audio correspondant à l’objet sélectionné.

2.3. Fonctionnement réel

Afin de mieux appréhender le fonctionnement réel souhaité de notre système, nous présentons un exemple de scénario proposé par Mme. GROUX dans le cahier de charge :

1. LISA déclenche les périphériques correspondants à l’instance (ambiance) de la cathédrale : son (écho) – odeur (encens + cire) – chaleur (froid et humide) – éclairage de l’ordinateur et son entourage (sombre)

2. Avec le profil de Melle Elise³, LISA déduit

- Sens à considérer en priorité : ouïe (1) et odorat (2)
- Sens 1 =ouïe
- Sens 2 = odorat
- Extrait de la base de données toutes les zones-intérêt qui correspondent à ouïe et odorat

3. LISA affiche l'ange dans Unity

- Sélectionne dans le profil : Champ de vision, Sensibilité lumière, Sensibilité ombre, Percep-Couleurs, Percep-gris
- Affichage adapté de l'ange pleureur
 - Zoom et centrage de l'ange (champ de vision).
 - Éclaircir les arrières plans pour ne voir que la forme (sensible au sombre).
 - Baisser la luminosité de la scène (et lumière).
 - Prendre une couleur jaune (pas de problèmes de couleurs)
 - Accentuer les contours.
 - Régler l'épaisseur du trait.

4. En temps réel et tant que non fin

- Détection action usager (axe regard)
 - Lorsque collision zone-interêt-3D
 1. Déclencher les périphériques correspondants au sens-1 si existe
 - a. Exemple / CASQUE qui déclenche la bande son d'un battement d'aile
 - b. Sinon Déclencher les périphériques correspondants au sens-2 si existe
 - a. Exemple / ODEUR d'une plume d'aile.
 - b. Sinon Déclencher une bande d'audiodescription.

³ Mademoiselle Elise est intéressée par la cathédrale d'Amiens, sa structure complète ainsi que son architecture, particulièrement ses sculptures, piliers, le trésor, la rosace et aussi la spatialisation de la cathédrale.

3. Architecture générale du système

Le schéma suivant décrit les ressources matérielles et logicielles qui composent le système :

