



HAL
open science

Naviguer en vision prothétique simulée : apport de la vision par ordinateur pour augmenter les rendus prothétiques de basse résolution

Victor Vergnieux

► **To cite this version:**

Victor Vergnieux. Naviguer en vision prothétique simulée : apport de la vision par ordinateur pour augmenter les rendus prothétiques de basse résolution. Environnements Informatiques pour l'Apprentissage Humain. Université Paul Sabatier - Toulouse III, 2015. Français. NNT : 2015TOU30323 . tel-01357149

HAL Id: tel-01357149

<https://theses.hal.science/tel-01357149>

Submitted on 29 Aug 2016

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

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



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier (UT3 Paul Sabatier)

Discipline : Informatique

Spécialité : Image, Information, Hypermédia

Présentée et soutenue par Victor Vergnieux
Le 2 décembre 2015

**Titre : Naviguer en vision prothétique simulée :
apport de la vision par ordinateur pour augmenter
les rendus prothétiques de basse résolution**

École doctorale : ED MITT

Unité de Recherche :

Institut de Recherche en Informatique de Toulouse (IRIT) – UMR 5505

Directeurs de thèse :

Christophe Jouffrais

Chercheur CNRS, Université Paul Sabatier & IRIT

Directeur de thèse

Marc Macé

Chercheur CNRS, Université Paul Sabatier & IRIT

Co-directeur de thèse

Jury :

Isabelle Viaud Delmon

*Directeur de Recherche CNRS, Université Pierre et Marie Curie &
IRCAM*

Rapporteur

Marie-Chantal Wanet-Defalque

Chercheur et Responsable de site CRIR-INLB, Université de Montréal

Rapporteur

Frédéric Chavane

Chercheur CNRS, Institut de Neurosciences de la Timone

Examineur

Jean-Pierre Jessel

Professeur, Université Paul Sabatier & IRIT

Examineur

Christophe Jouffrais

Chercheur CNRS, Université Paul Sabatier & IRIT

Directeur de thèse

Marc Macé

Chercheur CNRS, Université Paul Sabatier & IRIT

Co-directeur de thèse

Résumé

La cécité touche 39 millions de personnes dans le monde et génère de nombreuses difficultés dans la vie quotidienne. Plus précisément, les capacités de navigation (incluant orientation et mobilité) sont fortement diminuées, ce qui amène les personnes non-voyantes à limiter, voire à cesser leurs déplacements. Pour restaurer des sensations «visuelles», et par-delà, une certaine autonomie, il est possible de stimuler directement le système visuel résiduel d'une personne non-voyante à l'aide d'un implant administrant des micro-stimulations électriques. Le dispositif complet se compose d'une micro-caméra portée sur des lunettes et reliée à un ordinateur de poche, qui lui-même est connecté à l'implant. Lors des micro-stimulations, les sujets perçoivent des tâches grises, blanches ou jaunâtres appelées phosphènes. Ainsi la qualité de la vision restaurée est directement dépendante de la résolution et de la position de l'implant. Le nombre d'électrodes étant faible pour les implants en développement (moins d'une centaine), il est nécessaire de réduire drastiquement la résolution du flux vidéo pour la faire correspondre à la faible résolution de l'implant. Actuellement, l'Argus II de la société Second Sight est l'implant dont le développement est le plus avancé et sa résolution est de 60 électrodes, ce qui permet aux patients implantés de percevoir 60 phosphènes différents. Cette vision restaurée est donc très pauvre et un travail d'optimisation du signal est nécessaire pour pouvoir utiliser l'implant de manière fonctionnelle. Les sujets implantés sont impliqués dans des protocoles cliniques fermés ne permettant pas de les inclure dans d'autres expériences. Malgré cela, il est possible d'étudier les possibilités offertes par ces implants visuels en simulant la vision prothétique dans un visiocasque porté par des sujets voyants. Il s'agit du domaine de la vision prothétique simulée (VPS). La navigation n'a jamais été étudiée chez les patients implantés et très rarement en VPS. Il s'avère qu'avec des implants de très faible résolution, elle pose de grandes difficultés liées à la mobilité mais également des difficultés liées à l'orientation. Les travaux entrepris dans ce doctorat se concentrent sur l'étude de la navigation en VPS. Différentes théories en psychologie nous ont permis d'identifier les éléments importants pour les sujets afin qu'ils se repèrent et se construisent une représentation mentale fiable de l'environnement lors de la navigation. À partir de ces modèles, différents rendus prothétiques utilisant la vision par ordinateur ont été conçus et testés dans une tâche de navigation réalisée dans un environnement virtuel. Les expérimentations effectuées avaient pour objectif d'optimiser la perception et la compréhension de l'espace parcouru avec un implant de faible résolution. Ces évaluations reposaient sur la performance de temps des sujets pour effectuer la tâche de navigation et sur leur représentation mentale de l'environnement. Après la tâche de navigation, il leur était demandé de dessiner la carte des environnements explorés, afin d'évaluer ces représentations. Cette double évaluation a permis d'identifier les indices importants permettant de faciliter la perception et la mémorisation de la structure des

environnements dans une tâche de navigation en VPS. Pour améliorer les performances des personnes non-voyantes implantées, il apparaît notamment nécessaire de limiter la quantité d'information présentée, tout en préservant la structure de l'environnement grâce à des algorithmes de vision par ordinateur. Lorsque l'accès à des patients implantés sera plus ouvert, il deviendra nécessaire de valider ces différents résultats en les faisant naviguer en environnement virtuel puis en environnement réel.

Mots clés : Cognition Spatiale, Neuroprothèses Visuelles, Simulation de Vision Prothétique, Réalité Virtuelle, Navigation

Abstract

Blindness affects thirty nine millions people in the world and generates numerous difficulties in everyday life. Specifically, navigation abilities (which include wayfinding and mobility) are heavily diminished. This leads blind people to limit and eventually to stop walking outside. Visual neuroprosthesis are developed in order to restore such “visual” perception and help them to get some autonomy back. Those implants generate electrical micro-stimulations which are focused on the retina, the optic nerve or the visual cortex. Those stimulations elicit blurry dots called “phosphenes”. Phosphenes can be mainly white, grey or yellow. The whole stimulation device contains a wearable camera, a small computer and the implant which is connected to the computer. The implant resolution and position impact directly the quality of the restored visual perception. Current implants include less than a hundred electrodes so it is mandatory to reduce the resolution of the visual stream to match the implant resolution. For instance, the already commercialized Argus II implant from the company Second Sight (Seymar, California) is the leading visual implant worldwide and uses only sixty electrodes. This means that Argus II blind owners can perceive only sixty phosphenes simultaneously. Therefore this restored vision is quite poor and signal optimization is required to get to a functional implant usage. Blind people with implants are involved in restricted clinical trials and are difficult to reach. Yet, studying those implant possibilities is at our reach by simulating prosthetic vision and displaying it in a head mounted display for sighted subjects. This is the field of simulated prosthetic vision (SPV). Navigation was never studied with people with implant, and only a few studies approached this topic in SPV. In this thesis, we focused on the study of navigation in SPV. Computer vision allowed us to select which of the scene elements to display in order to help subjects to navigate and build a spatial representation of the environment. We used psychological models of navigation to conceive and evaluate SPV renderings. Subjects had to find their way and collect elements in a navigation task in SPV inspired by video games for the blind. To evaluate their performance we used a performance index based on the completion time. To evaluate their mental representation, we asked them to draw the environment layout after the task for each rendering. This double evaluation lead us to spot which elements can and should be displayed in low resolution SPV in order to navigate. Specifically those results show that to be understandable in low vision, a scene must be simple and the structure of the environment should not be hidden. When blind people with implant will become available we will be able to confirm or deny those results by evaluating their navigation in virtual and real environments.

Keywords: Spatial Cognition, Visual Neuroprosthesis, Simulated Prosthetic Vision, Virtual Reality, Navigation

Remerciements

Tout d'abord je souhaite remercier les membres du jury : Jean-Pierre Jessel, Frédéric Chavane, et les deux rapporteurs : Isabelle Viaud-Delmon et Marie-Chantal Wanet-Defalque. Merci beaucoup d'avoir pris le temps de lire ce mémoire en détails et pour vos précieux retours. Le travail pour arriver au doctorat se produit rarement en solitaire et je voudrais remercier mes deux encadrants : Christophe et Marc. Christophe, ta bonne humeur communicative et ta capacité à remotiver tes étudiants lorsqu'ils doutent m'ont bien aidé lors des périodes de creux. Merci également de m'avoir aiguillé dans le bon sens quand j'en avais besoin. Marc, merci beaucoup pour ta disponibilité pendant cette thèse. Ton appui quotidien ainsi que tes multiples relectures de mes productions m'ont été très précieux. Merci aussi à Grégoire qui a soutenu sa thèse un an avant moi sur un sujet voisin. Les résultats présentés dans ce mémoire lui doivent beaucoup.

Je tiens à remercier l'équipe Elipse de m'avoir accueilli en son sein pendant quasiment quatre années. La bonne ambiance qui y règne permet à chacun d'obtenir des retours constructifs sur son travail et de toujours trouver quelqu'un à qui parler en cas de coup dur. La plupart des doctorants n'ont pas la chance d'être dans une équipe qui fonctionne aussi bien. Pour n'oublier personne de l'équipe je ne vais pas citer de noms mais merci à tout le monde pour cette bonne dynamique ! Une pensée particulière pour Damien, Julie et Philippe qui ont dû me supporter et me supportent encore dans le bureau 270. Je pense aussi au Dr. Loupi qui m'a beaucoup aidé. Bon courage à tous ceux qui rédigent leur manuscrit en lisant ces mots !

À partir de ma deuxième année de thèse, j'ai eu la chance de participer à la vie administrative de l'École Doctorale MITT. Cette expérience – intense par moment – m'a beaucoup apporté sur le plan humain. Merci donc à toute l'équipe de l'EDMITT et en particulier à Agnès et Martine (les secrétaires infatigables), à Jean-Michel, Chantal et Jean-Paul avec qui nous avons géré tant bien que mal des médiations parfois difficiles. Je n'oublie pas Laure qui m'a transmis le flambeau et Mohamed qui a pris la suite.

Merci aux membres des associations Incognu et Fresco avec qui j'ai partagé de passionnants et excellents moments autour des sciences cognitives. Merci à tous les amis qui m'ont apporté beaucoup d'oxygène quand j'en avais besoin. Merci à Nanard, meilleur parmi les meilleurs et à Joseph pour ses rires et ses olives.

Merci à mes parents qui ont toujours été de bons conseils et m'ont beaucoup apporté. Je vous dois tant et tant, merci de m'avoir donné le goût de la recherche et pour votre soutien sans faille. Merci aussi à ma petite sœur préférée qui m'a montré ce que détermination voulait dire.

Enfin merci à Tiffany, mon épouse qui a beaucoup enduré lors de cette dernière année. Sans son soutien, sa participation à la correction du manuscrit et sa mise en page je n'aurais tout simplement pas pu rendre ce mémoire dans les délais que je m'étais fixé. Merci beaucoup pour tous les sacrifices que tu as faits, pour toutes les petites joies et les grands bonheurs que tu me donnes chaque jour. Avec tout mon amour, ce mémoire t'est dédié.

Lhèu veiram pas jamei la fin,
La libertat qu'ei lo camin.
Peut-être n'en verra-t-on jamais la fin,
La liberté, c'est le chemin.