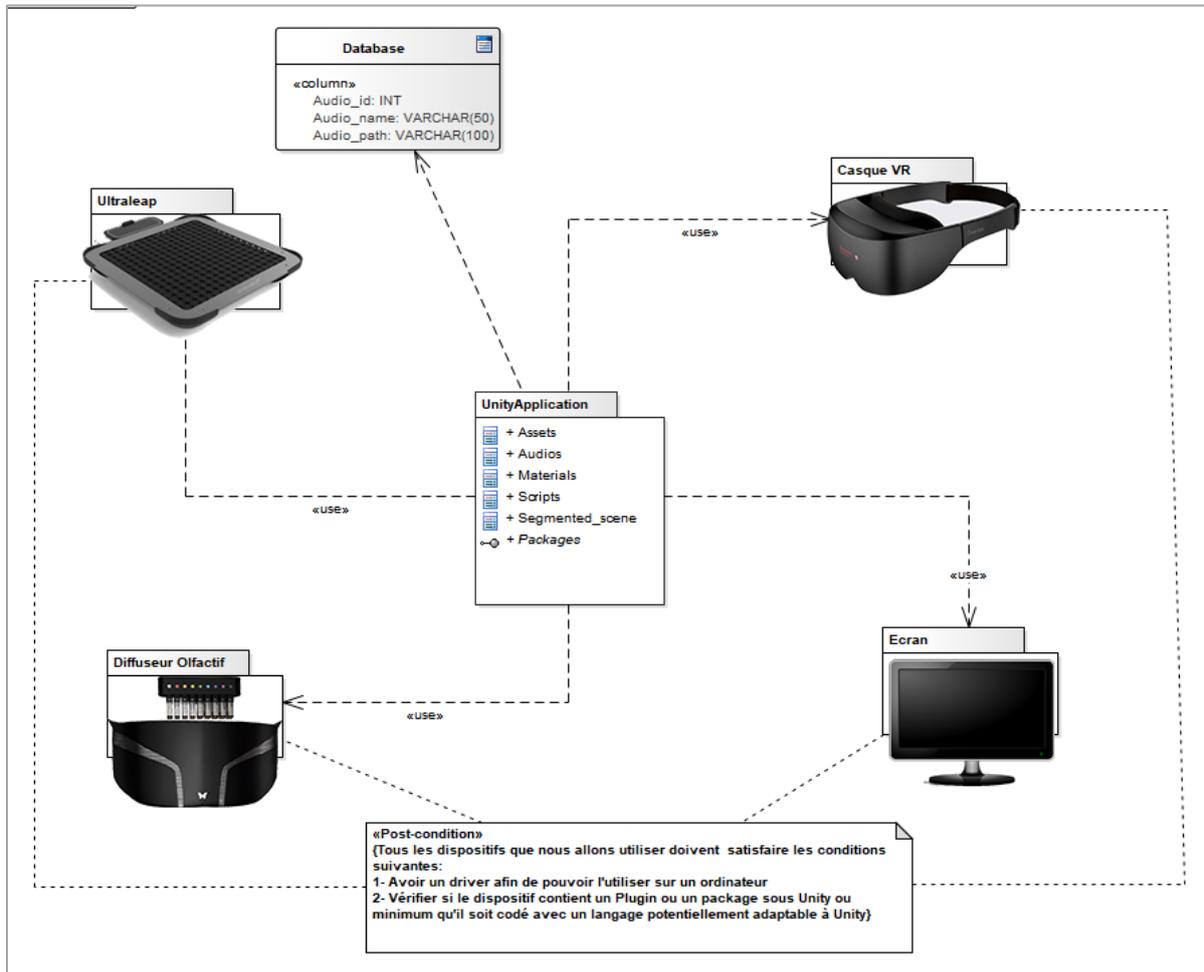


Figure 27: Schéma représentant l'architecture générale du système

Le système se constitue de plusieurs périphériques matérielles et ressources logicielles, nous présentons dans cette section une description de chaque composant :

Casque VR : casque de réalité virtuelle ou casque VR est une révolution technologique qui permet de vous immerger complètement dans des univers virtuels splendides et saisissants de réalisme. A travers ce dispositif nous pouvons visualiser les scènes 3D représentant le patrimoine architectural.

Ultraleap : UltraLeap est une fusion de deux marques : UltraHaptics et LeapMotion. UltraHaptics a racheté LeapMotion en 2019. Les interfaces proposées par la marque utilisent les technologies aussi bien d'UltraHaptics (les matrices de transducteurs à ultrasons pour transmettre des ultrasons aux mains). Les émetteurs à ultrasons font converger les ondes sur un point focal (même point dans l'espace, au même moment). Ce point peut être fixe dans l'espace ou suivre la position de la main. La force de ces ultrasons permet de ressentir une

sensation de relief sur la peau qui sera détecté par les mécanorécepteurs. Ces points de pression combinés permettent de créer plusieurs formes 3D et sensations sur notre peau.

Diffuseur olfactif : L'odorat étant directement lié à la mémoire profonde, il suscite facilement l'émergence d'émotions. C'est précisément cet aspect que l'utilisation d'un diffuseur olfactif vise à solliciter. Malheureusement, nous n'avons pas encore trouvé un diffuseur compatible aux besoins de notre système, mais ça reste une piste de recherche prometteuse que nous n'avons pas encore finalisé.

UnityApplication : représente l'ensemble des composants (Assets, matériaux, textures, scènes 3D, Scripts C#, packages, etc.) constituant notre application.

Database : Une base de données SQL qui contient une table des fichiers audios.

Comme nous l'avons déjà mentionné, notre projet est focalisé sur les deux sens (la vue et l'ouïe). Nous avons utilisé l'interface Unity affichée dans un écran d'ordinateur pour la vue, et des haut-parleurs pour le son.

Conclusion

Ce chapitre nous a permis dans un premier lieu de bien se situer dans le cadre du projet en mettant l'accent sur les objectifs et les missions de notre travail, ensuite nous avons identifié les acteurs ainsi que les cas d'utilisation, avant de clôturer avec une présentation de l'infrastructure générale du système. Le chapitre suivant sera consacré à la mise en place de la solution.

Chapitre 4

Mise en œuvre

Introduction

Pour mettre en place notre application, nous avons eu besoin d'un ensemble d'outils. Dans cette partie l'intérêt est porté aux descriptions des outils utilisés dans la phase de la réalisation du projet, les logiciels et les technologies, ainsi que l'acheminement des interfaces réalisées.

1. Identification des outils de travail

1.1. Unity

L'utilisation de Unity pour développer notre système n'est pas choisi au hasard. Non seulement qu'il est l'outil de développement 3D en temps réel le plus utilisé par les concepteurs et développeurs, mais aussi, qu'il permet d'y intégrer des packages externes des différents dispositifs que nous souhaitons utiliser dans notre système. De plus, il a été proposé par l'équipe du laboratoire, vu qu'ils l'ont déjà utilisé dans le projet PHARES pour mettre en œuvre le dispositif Ultraleap.

Unity est un moteur de jeu multiplateforme (smartphone, ordinateur, consoles de jeux vidéo et Web) développé par Unity Technologies. Il est l'un des plus fournis dans l'industrie du jeu vidéo, aussi bien pour les grands studios que pour les indépendants du fait de sa rapidité aux prototypages et qu'il permet de sortir les jeux sur tous les supports « *Code once, deploy everywhere* ».



Le moteur peut être utilisé pour créer des jeux tridimensionnels (3D) et bidimensionnels (2D), ainsi que des simulations interactives et d'autres expériences. Il a été adopté par des secteurs autres que celui des jeux vidéo, comme le cinéma, l'automobile, l'architecture, l'ingénierie et la construction. Unity offre une API de script primaire en C#, à la fois pour l'éditeur Unity sous la forme de plugins, et pour les jeux eux-mêmes, ainsi qu'une fonctionnalité de glisser-déposer (*Drag and drop*).

Le logiciel de conception développé d'abord pour la plate-forme Mac a été porté sous Windows et permet d'obtenir des applications compatibles avec Windows, Mac OS X, iOS, Android, PlayStation, Xbox 360, WebGL, etc.

Les formats supportés par Unity sont des fichiers images (.psd, .jpg, .png, .svg, etc.), des fichiers audios (.mp3, .ogg, .wav), ou des fichiers de structures 3D (.obj, .b3d, .fbx, .blend, etc.) [16].

1.2. Blender

Comme nous l'avons déjà énoncé dans le chapitre 2, nous avons eu besoin d'un outil pour segmenter et annoter les fichiers résultants des numérisations 3D qui sont sous la forme d'un nuage de point dense. Etant donné que Unity ne supporte pas la segmentation, nous étions obligés d'utiliser un outil externe. Dans un premier temps, nous avons essayé d'utiliser deux logiciels de traitement de nuages de points 3D (MeshLab et CloudCompare). Ils sont des outils très reconnus dans le domaine de la segmentation ainsi que l'annotation sémantique. Néanmoins, le problème que nous avons rencontré, c'est que nous ne pouvons pas utiliser les fichiers obtenus par ces logiciels sous Unity. Après quelques recherches, nous avons trouvé que Blender est le logiciel le plus répandu dans le domaine, ainsi qu'il est supporté par Unity. Le format des fichiers résultants (.blend) sont directement manipulables par Unity sans pour autant les convertir.

Blender est un logiciel libre de modélisation, d'animation par ordinateur et de rendu en 3D, créé en 1994. Il est actuellement développé par la Fondation Blender. Depuis 2019 le logiciel Blender est de plus en plus reconnu par les entreprises du secteur de l'animation 3D, comme Epic Games, Ubisoft et NVIDIA.



Il propose des fonctions avancées de modélisation (dont la sculpture 3D, le texturage et dépliage UV, etc), d'animation 3D (rigging, blend shape), et de rendu (sur GPU comme sur CPU). Il gère aussi le montage vidéo, la composition, la création nodale de matériaux, ainsi que diverses simulations physiques telles que les particules, les corps rigides, les corps souples et les fluides. Ses capacités sont par ailleurs très extensibles, grâce à un système des addons (ajout de scripts à l'aide du langage Python).

Disponible dans de nombreuses langues, Blender est également disponible sur plusieurs plates-formes telles que Microsoft Windows, MacOS, GNU/Linux, etc. De plus, c'est un programme extensible (ajout de scripts à l'aide du langage Python). Blender accepte des extensions en Python, mais n'est pas programmé lui-même dans ce langage pour des raisons de performance. Le code est presque entièrement écrit en C/C++ et compilé à l'aide d'un compilateur classique comme gcc.

Blender permet d'exporter une diversité de formats selon le cas d'étude : 3ds Max file (.3ds), Autodesk FBX (.fbx), Stanford PLY (.ply), Wavefront (.obj), etc. [17].

1.3. C#

Pour contrôler l'état et le comportement des *GameObjects*, nous utilisons le système de *component* de Unity. Un *component* peut être déjà prédéfini par Unity et il suffit de l'attacher à l'objet. Nous pouvons aussi personnaliser le comportement de nos *GameObjects* en ajoutant un script. Un script est une classe, mais lorsqu'il est attaché en tant que *component*, le *GameObject* obtient une instance de la classe.

Les scripts indiquent à nos *GameObjects* comment se comporter ; Ils doivent être attaché à un *GameObject* dans la scène afin d'être appelé par Unity. Les scripts dans Unity sont différents de la programmation pure, c'est-à-dire que nous n'avons pas besoin de créer le code qui exécute l'application, car Unity le fait pour nous. Au lieu de cela, nous nous concentrons sur le gameplay dans nos scripts. Les scripts sont écrits dans un langage spécial que Unity peut comprendre. C'est grâce à ce langage que nous pouvons parler au moteur et lui donner nos instructions. Au-delà de cela, les scripts peuvent être utilisés pour créer des effets graphiques, contrôler le comportement physique des objets ou même mettre en œuvre un système d'IA personnalisé pour les personnages du jeu.