Nadau - De cap tà l'immortèla (extrait)

Table des matières

Résumé.....	3
Abstract.....	5
Remerciements	7
Introduction	15
A. Causes de la déficience visuelle	15
B. Évaluation de la déficience visuelle.....	15
C. Handicaps dus à la cécité.....	16
D. Systèmes d'assistance ou de suppléance	16
PARTIE 1. ÉTAT DE L'ART.....	21
Chapitre I.Le système visuel.....	23
A. Structure de l'œil et trajet de la lumière.....	23
B. La rétine.....	24
C. Le nerf optique	25
D. Le cortex visuel.....	26
Chapitre II Neuroprothèses Visuelles	29
A. Premières expériences	29
B. Projets d'implants et essais cliniques	30
B.1. Cortical.....	31
B.1.a. Monash Vision Group (Australia).....	31
B.1.b. National Institute of Health (États-Unis)	31
B.1.c. Intracortical Visual Prosthesis	32
B.1.d. Utah Electrode Array	33
B.1.e. École Polytechnique de Montréal (Canada)	33
B.1.f. Projet Orion (Second Sight).....	34
B.2. Nerf optique.....	34
B.2.a. Veraart (Belgique)	34
B.2.b. Chinese Project for Sight (Chine)	36
B.2.c. AV-DONE (Japon)	37
B.3. Rétine.....	37
B.3.a. Second Sight	37

B.3.b. Retina Implant et le projet Alpha-IMS	40
B.3.c. Boston Retinal Implant Project	41
B.3.d. Intelligent Medical Implant et Pixium Vision.....	42
B.3.e. Bionic Vision Australia	43
B.3.f. Artificial Silicon Retina (ASR).....	43
B.3.g. EpiRet Gmbh.....	44
Chapitre III.Simulation de la vision prothétique	47
A. Principe	47
B. Acuité visuelle.....	49
C. Lecture	51
D. Reconnaissance et localisation de visages.....	53
E. Reconnaissance et saisie d'objets	54
F. Mobilité et navigation.....	58
Chapitre IV.Modèles de cognition spatiale	67
A. Introduction et définition	67
B. Représentations acquises lors de la navigation.....	70
B.1. L'intuition de Tolman	70
B.2. Le cadre dominant : modèle L-R-S.....	71
B.3. Dépasser le cadre dominant : apports de Montello.....	73
B.4. Network of Reference Frame theory	74
B.5. Cognition spatiale des non-voyants	77
PARTIE 2. CONTRIBUTIONS.....	81
Chapitre V. Approche par localisation appliquée à la navigation.....	83
A. Matériel et méthodes	84
A.1. Rendus.....	84
A.2. Matériel	87
B. Protocole	89
B.1. Sujets.....	89
B.2. Tâche.....	89
B.3. Variables dépendantes.....	90
B.4. Hypothèses.....	90
C. Résultats.....	91
C.1. Résultats quantitatifs	91
C.2. Résultats qualitatifs.....	92
D. Discussion	93

Chapitre VI. Contextualiser pour mieux naviguer en vision prothétique augmentée.....	97
A. Matériel et Méthodes	98
A.1. Environnements.....	98
A.2. Rendus.....	99
A.3. Tâche.....	100
A.4. Matériel	101
A.5. Sujets.....	102
A.6. Variables dépendantes.....	103
A.7. Hypothèses.....	103
B. Résultats.....	104
B.1. Résultats quantitatifs	104
B.2. Résultats qualitatifs.....	105
C. Discussion	106
D. Conclusion	108
Chapitre VII. Différentes méthodes pour augmenter la structure en VPS.....	109
A. Matériel et méthode.....	111
A.1. Pré-expérimentation.....	111
A.2. Rendus.....	111
A.3. Hypothèses.....	114
A.3.a. Rendus Irrlicht	114
A.3.b. Rendus VPS.....	114
A.3.c. Autres facteurs	115
A.4. Sujets.....	115
A.5. Matériel	115
A.6. Environnements.....	117
A.7. Tâche et variables observées	117
B. Résultats.....	119
B.1. Indice de performance	119
B.2. Charge cognitive	120
B.3. Évaluation des dessins.....	122
B.4. Sens de l'orientation (score au questionnaire de Santa Barbara)	123
B.5. Influence de la pratique des jeux vidéo et de la connaissance de la réalité virtuelle.....	124
C. Discussion	126
C.1. Résultats généraux	127
C.2. Avantage des joueurs	130
D. Conclusion	134
Chapitre VIII. Navigation et points de repères	137
A. Protocole	137

B. Sujets	138
C. Résultats.....	138
D. Discussion et Conclusion.....	139
PARTIE 3. DISCUSSION ET PERSPECTIVES.....	141
Chapitre IX. Principaux résultats	145
Chapitre X. Limitations des différents résultats observés	149
A. L'évaluation de la navigation en VPS	149
B. Validation de la Simulation de Vision Prothétique	151
C. Techniques d'interaction utilisées (clavier et casque) pour naviguer dans un environnement virtuel.....	151
D. Joueurs et non-joueurs, interaction et rendus utilisés.....	152
E. Variabilité inter-individuelle.....	153
F. Les résultats sont-ils applicables à des patients implantés ?	153
Chapitre XI. Discussion générale et perspectives	155
A. Utilisabilité des implants à basse résolution pour naviguer	155
B. Sujets implantés et sujets voyants.....	156
C. Importance des algorithmes de vision embarquée et des nouvelles technologies.....	157
BIBLIOGRAPHIE.....	161
TABLE DES FIGURES	179
TABLE DES TABLEAUX.....	186
ANNEXES	187

Introduction

La vision est le sens prédominant chez la plupart des primates diurnes et en particulier chez l'homme. La perte de ce sens implique donc un handicap majeur. D'après l'Organisation Mondiale de la Santé (OMS), deux cent quatre-vingt-cinq millions de personnes dans le monde sont déficientes visuelles et un peu moins de quarante millions d'entre elles sont aveugles [WHO 2014]. Quatre-vingt-dix pour cent des aveugles vivent dans des pays en voie de développement et quatre-vingt-deux pour cent des aveugles ont cinquante ans ou plus. Le nombre de personnes aveugles en raison de maladies infectieuses a diminué depuis une vingtaine d'années mais le vieillissement de la population dans de nombreux pays amène à une augmentation de la population touchée.

A. Causes de la déficience visuelle

La principale cause de déficience visuelle est la cataracte qui consiste en une opacification partielle ou totale du cristallin. Dans les pays développés, une simple opération chirurgicale permet de retrouver une vision normale, mais ce n'est pas encore le cas dans les pays en voie de développement, ce qui fait que cette maladie représente encore quarante-huit pour cent des cas de déficience visuelle. Dans douze pour cent des cas il s'agit du glaucome qui est une maladie dégénérative du nerf optique souvent associée à une pression intraoculaire abîmant les fibres du nerf optique. Viennent ensuite la dégénérescence maculaire liée à l'âge dans neuf pour cent des cas puis les opacités cornéennes dans cinq pour cent des cas. Enfin la rétinopathie diabétique représente également cinq pour cent des cas et la cécité congénitale quatre pour cent des cas. Il existe une grande diversité de malvoyances dues aux différentes maladies qui peuvent porter atteinte au système visuel et à leur état d'avancement. Pour évaluer l'état de malvoyance d'une personne, il est nécessaire de mesurer l'acuité visuelle de chaque œil.

B. Évaluation de la déficience visuelle

L'acuité visuelle commune est la distance minimale pour discerner deux points distincts de contraste maximum. Il s'agit en général d'une fraction dont le numérateur est la distance d'observation et le dénominateur la distance minimum des points discernables multipliée par une constante. En France l'acuité visuelle s'exprime en dixièmes. La classification statistique internationale des maladies et des problèmes connexes, publiée par l'OMS donne une définition de la malvoyance selon l'acuité visuelle [OMS 2015].

- ♦ Catégorie 0 : Déficience visuelle légère ou inexistante. L'acuité visuelle est supérieure ou égale à trois dixièmes.

- ◆ Catégorie 1 : Déficience visuelle modérée. L'acuité visuelle est comprise entre un dixième et trois dixième.
- ◆ Catégorie 2 : Déficience visuelle sévère. L'acuité visuelle est comprise entre un vingtième et un dixième.
- ◆ Catégorie 3 : Cécité. L'acuité visuelle est comprise entre un cinquantième et un vingtième. La personne peut compter les doigts d'une main à un mètre.
- ◆ Catégorie 4 : Cécité. L'acuité visuelle est inférieure à un cinquantième, la personne peut percevoir de la lumière mais ne peut pas compter les doigts d'une main à un mètre.
- ◆ Catégorie 5 : Cécité. Absence de perception de lumière.

C. Handicaps dus à la cécité

La déficience visuelle a bien sûr des conséquences très importantes sur le quotidien des personnes atteintes. Le maintien de cinq capacités en particulier est mis en avant par l'organisation European Blind Union [Den Brinker 2003]:

- ◆ La lecture et l'écriture
- ◆ L'orientation et la mobilité
- ◆ Les activités quotidiennes
- ◆ La communication
- ◆ Le maintien d'activités visuelles

En 2005 l'Institut National Canadien pour les Aveugles a mené une enquête pour connaître les besoins de personnes aveugles ou déficientes visuelles [Simson et al. 2005]. Le fait de pouvoir se déplacer et s'orienter de manière autonome permet d'accéder aux services disponibles dans les villes et permet de plus de réduire l'isolement social qui touche de nombreux non-voyants. Cette étude indique que le besoin non satisfait le plus fréquemment cité est le déplacement. De plus, les auteurs notent que les services d'aide et de soutien les plus souvent sollicités par les personnes aveugles sont les services d'orientation et de mobilité. La capacité de se déplacer et de s'orienter est fortement liée à notre sens visuel, les personnes aveugles rencontrent donc de grandes difficultés pour comprendre leur environnement et s'y mouvoir ce qui impacte très directement leur autonomie.

D. Systèmes d'assistance ou de suppléance

Pour permettre aux personnes non-voyantes de conserver au maximum leur autonomie, différentes stratégies sont aujourd'hui proposées. Les premières aides dont bénéficient les personnes malvoyantes et qui se sont généralisées sont la canne blanche ainsi que le chien guide. Ces aides à la mobilité et à la navigation ont apporté un vrai bénéfice à l'autonomie des non-voyants. De nombreux systèmes ont été produits pour aider les non-voyants à retrouver de l'autonomie. Il faut distinguer deux catégories dans ces dispositifs. Les premiers sont les systèmes « génériques » qui visent à substituer ou restituer la vision dans sa totalité. La seconde

catégorie concerne les autres systèmes dits « spécifiques », ces systèmes sont destinés à aider une personne dans l'accomplissement d'une tâche spécifique [Jouffrais 2011].

Les dispositifs électroniques spécifiques d'assistance à la navigation au sens large se classent en deux catégories. Certains systèmes viennent compléter des aides déjà existantes et sont surtout présents pour assister à la mobilité, il s'agit des Electronic Travel Aids (ETA). Ces systèmes se concentrent principalement sur l'évitement d'obstacle et la sécurité. Ces dispositifs sont conçus pour compléter la canne blanche afin de détecter des obstacles en dehors de sa portée comme le projet EyeCane [Ju et al. 2009; Buchs et al. 2015], voir [Roentgen et al. 2008] pour une revue des aides à la mobilité. D'autres systèmes sont conçus dans l'objectif d'assister les non-voyants dans la navigation elle-même. Ces systèmes dénommés Electronic Orientation Aids (EOA) apportent une aide aux non-voyants pour trouver leurs itinéraires. Des dispositifs utilisant des balises RFID disposées en ville ou à l'intérieur d'un bâtiment permettent de guider les non-voyants lors de leur navigation [Willis & Helal 2005]. D'autres dispositifs reposent sur des systèmes d'information géographique (SIG) externes (comme Open Street Map¹) ou disposent de leur propre SIG afin de fournir une assistance à l'orientation des non-voyants [Völkel et al. 2008]. Nous pouvons citer le projet Drishti [Ran et al. 2004] ainsi que le dispositif NAVIG développé au sein de notre équipe [Katz et al. 2012]. Une revue de la littérature sur les dispositifs d'aide à la navigation a été rédigée par Giudice et Legge (2008).

Les dispositifs génériques peuvent eux aussi se distinguer en deux sous-catégories. La première inclut les dispositifs procédant à une substitution de la vision. L'information visuelle est captée par une caméra et restituée via une stimulation tactile ou auditive. Paul Bach-y-Rita et collaborateurs développèrent le premier système appelé Tactile-Visual Sensory Substitution (TVSS). Ce dispositif comprenait une chaise de dentiste sur laquelle était fixée une grille de quatre cents picots qui stimulaient le dos des sujets par vibration (Figure 1). Une caméra située à proximité de la chaise était directement manipulée par les sujets. Grâce à ce système, six premiers sujets aveugles sont parvenus à discriminer des formes simples et à reconnaître des objets après un apprentissage d'une vingtaine d'heures au minimum [Bach-y-Rita et al. 1969].