Le langage utilisé dans Unity s'appelle C#. Comme tout langage, les langages de script ont une syntaxe, et les parties principales sont appelées variables, méthodes et classes. C# est un langage de programmation orienté objet, commercialisé par Microsoft depuis 2002 et destiné à développer sur la plateforme Microsoft .NET.

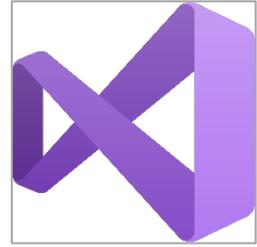


Il est un langage de programmation orienté objet, fortement typé, dérivé de C et de C++, ressemblant au langage Java dont il reprend la syntaxe générale ainsi que les concepts, y ajoutent des notions telles que la surcharge des opérateurs, les indexeurs et les délégués. Il est utilisé pour développer des applications web, ainsi que des applications de bureau, des services web, des commandes, des widgets ou des bibliothèques de classes. En C#, une application est un ensemble de classes où une des classes comporte une méthode principale, comme cela se fait en Java [18].

1.4. Visual studio Community

Unity a la particularité d'utiliser le code (C#) sur la plateforme « .NET » avec l'implémentation Mono. Son éditeur était autrefois basé sur MonoDevelop, via MonoDevelop-Unity, mais à partir de la version 2018.1, il est basé sur Visual Studio Community.

Visual Studio Community est l'un des produits parmi une suite de logiciels de développement pour Windows et MacOS conçue par Microsoft. La dernière version s'appelle Visual Studio Community 2019 et c'est ce que nous avons utilisé dans notre travail. Il est un environnement de développement intégré (IDE) extensible complet et gratuit pour créer des applications web ASP.NET, des services web XML, des applications bureautiques et des applications mobiles [19].



1.5. MySQL

Dans le but de déclencher une action sensorielle autre que la vue, à condition qu'elle soit en dehors de Unity, nous avons utilisé des fichiers audios issus d'une base de données externe, bien que nous puissions juste les intégrer directement dans les « Assets » de notre application. C'est pour cela que nous avons utilisé une base de données SQL pour stocker nos fichiers audios.

MySQL est un système de gestion de bases de données relationnelles (SGBDR). Il est distribué sous une double licence GPL et propriétaire. Il fait partie des logiciels de gestion de base de données les plus utilisés au monde, autant par le grand public (applications web principalement) que par des professionnels, en concurrence avec Oracle, PostgreSQL et Microsoft SQL Server [20].



2. Segmentation avec Blender

Avant de passer à la présentation des interfaces de l'application, nous présentons tout d'abord les étapes suivies pour obtenir un modèle 3D segmenté que nous pouvons manipuler par la suite sous Unity. Nous prenons à titre d'exemple le modèle de l'ange que nous avons déjà présenté.

La figure 28 ci-dessous représente le modèle 3D de l'ange importé dans la plateforme Blender avant la segmentation. Le modèle de l'ange est téléchargé à partir du site web www.Sketchfab.com dans le format (.fbx), qui est bien évidemment supporté par Blender.



Figure 28: Modèle 3D de l'ange avant segmentation

Ensuite, nous avons segmenté le modèle manuellement en différentes zones d'intérêt. Autrement dit, nous avons divisé l'ange en maillage de polygones homogènes en formant des régions significatives (Figure 29).

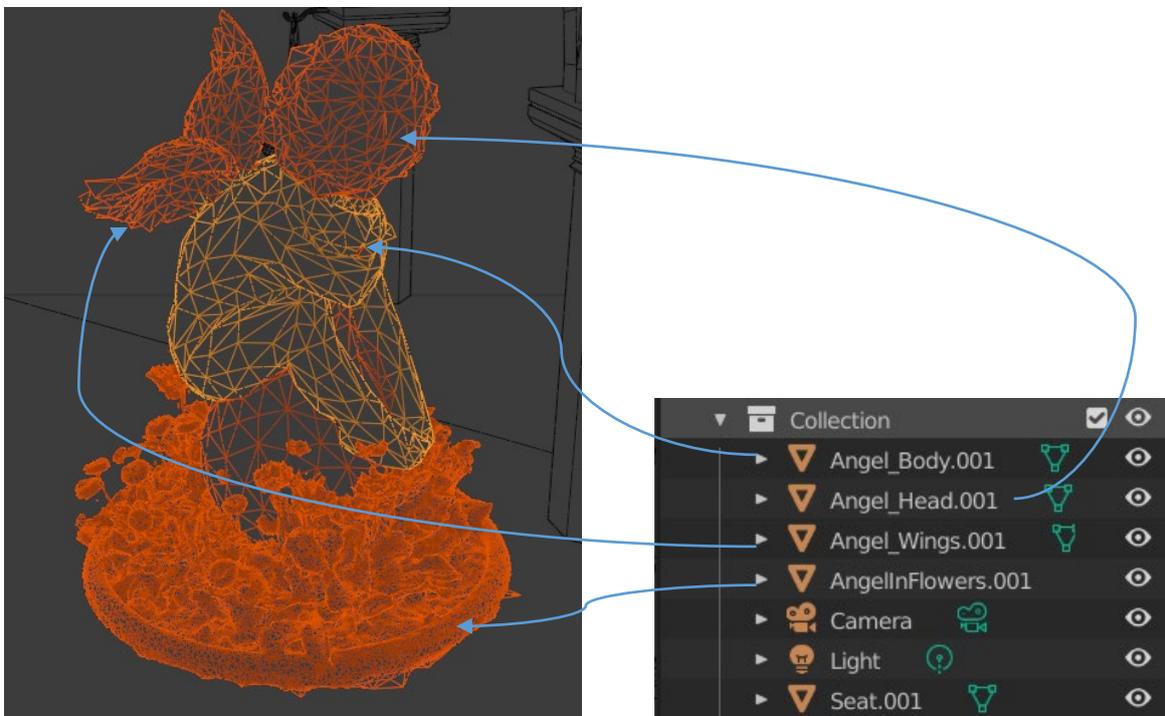


Figure 29: Modèle de l'ange 3D segmenté dans Blender

Etant donnée que Blender est supporté par Unity, alors nous avons enregistré notre modèle sous le projet Unity en format (.blend) et nous avons eu directement l'ange divisé en *GameObjects* sous Unity que nous pouvons manipuler facilement.

La figure 30 représente la fenêtre *Hierarchy* qui contient tous les *GameObjects* de jeu de la scène actuelle, éventuellement, les différents composants du modèle de l'ange.

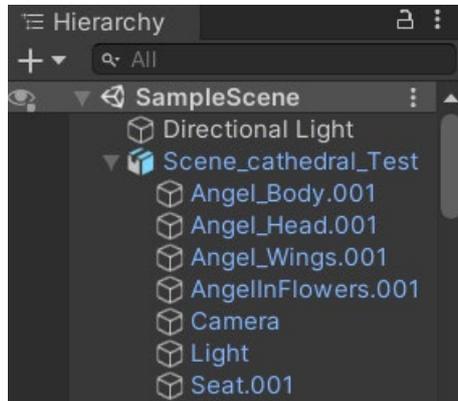


Figure 30: La fenêtre Hierarchy de Unity

3. Présentation de l'interface

Dans cette partie nous avons présenté les captures d'écran de la première version de notre application. Nous avons essayé de suivre un scénario logique, commençant par le premier lancement de l'application par l'utilisateur, passant par l'interface à travers laquelle l'utilisateur peut choisir sa pathologie ainsi que la scène ou l'objet 3D qu'il souhaite explorer.

3.1. Vue d'entrée

La figure 31 illustre l'interface d'accueil de l'application contenant un logo et un bouton pour lancer l'application.



Figure 31: Vue d'accueil de l'application

3.2. Vue de choix de pathologie

La figure 32 représente l'interface à travers laquelle l'utilisateur peut choisir la pathologie. Comme nous l'avons déjà énoncé dans le chapitre 2, il existe plusieurs pathologies, mais nous avons travaillé seulement sur la cataracte, DMLA et le glaucome.

The screenshot shows the LISA application logo at the top left. Below it, a light gray rectangular box contains the text "Veuillez choisir une pathologie:". Underneath this text is a white dropdown menu with "Cataracte" selected and a downward arrow. To the right of the dropdown is a blue button with the text "Sélectionner" in white.

Figure 32: Vue de choix de pathologie

3.3. Vue de choix de la scène

L'interface de la figure 33 permet à l'utilisateur de choisir la scène ou l'objet 3D qu'il veut explorer. Eventuellement, la scène sera divisée en plusieurs objets 3D, et l'utilisateur choisira l'objet ou la partie de la scène qu'il souhaite scruter.

The screenshot shows the LISA application logo at the top left. Below it, a light gray rectangular box contains the text "Veuillez choisir une scène:". Underneath this text is a white dropdown menu with "Scène 1" selected and a downward arrow. To the right of the dropdown is a blue button with the text "Sélectionner" in white.

Figure 33: Vue de choix de scène ou objet 3D

3.4. Vues de la scène 3D

Dans cette section nous présentons les différentes vues de la scène que nous avons utilisée dans notre travail. La scène est un modèle 3D d'une cathédrale et un ange.