Ce même groupe a développé par la suite le Tongue Display Unit (TDU), ce dispositif disposait dans sa première version d'une matrice de quarante-neuf électrodes pour stimuler la langue [Bach-y-Rita et al. 1998]. Son fonctionnement est analogue au TVSS, le signal d'une caméra est traité afin de le transformer en un signal d'électrostimulation. Le TDU est un système mobile, ce qui permet aux utilisateurs de se déplacer. Une autre version de ce dispositif contenant une centaine d'électrodes a été utilisée dans une étude portant sur la navigation (Figure 2). Les sujets devaient traverser un couloir en U en évitant des obstacles placés sur leur chemin [Chebat et al. 2011]. La dernière version de ce système contient cent quarante-quatre électrodes [Kaczmarek 2011].

1 <https://www.openstreetmap.org> (Dernier accès le 03-07-2015)

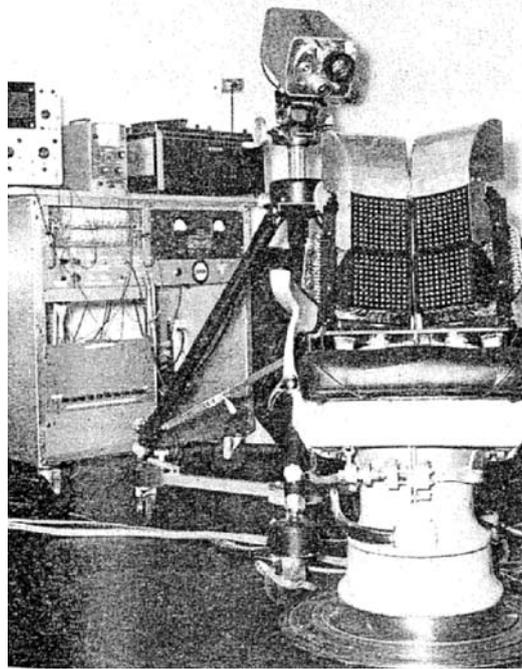


Figure 1 : Dispositif utilisé par Bach-y-Rita. La grille de stimulation est fixée sur le dossier de la chaise. La caméra est disponible à proximité de la chaise. Source : [Bach-y-Rita et al. 1969]

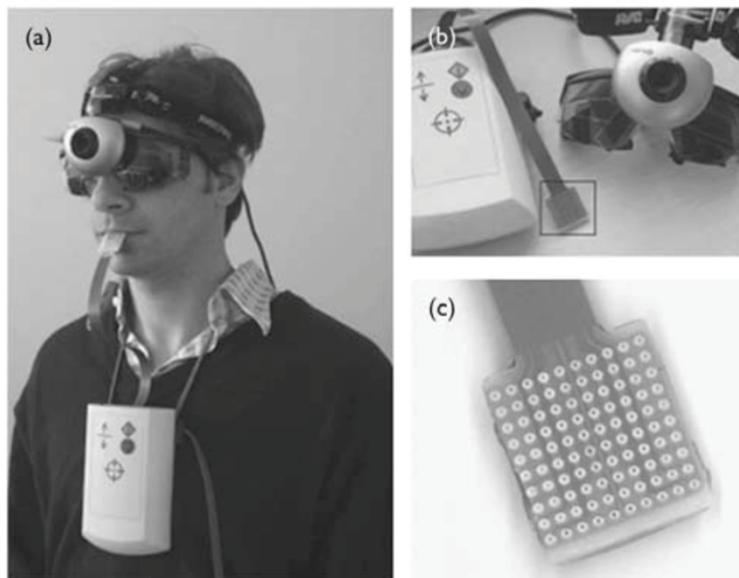


Figure 2 : Le Tongue Unit Display évalué en tâche de navigation. Le sujet porte une caméra dont le signal est traité pour fournir une séquence d'électrostimulation à la matrice d'électrodes. Source : [Chebat et al. 2011]

Un autre sens très utilisé dans la substitution sensorielle est l'audition. En 1992, Peter Meijer a dévoilé son système baptisé « The vOICE » [Meijer 1992]. Ce dispositif restitue une image en niveaux de gris chaque seconde selon le processus suivant :

- ◆ Une image est capturée par la caméra chaque seconde et traitée par le système.
- ◆ Pour la restitution, l'image est balayée de gauche à droite en une seconde. Ainsi le son entendu au début de la restitution correspond au bord gauche de l'image, celui entendu une seconde plus tard correspond au bord droit de l'image. Le rendu est également stéréo,

le son généré circulant de la gauche à la droite en une seconde. Un bref silence sépare le rendu sonore de chaque image.

- ◆ La fréquence du son traduit la position en ordonnée de l'élément. À un instant t , un son grave traduit la présence d'un élément en bas de l'image alors qu'un son aigu signale la présence d'un élément en haut de l'image.
- ◆ L'intensité lumineuse est restituée par le volume sonore.

En reconstruisant une image à partir du signal audio généré, il a été évalué que le système « the vOICe » dispose d'une « résolution » qui peut aller de neuf cents à trois-mille-six-cent pixels. Bien que ce système ait déjà plus de vingt ans, il continue à être évalué [Ward & Meijer 2010] et est utilisé quotidiennement par certains non-voyants². La principale limite de ce dispositif est l'apprentissage qui peut être très long avant d'en voir effectivement les bénéfices dans la vie quotidienne. De plus le rendu sonore masque les bruits environnants ce qui peut fortement gêner les sujets et cette stratégie ne permet pas de rendre plus d'une image par seconde, ce qui peut limiter ses applications, en particulier dans le cadre de la navigation.

D'autres systèmes ont été mis au point pour retranscrire une information visuelle via l'audition. En particulier le système PSVA élaboré par le groupe de Veraart à Bruxelles permet de restituer une image via un son en un temps de 100ms soit dix fois plus vite que le système the vOICe [Capelle et al. 1998]. L'idée théorique principale derrière ce prototype est de connecter un modèle du sens que l'on souhaite suppléer (ici la vision) à un modèle inverse du sens que l'on utilise pour la restitution (ici l'audition). Le fait de modéliser ainsi les différents sens impliqués dans la substitution sensorielle permettrait d'optimiser les systèmes en question. Comme le système the vOICe, le système PSVA utilise la fréquence pour coder la hauteur des pixels et l'amplitude pour coder la luminosité. En revanche, une distribution plus fine des fréquences audio est également mise en place pour distinguer la droite de la gauche. Ainsi les images ne sont pas balayées dans le temps et peuvent être rendues à une fréquence de 10Hz. Pour se rapprocher du modèle de la vision, il est proposé de densifier le centre de l'image en pixel pour faire une analogie avec la vision centrale et de laisser la périphérie de l'image avec une densité faible de pixel. Au final l'image est décomposée en 124 pixels ayant chacun une sinusoïdale propre. Il a été montré que ce système permettait à des sujets voyants de reconnaître des symboles simples de manière efficace avec seulement une dizaine de sessions d'entraînement d'une heure [Arno et al. 1999]. Il a également été montré que l'utilisation de ce système pour reconnaître des symboles impliquait chez des non-voyants l'utilisation du cortex visuel [Arno et al. 2001]. En particulier, une partie dorsale du cortex extrastrié (zone 18 de Brodmann) impliquée dans l'interprétation des images chez les voyants semble être utilisée par les non-voyants congénitaux lorsqu'ils utilisent le système PSVA pour reconnaître des symboles [Collignon et al. 2006]. Il est à noter que ce système a également été utilisé pour explorer des environnements virtuels afin d'y rechercher des informations de distance ce qui a également activé des zones dédiées à la vision chez des sujets voyants [Renier et al. 2005].

2 <http://www.seeingwithsound.com/>

Cette première catégorie de dispositifs génériques contient donc tous les systèmes permettant de restituer de l'information visuelle par une autre modalité sensorielle. La seconde catégorie contient les dispositifs de restauration de la modalité visuelle. Ces dispositifs sont des implants qui viennent stimuler électriquement le système visuel résiduel afin de provoquer l'apparition de tâches lumineuses appelées phosphènes. Un apprentissage est alors nécessaire pour que les personnes implantées s'approprient cette nouvelle vision et puissent l'utiliser.

Certaines neuroprothèses visuelles sont déjà utilisées quotidiennement par des personnes non voyantes qui s'expriment régulièrement dans les médias et indiquent parvenir à percevoir leurs proches ou à apprécier des feux d'artifices. En revanche peu d'informations sont disponibles sur des gains d'autonomie en ce qui concerne la navigation. Ceci peut être dû aux faibles résolutions des implants ou à d'autres paramètres comme les différents rendus proposés par l'implant. Afin d'évaluer les possibilités offertes par les implants actuels et à venir, il est possible de simuler cette vision prothétique à l'aide de l'outil informatique.

Les travaux menés durant ces trois ans de doctorat s'articulent autour de la vision prothétique simulée et de la navigation. Les connaissances nécessaires pour mener ces travaux sont présentées dans la première partie. Ainsi ce manuscrit présente dans son premier chapitre le fonctionnement du système visuel humain pour introduire les différentes neuroprothèses visuelles présentées dans le chapitre 2. Le principe de la vision prothétique simulée est présenté dans le chapitre 3 avec les différentes expérimentations effectuées grâce à cet outil. Le chapitre 4 précise ce qu'est une tâche de navigation et quelles sont les représentations mentales associées à cette tâche. La deuxième partie de ce manuscrit contient les différentes expérimentations menées. Les trois premiers chapitres de cette partie présentent les expérimentations menées sur la navigation en vision prothétique simulée. Le quatrième chapitre présente un questionnaire étudiant les représentations mentales que se font les voyants et non-voyants d'un trajet connu. La troisième partie contient le cinquième et dernier chapitre, les résultats des différentes expérimentations sont analysés pour en extraire des lignes directrices permettant d'élaborer des rendus pertinents pour la tâche de navigation en vision prothétique.

PARTIE 1 ÉTAT DE L'ART

Cette partie présente un état de l'art des différents domaines qui ont été mobilisés afin de mener à bien les travaux présentés dans la deuxième partie. Le fonctionnement du système visuel est d'abord présenté pour permettre d'expliquer clairement les différences entre les différents types de neuroprothèses visuelles décrites dans le chapitre suivant. À partir des descriptions faites par les sujets portant des implants, il est possible de mener des expérimentations portant sur différentes tâches quotidiennes en utilisant une simulation de vision prothétique. Ces différentes expérimentations sont présentées dans un chapitre et un soin particulier est porté aux tâches de navigation. La tâche de navigation est enfin caractérisée dans un chapitre portant sur les différents modèles psychologiques expliquant ce qu'est la navigation et ce qu'elle implique au niveau des représentations mentales utilisées.

Chapitre I

Le système visuel

Le système visuel humain est très complexe. Il est étudié extensivement depuis des décennies et nous ne ferons ici qu'une présentation succincte de son fonctionnement. La lumière traverse les différentes structures optiques de notre œil avant d'être captée par les photorécepteurs. Sous l'action des photons, les photorécepteurs génèrent et transmettent des signaux (via les cellules bipolaires) aux cellules ganglionnaires qui constituent le nerf optique. Les nerfs optiques amènent l'information visuelle jusqu'au corps genouillé latéral d'où elle est relayée jusqu'au cortex visuel dans le lobe occipital. L'information visuelle peut alors être traitée par le cerveau pour construire la perception visuelle.

A. Structure de l'œil et trajet de la lumière

Pour être perçu, un rayon lumineux doit tout d'abord traverser la conjonctive et la cornée situées à l'avant de l'œil (Figure 3). Ensuite vient l'humeur aqueuse qui sert de milieu nutritif à la cornée. La lumière traverse l'iris par l'ouverture de la pupille et rencontre le cristallin qui sert de lentille organique et permet de concentrer la lumière sur la rétine au fond de l'œil. Le globe oculaire est rempli d'une substance transparente et gélatineuse appelée le corps vitré, il est constitué à quatre-vingt-dix-huit pour cent d'eau et permet de maintenir une pression constante à l'intérieur de l'œil. Après le corps vitré, la rétine peut capter la lumière grâce aux photorécepteurs.