3.4.1. Vues de l'intérieure de la scène

Il est présenté ci-dessous la vue à partir de différents angles :



Figure 34: Illustration de quelques angles de vue de la scène 3D

3.4.2. Modèle de l'ange

Il est représenté ci-dessous la vue principale de l'application avec un menu que nous pouvons masquer en cliquant sur le bouton « Menu ». Le menu des paramètres permet à l'utilisateur d'adapter la scène ou l'objet 3D à sa vue. Il contient jusqu'à présent, un *slider* pour modifier la luminosité et un sous-menu pour régler les contours. Ce sous-menu contient un *slider* pour l'épaisseur et un sélecteur de couleur pour la couleur des contours.



Figure 35: Modèle 3D de l'ange

La figure 36 représente le bon fonctionnement d'une des missions principales de notre travail. Il était demandé de déterminer où l'utilisateur regarde, autrement dit, définir l'axe du regard (représenté par le rayon) de l'utilisateur (représenté dans notre cas par le cylindre rouge) et son intersection avec les objets de la scène. En temps réel, à chaque collision entre la ligne de regard de l'utilisateur et un objet, nous distinguons ce dernier en modifiant la couleur de sa texture en vert.

Nous pouvons remarquer dans cette illustration que le regard de l'utilisateur est orienté vers les ailes de l'ange.

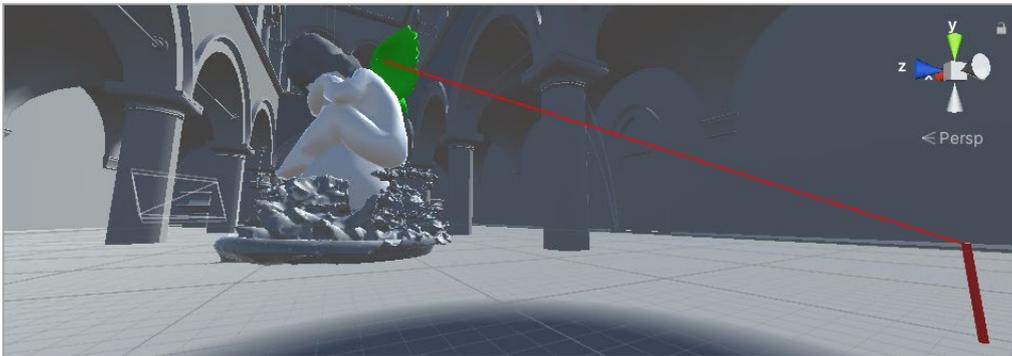


Figure 36: Illustration représentant la ligne de regard de l'utilisateur

De plus, nous avons réussi à synchroniser la vue à un autre sens qui est dans notre cas « l'ouïe », c'est-à-dire que nous pouvons lancer un fichier audio issu d'une base de données selon la partie de l'objet/scène observée. Les deux figures 37 et 38 représentent le bon fonctionnement du traitement.

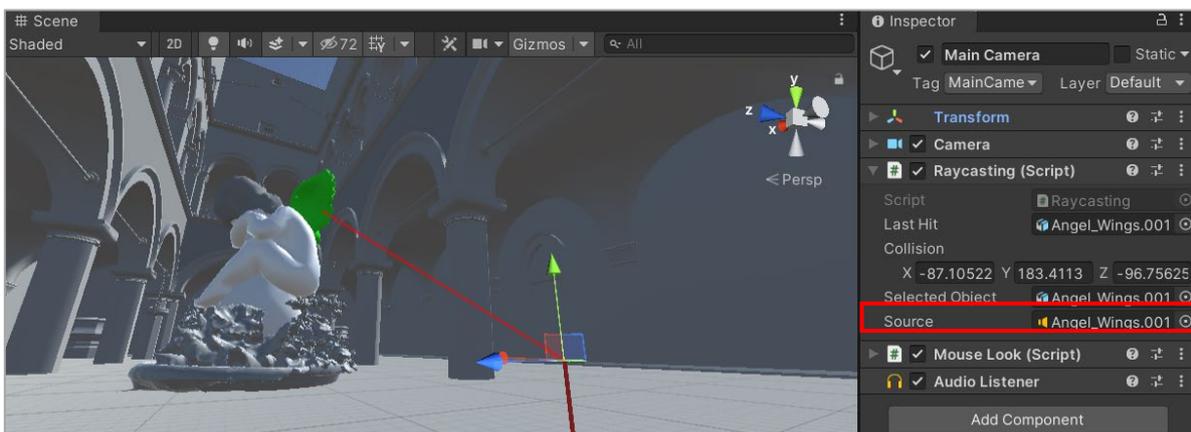


Figure 37: Illustration du bon fonctionnement de la synchronisation de la vue et l'audition

	id	objectName	audioPath
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	1	wings	C:/xampp/htdocs/sqlconnect/sound/

Figure 38: La table de la base de données contenant le fichier audio récupéré

Conclusion

Ce dernier chapitre porte sur la réalisation des objectifs de notre travail, en partant de la segmentation du modèle 3D jusqu'à la synchronisation de la vue avec l'audition, passant par l'illustration des différentes vues de la scène utilisée dans notre projet.

Notre contribution dans ce projet jusqu'à présent ne concerne que les deux sens (la vue et l'ouïe). Le travail qui reste consiste à essayer d'intégrer d'autres sens (le toucher et l'odorat) en continuant l'étude des autres dispositifs cités dans le chapitre précédant. Nous essaierons de réaliser ce travail dans les temps qui suivent.

Conclusion & perspectives

Ce mémoire a été le fruit de mon projet de fin d'études que j'ai réalisé au sein du laboratoire MIS d'Amiens, France. Dans ce projet de recherche et développement, nous avons participé au sujet de recherche PHARES, qui a comme but d'apporter une solution multisensorielle aux malvoyants afin de leur permettre l'accès au patrimoine 3D par une expérience qui diffère des expériences actuelles.

Notre propre contribution comprend une réponse au cahier des charges fourni par l'équipe du laboratoire. Il contient comme missions principales, d'une part, le traitement des fichiers résultants des numérisations 3D réalisées par le laboratoire pour qu'ils soient maniables. D'une autre part, la conception et la réalisation d'une interface sous Unity qui intègre deux sens jusqu'à présent (la vue et l'ouïe).

La mise en place de cette solution a nécessité dans un premier temps, une identification du système visuel humain ainsi que les déficiences visuelles affectants la vision. Ensuite, nous avons étudié les méthodes de traitement des données 3D et en parallèle, nous avons suivi une formation sur Unity afin de pouvoir implémenter la solution.

Le travail réalisé n'est que la première version du système. Il reste encore plusieurs voies d'améliorations. Nous pouvons ajouter à l'interface d'autres sens (le toucher et l'odorat). D'ailleurs, nous avons déjà entamé cette piste de recherche par l'étude du dispositif Ultraleap et nous avons aussi commencé à chercher un diffuseur olfactif qui peut être compatible à notre système.

En effet, cette expérience m'a permis de s'initier au domaine de la recherche. De plus, c'était une opportunité pour acquérir des nouvelles compétences et connaissances sur la 3D et de maîtriser des nouveaux outils très intéressants, surtout le moteur de jeu Unity et le logiciel de modélisation 3D Blender.

Par le présent rapport, nous présentons le projet qui marque la fin de notre cycle Master Systèmes intelligents et Réseaux.

Bibliographie et Webographie

- [1] « Anatomie et fonctionnement de l'œil | Dr Leininger ». <https://dr-leininger.fr/loeil-et-la-vision/anatomie-de-loeil> (consulté le juin 18, 2021).
- [2] G. DENIS, « Apport de la vision par ordinateur dans l'utilisabilité des neuroprothèses visuelles », Thèse de doctorat, Université de Toulouse 3 Paul Sabatier, Toulouse, France, 2014. [En ligne]. Disponible sur: <http://thesesups.ups-tlse.fr/2593/>
- [3] illusion-optique-07, « I. 4) les cônes et les bâtonnets », *Skyrock*, févr. 02, 2010. <https://illusion-optique-07.skyrock.com/2774246150-I-4-les-cones-et-les-batonnets.html> (consulté le juin 18, 2021).
- [4] H. N. Dahal, « VISUAL FIELD TESTING AND INTERPRETATION », p. 82.
- [5] « Cécité et déficience visuelle ». <https://www.who.int/fr/news-room/fact-sheets/detail/blindness-and-visual-impairment> (consulté le juin 18, 2021).
- [6] J. D. Steinmetz *et al.*, « Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study », *Lancet Glob. Health*, vol. 9, n° 2, p. e144-e160, févr. 2021, doi: 10.1016/S2214-109X(20)30489-7.
- [7] « Eye Disease Simulations | National Eye Institute (NEI) Media Library ». <https://medialibrary.nei.nih.gov/search?keywords=&f%5B0%5D=category%3A8> (consulté le juin 18, 2021).
- [8] « Glaucome : Types, symptômes, diagnostic, suivi, traitements », *Santé sur le net*. <https://www.sante-sur-le-net.com/maladies/ophtalmologie/glaucome/> (consulté le juin 18, 2021).
- [9] K. Jacquot, « Numérisation et restitution virtuelle des maquettes de la collection de Louis XIV, le cas des fortifications bastionnées », Thèse de doctorat, Université de Lorraine, Nancy, France, 2014. [En ligne]. Disponible sur: <https://hal.univ-lorraine.fr/tel-01750670>
- [10] V. Abergel, « Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte », Thèse de doctorat, HESAM Université, Aix-en-Provence, France, 2020. [En ligne]. Disponible sur: <https://tel.archives-ouvertes.fr/tel-02950598>
- [11] A. MANUEL, « Annotation sémantique 2D/3D d'images spatialisées pour la documentation et l'analyse d'objets patrimoniaux », Thèse de doctorat, Arts et Métiers, ParisTech, Aix-en-Provence, France, 2016. [En ligne]. Disponible sur: <https://pastel.archives-ouvertes.fr/tel-01304776>
- [12] M. Pierrot-Deseilligny, L. De Luca, et F. Remondino, « Automated Image-Based Procedures for Accurate Artifacts 3D Modeling and Orthoimage Generation », *Geoinformatics FCE CTU*, vol. 6, p. 291-299, déc. 2011, doi: 10.14311/gi.6.36.