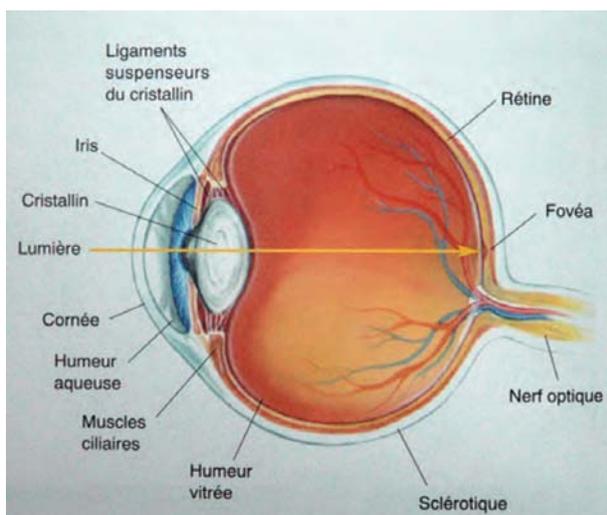


Figure 3 : Schéma de l'œil. Source : Jérôme Dupire

Quatre membranes principales permettent de structurer l'œil : la conjonctive, la sclérotique, la choroïde et la rétine. La conjonctive située entre l'œil et le milieu extérieur le couvre et se replie pour tapisser l'intérieur des paupières. La sclérotique qui prolonge la cornée est une membrane de couleur blanche : elle forme la paroi dure du globe oculaire. La choroïde située entre la sclérotique et la rétine permet d'alimenter la rétine et l'iris en nutriments. Cette membrane est fortement vascularisée et de couleur très

sombre, elle donne la couleur noire au fond de notre œil. Enfin la rétine est plaquée contre la choroïde grâce à la pression exercée par le corps vitré, sa structure est détaillée dans la partie suivante.

B. La rétine

La rétine est une membrane complexe, constituée de plusieurs couches de cellules (Figure 4). Les cellules photoréceptrices situées au fond de la rétine sont les cônes et les bâtonnets. Les cônes sont des cellules sensibles à une bande précise de longueur d'ondes de la lumière. Il en existe de trois types : les cônes sensibles au rouge, au vert et au bleu. Chaque œil contient entre cinq et sept millions de cônes. Les bâtonnets réagissent à une bande de fréquences plus large mais sont également beaucoup plus sensibles à la lumière et peuvent s'activer pour des intensités lumineuses très faibles. Chaque rétine dispose de cent dix millions de bâtonnets. Le signal émis par les photorécepteurs est transmis d'abord aux cellules bipolaires et horizontales avant d'arriver aux cellules amacrines et ganglionnaires. Les axones des cellules ganglionnaires se rassemblent à l'arrière de l'œil et forment le nerf optique.

Les photorécepteurs ne sont pas uniformément répartis dans la surface formée par la rétine et deux zones distinctes se démarquent dans la rétine :

- ♦ La rétine centrale contient en immense majorité des cônes et permet la vision des formes et des couleurs. Cette zone est extrêmement réduite et couvre environ un degré du champ de vision.
- ♦ La rétine périphérique a une densité de cônes beaucoup plus faible et contient essentiellement des bâtonnets. La vision en périphérie est donc beaucoup moins nette mais permet de détecter avec précision des changements d'intensité lumineuse.

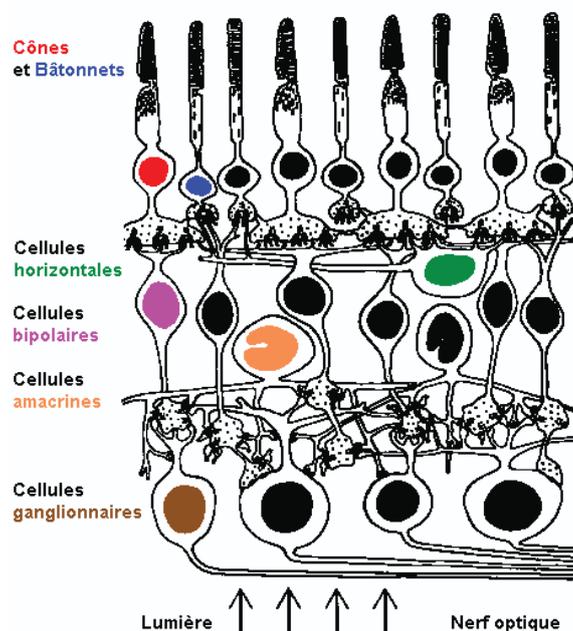


Figure 4 : Les différentes couches cellulaires de la rétine.

Source : <http://ophtalmologie.pro>

Comme le montre la Figure 4, la lumière doit traverser toutes les couches cellulaires de la rétine avant d'atteindre les photorécepteurs. Cela signifie que les axones des cellules ganglionnaires formant le nerf optique traversent la couche des photorécepteurs pour sortir de l'œil. La rétine dispose ainsi d'une petite zone ne contenant aucun photorécepteur, il s'agit

de la papille optique également appelée tâche aveugle. Par cette zone transitent les axones formant le nerf optique ainsi que les vaisseaux sanguins qui irriguent l'œil. L'évolution de la densité des photorécepteurs sur l'ensemble de la rétine est visible sur la Figure 5.

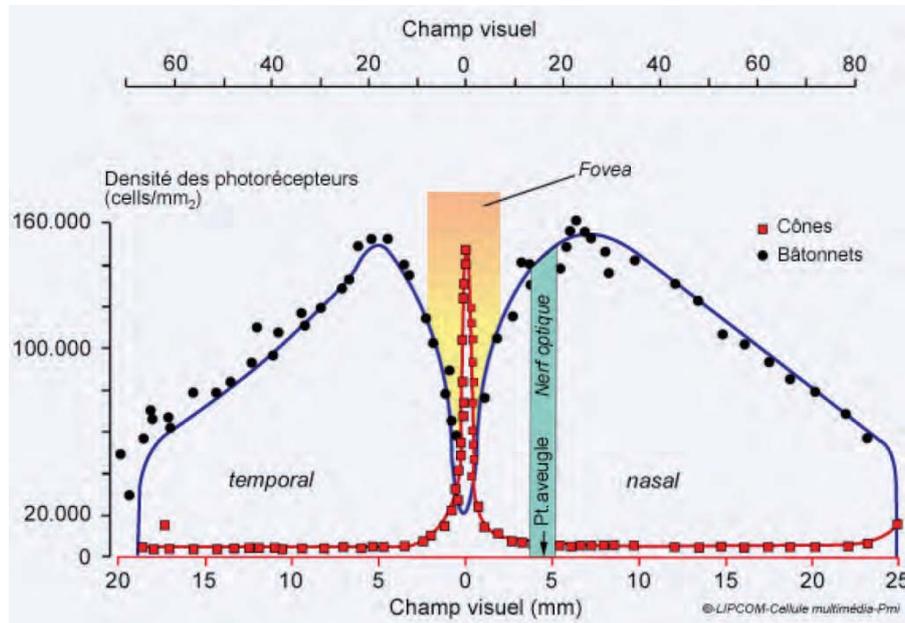


Figure 5 : Évolution de la densité des photorécepteurs sur l'ensemble de la rétine. La fovéa ainsi que le point aveugle sont visibles. Source : cours de R. Puiol

C. Le nerf optique

Le nerf optique est constitué d'un million deux cents mille axones des cellules ganglionnaires et transfère l'information à partir de la rétine. Au niveau cortical, l'information provenant de l'hémi-champ visuel droit (respectivement gauche) est traitée par l'hémisphère gauche (respectivement droit). Ainsi les deux nerfs optiques se divisent chacun en deux pour alimenter les deux hémisphères. Cette séparation s'effectue au sein du chiasma optique (Figure 6).

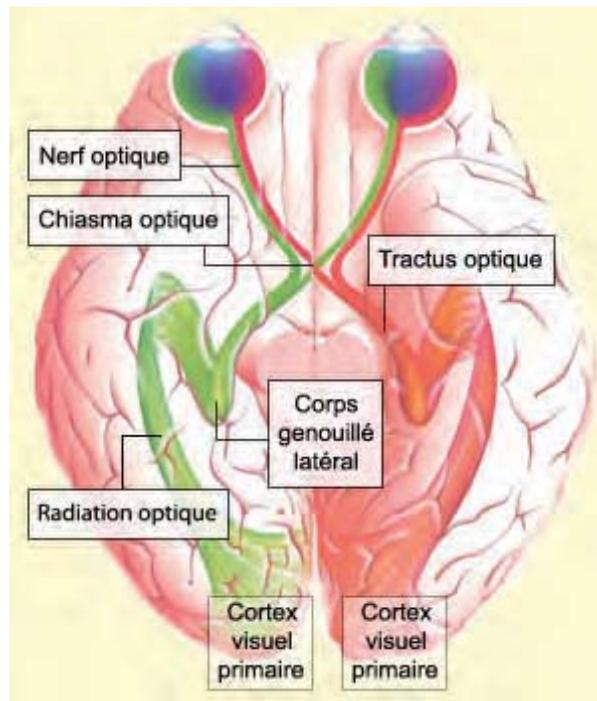


Figure 6 : Chemin parcouru par l'information issue du champ visuel gauche et droit pour chaque œil.
Source : <http://lecerveau.mcgill.ca>

À la sortie du chiasma optique, les axones se dirigent à travers le tractus optique pour alimenter le thalamus et plus précisément le corps géniculé (dit aussi genouillé) latéral. Ensuite les neurones de projection du corps géniculé latéral transmettent l'influx nerveux provenant du nerf optique au lobe occipital via les radiations optiques.

D. Le cortex visuel

Les neurones du corps géniculé latéral se projettent dans l'aire visuelle V1 située dans la partie postérieure du cortex occipital, de chaque côté de la scissure calcarine. Cette aire traite l'information de l'hémi-champ contralatéral et a la particularité d'être organisée de manière rétinotopique. Cela signifie que cette aire présente une structure analogue à la rétine même si une déformation est présente. Cette propriété de l'aire V1 est fondamentale pour les neuroprothèses visuelles comme nous allons le voir par la suite. Cette aire est connectée principalement à l'aire corticale visuelle secondaire (V2) qui l'entoure (Figure 7). Après cela, le traitement de l'information visuelle se complique considérablement, deux voies principales permettent de traiter l'information en parallèle. La première s'étend dans le lobe temporal et s'occupe notamment de la reconnaissance des objets, la seconde se projette vers le lobe pariétal et est essentielle à la reconnaissance des mouvements et à la localisation des objets.

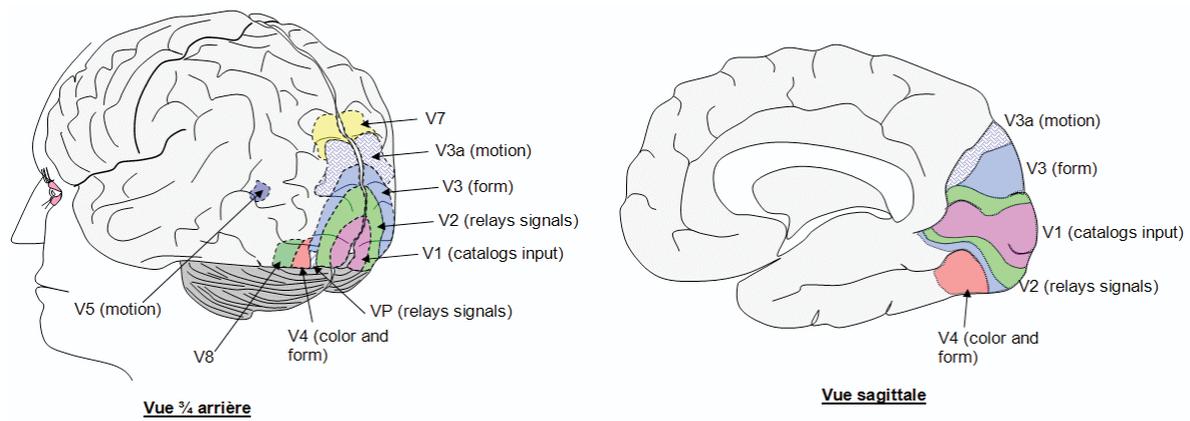


Figure 7 : Positions des aires corticales visuelles. Source : académie de Dijon

Ce chapitre a permis de présenter rapidement le système visuel dont les trois principaux constituants à retenir sont la rétine, le nerf optique et le cortex visuel. Ces trois éléments du système visuel sont les principaux lieux de stimulation pour réaliser des neuroprothèses visuelles.

Chapitre II

Neuroprothèses Visuelles

La connaissance même sommaire du système visuel a permis très tôt d'imaginer une restauration de la vision par stimulation électrique directe de ce système. Cependant, il a fallu attendre les progrès considérables effectués en électronique dans la miniaturisation des composants et la puissance de calcul pour réellement voir se développer des neuroprothèses visuelles utilisables quotidiennement.

A. Premières expériences

Dès le XVIII^{ème} siècle une tentative documentée a été faite par le médecin Charles Leroy pour tenter de restaurer par stimulation électrique l'ouïe, la vision et la mobilité de membres [Leroy 1755]. Dans cette première expérience, un sujet aveugle atteint de « goutte sereine » (ou amaurose) a subi plusieurs séries de stimulations électriques. Lors de la première série, durant laquelle le courant traversait tout le corps, le sujet a vu « comme une flamme qui paraissait passer rapidement et en descendant devant ses yeux ». La deuxième série de stimulations, se focalisant plus précisément sur la tête lui a permis de voir « une lumière bien plus forte que de coutume ». Encouragé par ces résultats, la dernière série de stimulations se concentrait exclusivement sur la tête en faisant partir le courant des yeux vers le lobe occipital de façon à ce que le courant « traverse la tête seule dans la route et la direction des nerfs optiques ». Lors de cette série, le sujet a indiqué voir « un peuple rangé devant lui et un spectacle admirable ». Il semble donc que lors de cette expérience, Charles Leroy a réussi à provoquer l'apparition de percepts lumineux par l'application de courants électriques.

Au début du XX^{ème} siècle, des chirurgiens allemands ont provoqué l'apparition de phosphènes en stimulant le cortex visuel de patients aveugles [Foerster 1929; Krause & Schum 1931]. Il a fallu attendre 1968, pour que deux scientifiques de l'université de Cambridge produisent un premier prototype de neuroprothèse visuelle [Brindley & Lewin 1968]. Ils ont implanté de façon aiguë une matrice de quatre-vingts électrodes à la surface du lobe occipital droit d'un patient de cinquante-deux ans pour stimuler le cortex visuel primaire. L'activation de ces électrodes a provoqué la perception de tâches lumineuses chez le patient, validant la possibilité de construire un implant visuel venant stimuler directement le cortex. L'américain William H. Dobelle a travaillé également sur le sujet dès les années 70 et a confirmé par ses travaux l'hypothèse de Brindley [Dobelle et al. 1974; Dobelle & Mladejovsky 1974]. Après avoir effectué des stimulations aiguës du lobe occipital, quatre volontaires aveugles ont été implantés, deux en 1974 et deux en 1978 [Dobelle 2000]. La FDA (Food and Drug Administration) n'avait pas donné son accord pour l'implantation de ces prothèses corticales et Dobelle s'exila au Portugal en 1983 pour y poursuivre

ses travaux. Seize nouveaux patients furent implantés par cette équipe peu avant le décès de Do-belle en 2004 mais ces implantations n'ont pas donné lieu à de nouvelles publications. Au final, peu de publications évaluent les performances de cet implant. Onze ans plus tard, il semble que ces travaux à la fois novateurs et controversés n'ont pas été poursuivis.

B. Projets d'implants et essais cliniques

Les implants corticaux ont été les premiers envisagés. La grande taille de l'aire visuelle V1 et sa rétinitopie en font un bon candidat pour l'implantation d'électrodes de stimulations. Une revue récente de ces implants est proposée par [Lewis et al. 2015]. En amont dans les voies visuelles, si le nerf optique est préservé il peut également être le lieu de stimulation. Enfin certaines pathologies atteignent spécifiquement des cellules de la rétine (comme la rétinopathie pigmentaire qui ne touche que les photorécepteurs). Pour les personnes atteintes de ces pathologies il est possible de stimuler la rétine et de faire percevoir des phosphènes. Trois types d'implants rétiniens ont été développés : les implants épi-rétiniens, subrétiniens et supra-choroïdiens (voir Figure 8).

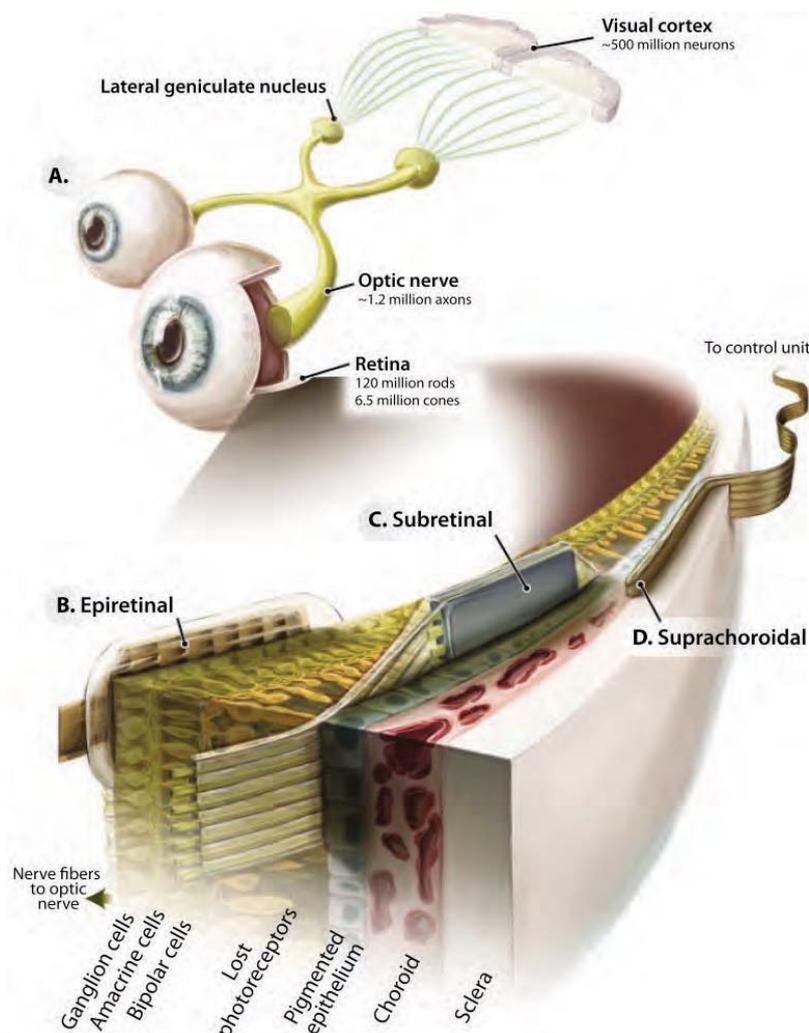


Figure 8 : Différents lieux d'implantation pour les neuroprothèses visuelles.

- A. Cortex visuel et nerf optique. B. Implants épirétiniens, au contact des cellules ganglionnaires.
 C. Implants sous-rétiniens, insérés à la place des photorécepteurs. D. Les implants supra-choroïdiens ont comme avantage d'être moins difficiles à implanter. Source : [Zrenner 2013]

B.1. Cortical

B.1.a. Monash Vision Group (Australia)

En 2010, un projet financé sur quatre ans fut initié en Australie sous le nom de Monash Vision Group. Ce projet regroupe une université, un hôpital et deux entreprises et étudie la possibilité de créer un implant cortical composé de sept à onze tuiles contenant chacune quarante-trois électrodes actives. Le nombre total d'électrodes pénétrantes utilisables serait situé entre trois cents et quatre cent soixante-treize [Lowery 2013]. Un système complet est visible dans la Figure 9, ses éléments seraient une caméra montée sur lunettes, un boîtier à six boutons traitant le signal de la caméra ainsi que l'implant lui-même, placé dans l'aire visuelle V1 du cortex. Pour le moment des travaux ont été faits sur le traitement de l'image vidéo en simulant la vision prothétique [Li et al. 2013]. Le financement de ce projet est arrivé à son terme à la fin de l'année 2014, avant que des essais cliniques ne soient lancés.

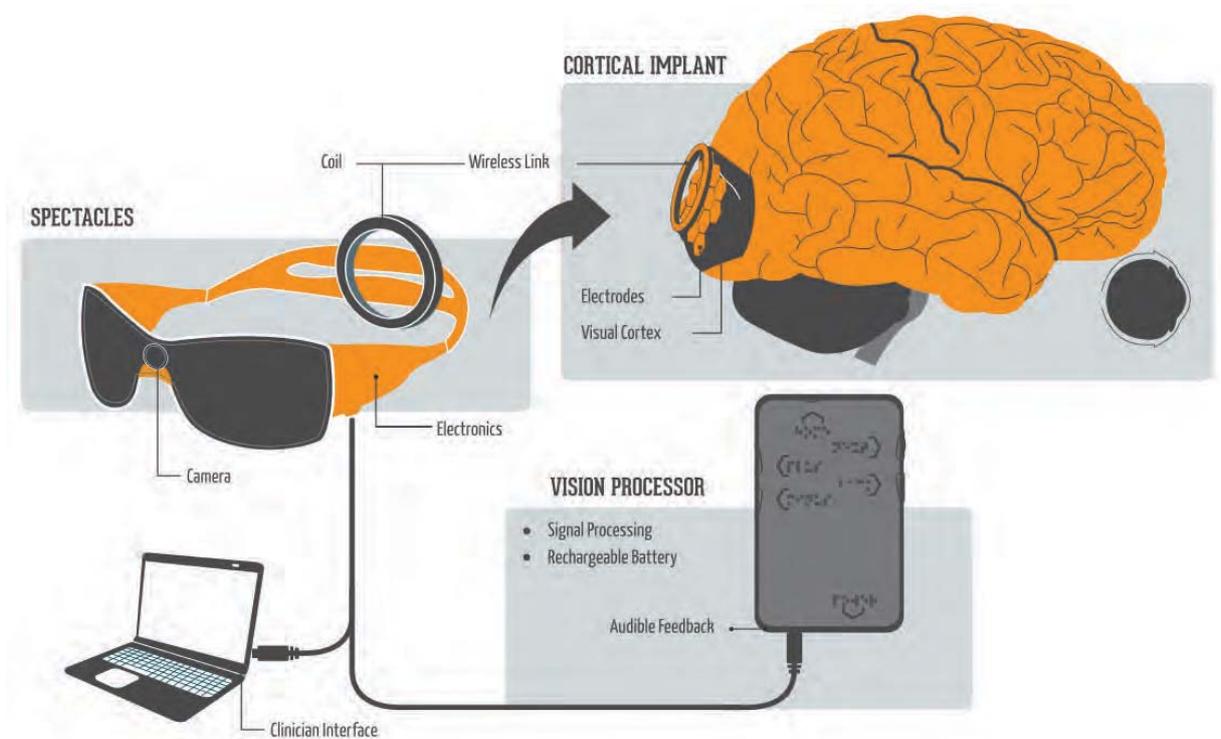


Figure 9 : Système en développement par le Monash Vision Group. Des lunettes contiennent une caméra, ce signal est traité par un processeur avec lequel les sujets peuvent interagir, l'information et l'énergie sont envoyées à l'implant grâce à une antenne située au niveau du lobe occipital. Source : [Monash Vision Group 2013].