-
- [13] F. Remondino, S. Del Pizzo, T. P. Kersten, et S. Troisi, « Low-Cost and Open-Source Solutions for Automated Image Orientation – A Critical Overview », in *Progress in Cultural Heritage Preservation*, vol. 7616, M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, et R. Caffo, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, p. 40-54. doi: 10.1007/978-3-642-34234-9_5.
- [14] *Angel relaxes in flowers - Download Free 3D model by zarGByte (@zarGByte) [be9bd91]*. Consulté le: juin 23, 2021. [En ligne Video]. Disponible sur: <https://sketchfab.com/models/be9bd91a129a4ded92fe63b52ea962ab/embed?autostart=1>
- [15] ANDREWS, P., ZAIHRAYEU, I. et PANE, J., 2012. A classification of semantic annotation systems. In : *Semantic Web*. n° 3, p. 223–248. DOI 10.3233/SW-2011-0056.
- [16] « Unity (moteur de jeu) », *Wikipédia*. juin 04, 2021. Consulté le: juill. 03, 2021. [En ligne]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=Unity_\(moteur_de_jeu\)&oldid=183548063](https://fr.wikipedia.org/w/index.php?title=Unity_(moteur_de_jeu)&oldid=183548063)
- [17] « Blender », *Wikipédia*. juin 05, 2021. Consulté le: juill. 03, 2021. [En ligne]. Disponible sur: <https://fr.wikipedia.org/w/index.php?title=Blender&oldid=183561954>
- [18] « C sharp », *Wikipédia*. mai 22, 2021. Consulté le: juill. 04, 2021. [En ligne]. Disponible sur: https://fr.wikipedia.org/w/index.php?title=C_sharp&oldid=183137863
- [19] « Microsoft Visual Studio », *Wikipédia*. févr. 03, 2021. Consulté le: juill. 05, 2021. [En ligne]. Disponible sur: https://fr.wikipedia.org/w/index.php?title=Microsoft_Visual_Studio&oldid=179539161
- [20] « MySQL », *Wikipédia*. mars 22, 2021. Consulté le: juill. 08, 2021. [En ligne]. Disponible sur: <https://fr.wikipedia.org/w/index.php?title=MySQL&oldid=181111841>

RÉALISATION D'UN PROTOTYPE D'UNE INTERFACE MULTISENSORIELLE POUR LES MALVOYANTS

Résumé

Afin de mettre à disposition des malvoyants le patrimoine architectural, l'équipe PR du laboratoire MIS d'Amiens a initié le projet PHARES (Percevoir grâce à une interface HAptique pour IEs malvoyantS) qui vise à utiliser les modèles 3D du patrimoine pour apporter aux malvoyants une solution multisensorielle qui diffère des expériences actuelles.

Dans ce cadre vient ce projet qui a comme but, d'initier la conception et la réalisation du système en réalisant un prototype sous Unity, qui permet de détecter l'axe de regard de l'utilisateur ainsi que la synchronisation de la vue à un autre sens (celui de l'audition) et lancer un fichier audio issu d'une base de données selon la partie de l'objet 3D observée.

Mots clés : patrimoine architectural, multisensorielle, déficience visuelle, modèle 3D, Unity.

REALIZATION OF A PROTOTYPE OF A MULTISENSORY INTERFACE FOR THE VISUALLY IMPAIRED

Abstract

In order to make the architectural patrimony available to the visually impaired, the PR team of the MIS laboratory in Amiens has initiated a project intitled PHARES (Percevoir grâce à une interface HAptique pour IEs malvoyantS). The project aims to use 3D models of the architectural patrimony to provide the visually impaired with a multisensory experience that differs from current experiences.

Within this context, comes this project which aims to initiate the design and realization of the system by creating a prototype using the game engine Unity. The implemented prototype allows to detect the user's line of sight, as well as the synchronization of the sight with another sense (the hearing one) and to launch an audio file from a database depending on the observed part of the 3D object.

Keywords: architectural patrimony, multisensory, visual impairment, 3D model, Unity