B.1.b. National Institute of Health (États-Unis)

L'équipe du National Institute of Health située à Bethesda dans le Maryland est la première équipe à stimuler le cortex visuel avec des électrodes pénétrantes. Lors d'une première expérience, trois sujets épileptiques ont subi une intervention sous anesthésie locale afin d'exciser des foyers épileptogènes. Lors de ces trois interventions, l'équipe a pu insérer dans le cortex visuel des électrodes à différentes profondeurs afin de procéder à des stimulations. Les résultats d'un sujet ont montré qu'à une profondeur légèrement supérieure à 2mm, seulement

20 μ A étaient nécessaires pour évoquer des phosphènes alors que des électrodes de surface nécessitent une intensité cinquante fois plus intense [Bak et al. 1990]. Forts de ces résultats, cette équipe a procédé à l'implantation pendant quatre mois d'une patiente âgée de quarante-deux ans et aveugle depuis ses vingt ans [Schmidt et al. 1996]. Trente-huit électrodes ont été implantées et trente-quatre ont permis d'évoquer des phosphènes chez la patiente. Lorsque la stimulation était proche du seuil de stimulation, les phosphènes pouvaient apparaître jaunes, bleus ou rouges. Lorsque la stimulation était plus importante, ils devenaient blancs, gris ou jaunes. La taille des phosphènes variait d'un simple point à une pièce de monnaie tenue au bout du bras. Les phosphènes suivaient également le déplacement des yeux. De plus les électrodes de profondeur pouvaient être cinq fois plus rapprochées que les électrodes de surface sans provoquer de fusion entre phosphènes adjacents. Ces résultats ont amené l'équipe de Schmidt à conclure qu'il était possible de développer un implant cortical en utilisant des électrodes de micro-stimulations pénétrantes.

B.1.c. Intracortical Visual Prosthesis



Figure 10 : Implant du projet Intracortical Visual Prosthesis.
 Les modules utilisés sont visibles à droite sur fond jaune.
 Source : Institut Technologique de l'Illinois³

L'équipe de Philip Troyk et du laboratoire de prothèse neurale de l'institut technologique de l'Illinois a proposé également son implant cortical dénommé Intracortical Visual Prosthe-

3 <http://neural.iit.edu/>

sis. Cet implant est composé de différents modules chacun contenant seize électrodes. Ces recherches se sont concentrées en premier lieu sur l'animal. Un implant composé de vingt-quatre modules de huit électrodes fut implanté chez un macaque [Troyk et al. 2003]. Plus de quatre-vingts pour cent des électrodes en place ont provoqué l'apparition de phosphènes et toutes les électrodes fonctionnant à la première séance fonctionnaient encore à la dernière séance. Par la suite d'autres études ont été menées chez le singe [Bradley et al. 2005] ainsi que chez le chat pour évaluer les performances des électrodes de stimulation [Kane et al. 2013]. De plus cette équipe a commencé à travailler sur la conversion du signal vidéo obtenu par une caméra en stimulations délivrées par les électrodes [Srivastava & Troyk 2005]. Enfin ce groupe cherche à identifier les facteurs qui vont permettre de sélectionner les meilleurs candidats à l'implantation [Lane et al. 2012]. La Figure 10 présente l'objectif de ce projet de prothèse. Un des avantages de cet implant est que son aspect modulaire permet d'en réduire le coût.

B.1.d. Utah Electrode Array

L'équipe de Richard Norman à l'université d'Utah développe un implant contenant cent électrodes de 1,5mm de long visible dans la Figure 11. Initialement cet implant était prévu pour stimuler les aires visuelles afin de restaurer la vision [Normann et al. 1999].

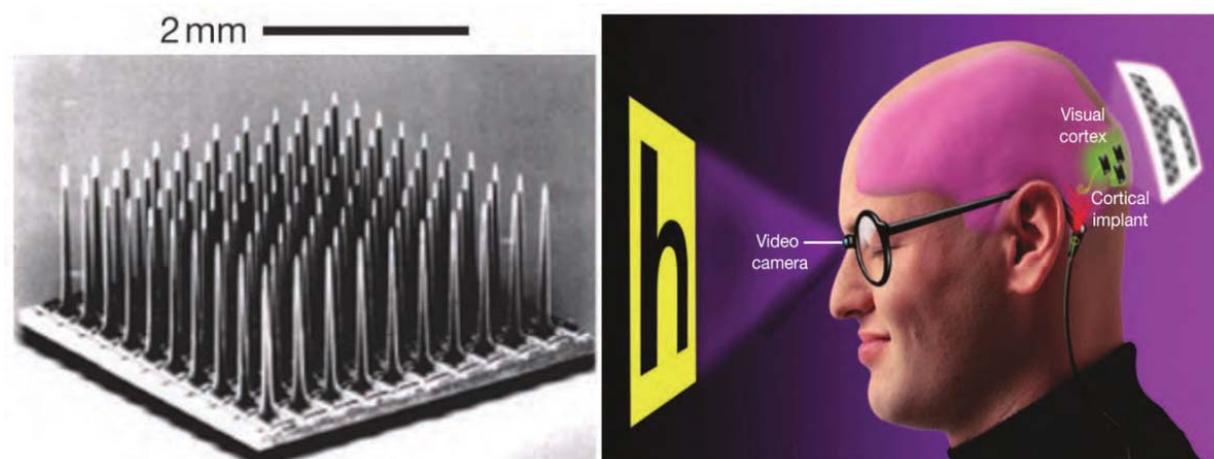


Figure 11 : À gauche : Utah Electrode Array, implant d'électrodes de microstimulations développé par Norman. À droite le système initialement prévu pour restaurer la vision chez les aveugles à l'aide de cet implant. Source : [Normann 2007]

La stimulation des aires visuelles a été testée chez le chat et le singe avec une évaluation des réponses comportementales. Ces études devaient permettre d'ouvrir la voie à des essais cliniques chez l'humain [Normann et al. 2009]. Cet implant a reçu l'accréditation de la Food and Drug Administration et des essais cliniques portant sur l'épilepsie ont été effectués chez l'homme en collaboration avec des universités espagnoles [Fernandez et al. 2014]. Pour le moment les résultats de ces recherches portent sur l'interface entre l'implant et le tissu neuronal, il n'y a pas encore d'essais cliniques portant sur la restauration de la vision chez l'homme.

B.1.e. École Polytechnique de Montréal (Canada)

L'école polytechnique de Montréal développe un implant cortical contenant plusieurs modules de seize électrodes chacun, les électrodes étant fonctionnellement regroupées par

quatre. Les avantages de l'approche par module cités par les auteurs sont divers : cette approche permet l'utilisation d'implants de petites tailles plus faciles à implanter et la multiplicité des modules permet de couvrir une plus large portion du champ visuel [Coulombe et al. 2007]. Les essais cliniques n'ont pas encore démarré pour cet implant. Cependant, cette équipe travaille également sur différents algorithmes permettant d'optimiser le signal. En particulier ils ont proposé un rendu permettant d'aider à la mobilité des personnes non-voyantes en indiquant les obstacles immédiats grâce à des cartes de profondeur [Mohammadi et al. 2012].

B.1.f. Projet Orion (Second Sight)

L'entreprise Second Sight est pour le moment la plus avancée sur le développement et la commercialisation des implants rétiniens. Cette entreprise a lancé également un projet d'implant cortical. Il a été annoncé une première implantation chez l'animal d'un prototype physique d'implant cortical dénommé Orion I⁴. Les évaluations fonctionnelles cliniques chez l'animal devraient se dérouler en 2016 pour commencer les essais cliniques chez l'homme durant l'année 2017.

B.2. Nerf optique

B.2.a. Veraart (Belgique)

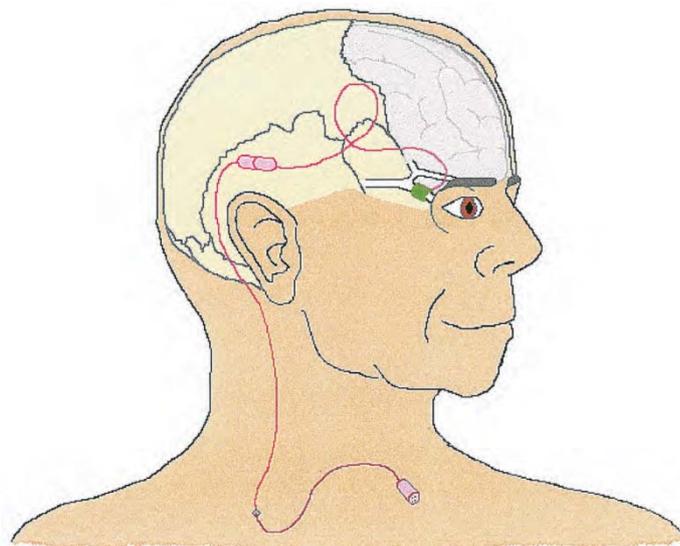


Figure 12 : Schéma de fonctionnement de l'implant développé dans le cadre du projet MiViP. Un connecteur transcutané assure la liaison entre l'implant et le dispositif externe de stimulation. Source : [Veraart et al. 1998]

En Belgique, le projet Mi-ViP coordonné par le professeur Charles Trullemans a permis le développement d'un implant stimulant le nerf optique. Cet implant est constitué d'une électrode à manchon spiral capable d'activer de manière sélective quatre zones différentes (Figure 12). Le professeur Claude Veraart, responsable de la dimension biomédicale du projet a effectué des tests pré-cliniques sur une patiente de cinquante-neuf ans, atteinte de rétinite pigmentaire et aveugle profonde depuis deux ans [Veraart et al. 1998]. Deux fois par

semaine, la patiente s'est rendue au laboratoire pour tester différents types de stimulations sur les quatre zones atteignables par l'implant. Au bout de quatre mois mille quatre cent soixante-cinq phosphènes en tout ont été répertoriés.

En 2000, la patiente a subi une nouvelle opération pour lui implanter une nouvelle version du système contenant une antenne afin que le dispositif interne puisse transmettre les informations de stimulation à l'implant via des ondes radio. En analysant les phosphènes évoqués en fonction des stimulations effectuées, l'équipe a mis au point un modèle prédictif pour que la stimulation évoque les phosphènes voulus à la position voulue. Ce modèle a introduit la notion de force de stimulation qui dépend de l'amplitude, de la durée, de la fréquence ainsi que du nombre d'impulsions dans le train de stimulations [Delbeke et al. 2003]. En utilisant ce modèle, l'équipe a fait une nouvelle expérience de façon à transmettre de l'information issue d'une caméra vidéo. Selon le modèle de stimulation, certains phosphènes nécessitent des trains de stimulation de durées plus ou moins longues pour apparaître. Pour transmettre un signal vidéo à l'utilisateur il est important de sélectionner les phosphènes qui peuvent apparaître le plus rapidement possible. Ainsi pour que l'implant puisse être utilisé avec une caméra vidéo [Brelén et al. 2005], le nombre de phosphènes exploitables est limité à cent neuf, leurs positions dans le champ visuel sont indiquées dans la Figure 13.

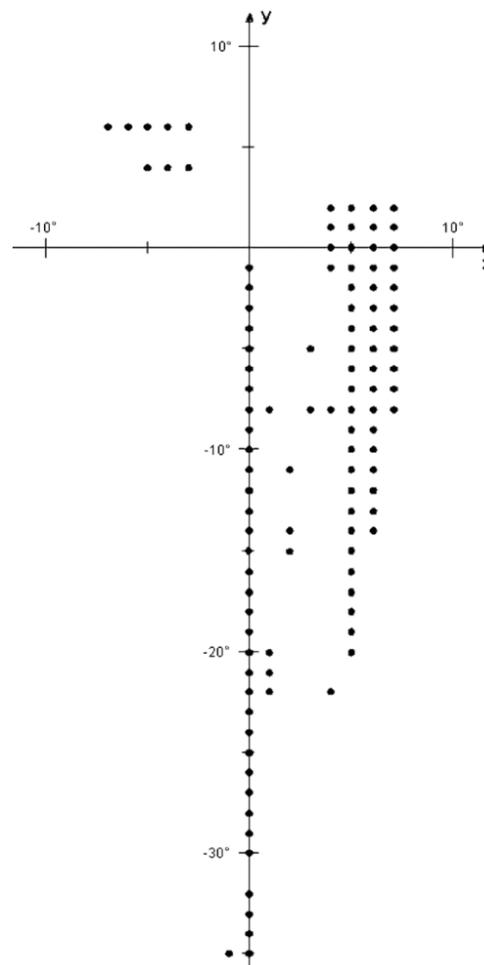


Figure 13 : Carte des phosphènes utilisés pour transmettre l'information visuelle d'une caméra au nerf optique. Chaque point représente la position d'un phosphène dans le champ visuel du sujet. Source : [Brelén et al. 2005].

En 2005 un patient âgé de soixante-huit ans a reçu une nouvelle version de l'implant développé par ce groupe. Cet implant contenait une électrode avec huit zones de contacts sur le nerf optique. Le lieu d'implantation était intraoculaire, facilitant ainsi la chirurgie [Brelén et al. 2006]. Les deux sujets ont participé à une évaluation plus fine des stimulations grâce à l'utilisation d'électrorétinogrammes et à l'enregistrement de potentiels évoqués (électroencéphalogramme) en regard des aires occipitales [Brelén et al. 2010]. Cette expérimentation a permis de chronométrer le temps précis entre la stimulation et l'apparition des potentiels évoqués à la surface du crâne. Ce groupe a prévu de faire des expérimentations chez l'animal pour valider son approche.

B.2.b. Chinese Project for Sight (Chine)

L'implant développé dans le cadre du projet C-Sight en Chine n'a pour le moment pas eu d'essai clinique chez l'homme. En revanche, des expériences ont montré une réponse corticale aux stimulations induites par leur implant sur le nerf optique chez l'animal [Chai et al. 2008]. Ces études ont permis d'optimiser la disposition des électrodes [Wu et al. 2010]. Il a également été observé qu'avec des électrodes de stimulation espacées de $150\mu\text{m}$ des réponses corticales peuvent être différenciées [Sun et al. 2011]. Contrairement aux autres projets, ce groupe prévoit d'implanter une micro-caméra directement à l'intérieur de l'œil plutôt que de la fixer sur des lunettes externes comme cela est montré dans la Figure 14. Avec une caméra externe fixée sur des lunettes, le champ caméra est asservi à la position de la tête et ne dépend pas de la position des yeux. Le fait d'implanter la caméra à l'intérieur de l'œil devrait permettre aux patients de contrôler plus aisément le champ caméra qui sera alors directement dépendant de la position de l'œil contenant l'implant. Le premier implant prévu chez l'homme par ce groupe serait une matrice composée de dix-huit électrodes [Ren 2014].

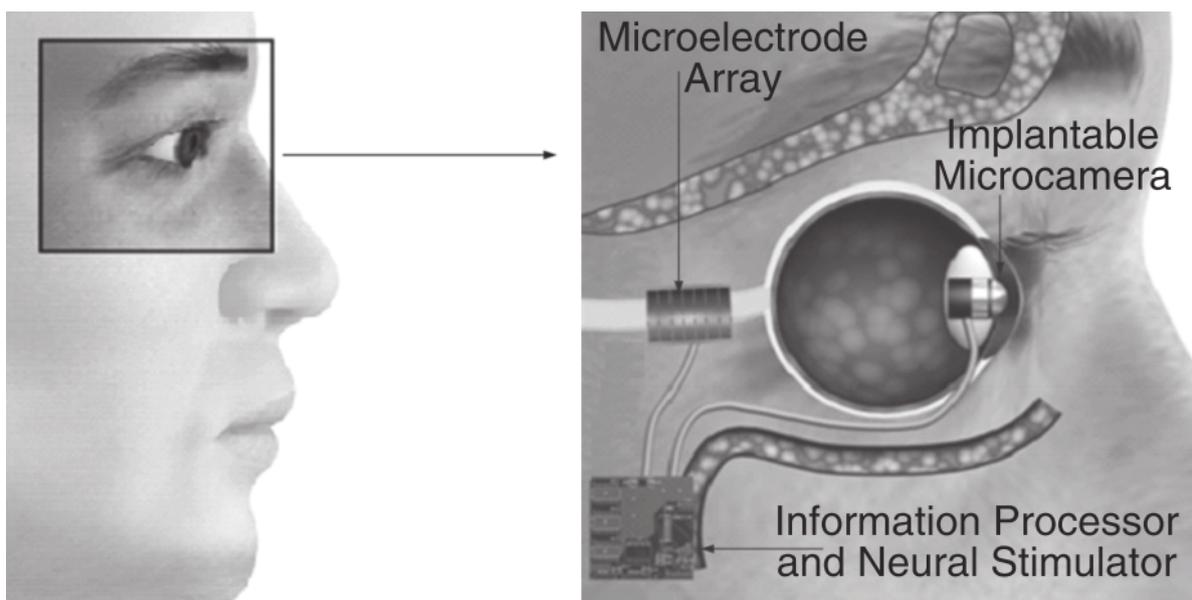


Figure 14 : Projet C-Sight d'implant du nerf optique. Une caméra implantée envoie l'information visuelle à un processeur qui transmet ensuite le signal adapté aux électrodes situés sur le nerf optique.

Source : [Chai et al. 2008]

B.2.c. AV-DONE (Japon)

Le Japon développe également un implant stimulant la papille optique (tête du nerf optique) via le projet AV-DONE (voir Figure 15). Après avoir effectué des essais précliniques chez le lapin [Sakaguchi et al. 2004; Fang et al. 2004]. Une première patiente âgée de trente-cinq ans a été implantée avec un prototype contenant trois électrodes de stimulation. Cette patiente a pu percevoir des phosphènes pour chacune des électrodes [Sakaguchi et al. 2009]. Une nouvelle version de cet implant contenant sept électrodes a été développée. Un autre essai clinique a eu lieu sur un patient de quarante-quatre ans [Nishida et al. 2015]. Lors de cette opération, dix minutes ont été nécessaires pour insérer l'implant contre plus d'une heure pour le premier essai clinique. Cette méthode rapide d'implantation semble intéressante, malheureusement neuf mois après implantation, une électrode était défectueuse et devenue inutilisable. Seize mois plus tard une autre électrode était également devenue inutilisable.

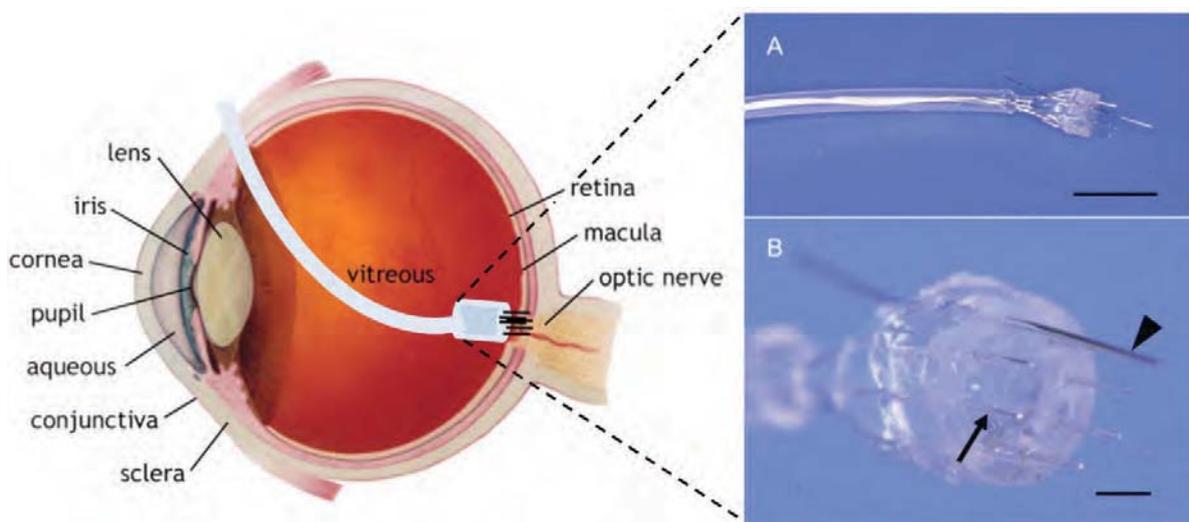


Figure 15 : Implant développé par le projet AV-DONE. Les électrodes sont insérées dans la tête du nerf optique. Sept électrodes de stimulation et une électrode de retour sont présentes. Source : [Nishida et al. 2015]

B.3. Rétine

Comme annoncé en introduction de ce chapitre, les implants peuvent être placés en différents lieux dans la rétine. Les implants épi-rétiniens se placent contre la rétine à l'intérieur de l'œil et sont donc en contact avec les cellules ganglionnaires qu'ils stimulent. Les implants sous-rétiniens sont fixés à la place des photorécepteurs défectueux et viennent directement stimuler les cellules bipolaires et les cellules horizontales.

B.3.a. Second Sight

L'entreprise Second Sight, située à Sylmar en Californie est l'entreprise qui dispose de l'implant dont le développement est le plus avancé. Cette entreprise a produit deux implants épi-rétiniens. Le premier dispositif, nommé Argus I et visible dans la Figure 16, contenait seize disques de platine servant d'électrodes. Ces électrodes de 520 μ m de diamètre étaient disposées en une grille de quatre lignes et quatre colonnes (voir Figure 17) espacées de 720 μ m. Un sujet implanté en 2003 est parvenu à percevoir dix nuances de gris sur chacune des électrodes.

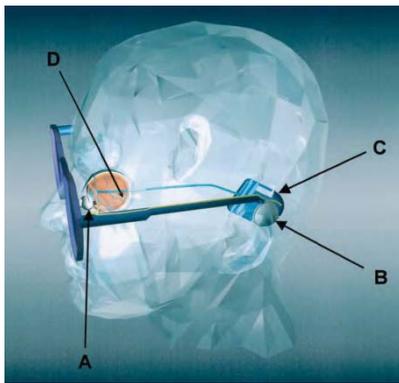


Figure 16 : Système complet de l'Argus I.
 (A) Caméra montée sur lunette; (B) Antenne de transmission; (C) Boîtier récepteur; (D) implant intraoculaire.
 Source : [Humayun et al. 2003].

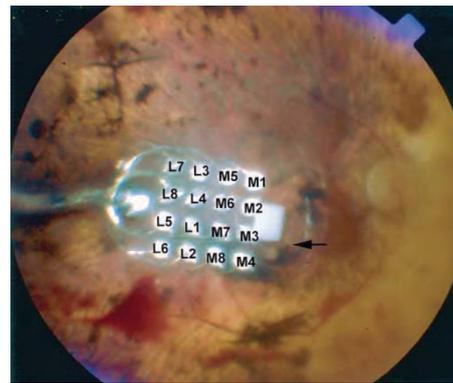


Figure 17 : Matrice d'électrodes située à l'intérieur de l'œil au contact de la rétine.
 Source : [Humayun et al. 2003].

Par la suite, une nouvelle version de cet implant a été mise au point avec la moitié des électrodes ayant un diamètre de $260\mu\text{m}$, l'autre moitié un diamètre de $520\mu\text{m}$. Entre 2002 et 2009 six sujets ont été implantés et ont réussi à repérer des changements de luminosité, à discriminer l'orientation de signaux transmis par leur implant, jusqu'à détecter des objets contrastés avec l'environnement [Caspi et al. 2009; Weiland et al. 2003; Weiland et al. 2004; Mahadevappa et al. 2005; De Balthasar et al. 2008].

L'implant de génération suivante, l'Argus II, est composé d'une matrice de soixante électrodes (6×10) de $200\mu\text{m}$ de diamètre, espacées de $525\mu\text{m}$. L'ensemble de l'implant couvre une surface visuelle de vingt degrés [Zhou et al. 2013]. Contrairement à l'Argus I, le signal de la caméra est envoyé vers un boîtier externe qui traite le flux d'information vidéo avant de l'envoyer à l'implant via une antenne située sur le côté externe de l'œil. Ce boîtier externe permet également aux utilisateurs de modifier en partie le signal de l'implant pour par exemple l'adapter à la luminosité ambiante.



Figure 18 : L'implant Argus II. L'antenne située au niveau de la tempe permet de récupérer l'information provenant du dispositif externe. Source : Second Sight⁵

L'Argus II est déjà commercialisé en Europe. Une étude clinique a été menée sur trente patients situés dans onze centres répartis à travers le monde. Cette étude avait comme principaux objectifs d'étudier l'acuité offerte par l'implant et de contrôler les risques liés à une implantation chronique. Les objectifs secondaires étaient d'étudier l'impact de l'implant sur les activités quotidiennes ainsi que d'évaluer la qualité de vie, l'orientation, la mobilité et la vision spatiale⁶ après implantation. La stabilité de l'implant et son fonctionnement ont également été surveillés pendant cette période.

Les résultats sont assez similaires à la tâche précédente, avec un taux de succès de vingt-cinq pour cent avec le système éteint et de cinquante pour cent avec le système en fonctionnement. Ainsi ces deux tâches restent assez difficiles à accomplir avec l'aide de l'implant. Ces deux tâches de mobilité sont représentées dans la Figure 19.

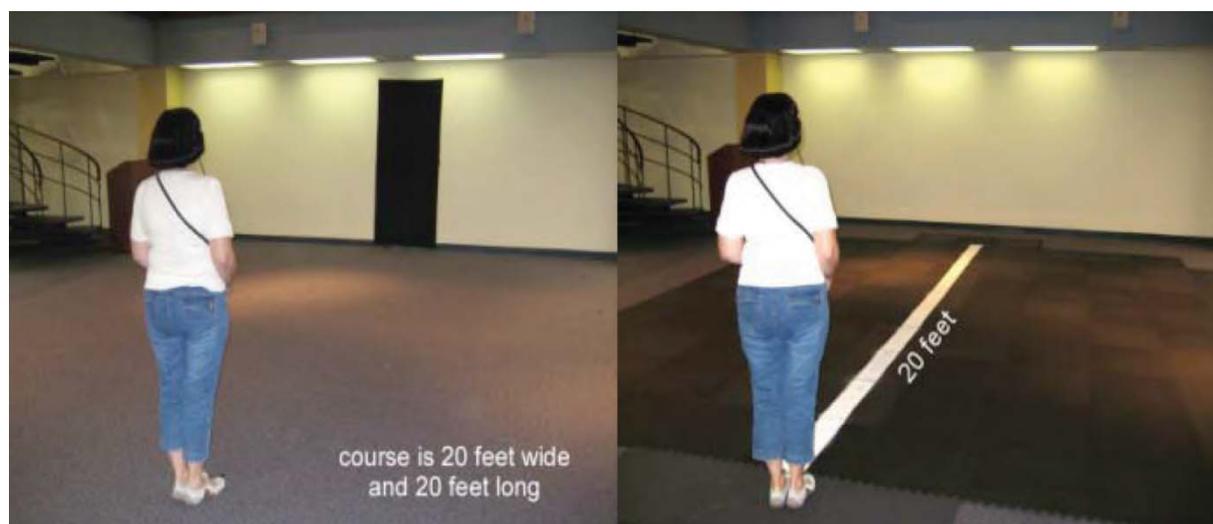


Figure 19 : Tâche de mobilité effectuée par les premiers patients ayant reçu l'Argus II. À gauche, les sujets doivent se diriger vers une porte sombre située sur un mur clair. À droite, les sujets doivent marcher le long d'une ligne blanche jusqu'au bout. Source : [Humayun et al. 2009]

En 2012 a été publiée une série de résultats intermédiaires sur les trente premiers sujets implantés [Humayun et al. 2012]. Ces résultats ont été compilés deux ans et demi après la première implantation et six mois après l'implantation la plus récente. Ils portent principalement sur le suivi médical des patients et l'évaluation de leur acuité visuelle. Ainsi sur les trente patients implantés, la meilleure acuité visuelle mesurée est de 20/1262, ce qui est resté bien en dessous du seuil légal français de cécité fixé à un vingtième. Les tâches de mobilité précédentes ont également été de nouveau évaluées et ces résultats sont présentés dans la Figure 20.

6 <https://clinicaltrials.gov/show/NCT00407602> (Dernière visite : le 17/08/2015)

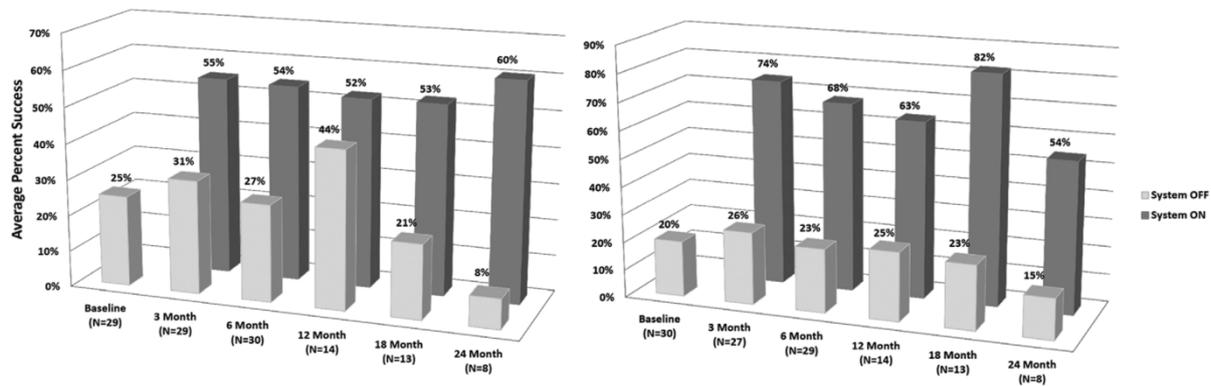


Figure 20 : Gauche : Pourcentages de succès pour l'orientation et la marche vers la porte. Droite : Pourcentages de succès pour le suivi de ligne. Source : [Humayun et al. 2012]

Ces trente premiers patients ont permis d'évaluer sérieusement l'apport de l'Argus II sur l'acuité visuelle [Dorn et al. 2012; Caspi & Zivotofsky 2015], la coordination visuo-motrice [Barry & Dagnelie 2012; Ahuja et al. 2011; Kotecha et al. 2014; Luo et al. 2014; Sabbah et al. 2014] et la lecture [da Cruz et al. 2013].

Le coût de l'implant a également été étudié [Vaidya et al. 2014] et durant l'été 2014, l'état français a décidé de rembourser chaque année le coût de ce dispositif pour une trentaine de personnes atteintes de rétinite pigmentaire⁷. En février 2015, Second Sight a annoncé que les trois centres français (situés à Paris, Bordeaux et Strasbourg) avaient commencé les implantations⁸. La commercialisation de l'implant a également été acceptée au Canada en Janvier 2015⁹. Des patients ont déjà été implantés à Toronto et Montréal.

B.3.b. Retina Implant et le projet Alpha-IMS

En 1995, le consortium « Subretinal Microphotodiodes » a été fondé avec à sa tête le professeur Eberhart Zrenner. Contrairement aux implants précédemment présentés, celui-ci ne dépend pas d'une caméra externe. Mille cinq cents photodiodes placées à l'emplacement des photorécepteurs captent la lumière pour la transmettre ensuite aux cellules bipolaires via une électrode. L'implant couvre 9mm², ce qui correspond à un champ visuel compris entre quinze et vingt-cinq degrés, et a une épaisseur de 70µm [Zrenner et al. 2009]. Les électrodes de stimulations ont chacune une surface de contact de 50µm x 50µm et sont espacées de 70µm. Seize autres électrodes dont la stimulation est contrôlable depuis l'extérieur et indépendamment de la lumière sont également présentes dans le prolongement de l'implant. La lumière reçue par les photodiodes n'apporte pas assez d'énergie pour stimuler directement les cellules rétiennes, cet implant dispose donc d'une alimentation sous la forme d'un boîtier externe relié à l'implant par un câble passant derrière l'oreille. Le boîtier externe permet d'agir sur les paramètres de stimulation (contrastes et la luminosité de l'image) grâce à deux potentiomètres. Une première étude portant sur neuf patients [Stingl et al. 2013] a montré que huit patients sur

7 Arrêté du 4 août 2014 relatif à la prise en charge au titre de l'article L. 165-1-1 du code de la sécurité sociale du système de prothèse épitréiniennne ARGUS II (<http://www.legifrance.gouv.fr>)

8 <http://www.businesswire.com/news/home/20150224005542/en/Sight-Announces-Commercial-Implants-Argus-II-Retinal>

9 <http://investors.secondsight.com/releasedetail.cfm?ReleaseID=890714>

neuf parvenaient à voir de la lumière, sept d'entre eux à la localiser, et cinq étaient capables de détecter des mouvements.

Vingt personnes dans le monde atteintes de rétinite pigmentaire ont été incluses dans un protocole d'expérimentation clinique et implantées pour une durée d'un an. En septembre 2014 au congrès annuel de la European Society of Retina Specialists (EURETINA), le professeur Zrenner a présenté les principaux résultats de cet essai clinique¹⁰. Cette étude a montré que quatre-vingt-six pour cent des patients ont perçu des phosphènes, cinquante-neuf pour cent des patients sont parvenus à détecter une source de lumière et quasiment la moitié d'entre eux ont rapporté avoir trouvé un intérêt à l'implant dans leur vie quotidienne. Les premiers résultats sur neuf patients sont sortis en 2014 [Kitiratschky et al. 2014] et en 2015 l'étude clinique portant sur les vingt-neuf premières personnes implantées a été publiée [Stingl et al. 2015]. Comme résultats particuliers, les auteurs insistent sur la localisation et la reconnaissance des objets qui sont améliorées par l'implant.

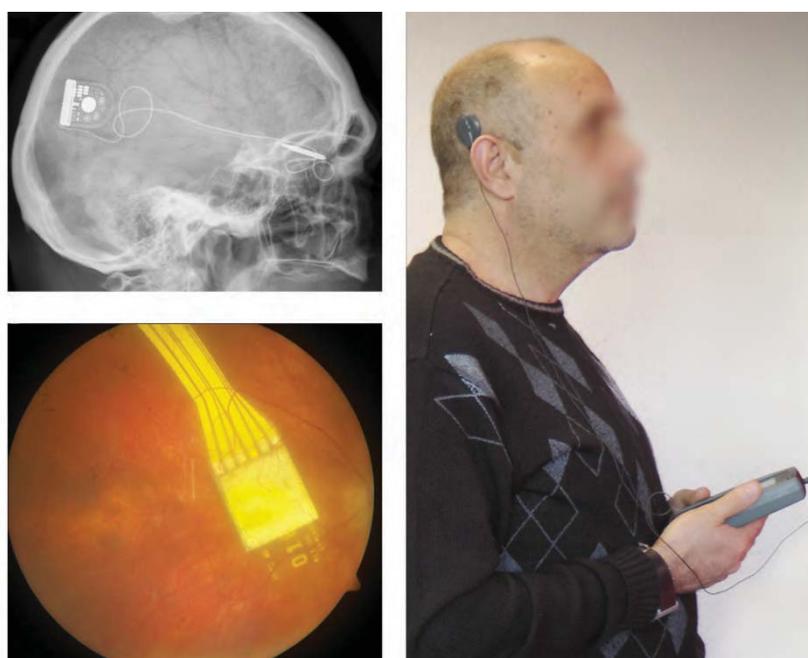


Figure 21 : Implant Alpha IMS. L'énergie nécessaire ainsi que le contrôle de la luminosité passent par une antenne située derrière l'oreille.

Source : [Stingl et al. 2015]

B.3.c. Boston Retinal Implant Project

Joseph Rizzo III et John Wyatt sont à la tête d'un consortium américain portant sur le développement d'un implant sous-rétinien, le Boston Retinal Implant Project. Après des études portant sur la stimulation de la rétine in-vitro [Grumet et al. 2000], des sujets volontaires ont pu avoir leur rétine stimulée lors d'opérations chirurgicales [Rizzo III et al. 2003b ; Rizzo III et al. 2003a]. L'élaboration d'un implant est maintenant en cours [Rizzo 2011]. Pour le moment aucun sujet humain n'a été implanté mais un implant contenant quinze électrodes a déjà été implanté chez des petits porcs [Shire et al. 2009][Kelly et al. 2009]. L'équipe se refuse à implanter

des sujets humains avec un implant contenant moins de deux cents électrodes, ce seuil étant pour eux un minimum à assurer pour avoir un gain fonctionnel réel chez les patients. Afin d'arriver à ce nombre d'électrodes, il est impératif d'optimiser l'espace occupé par le boîtier et la transmission d'énergie pour l'alimentation des électrodes.

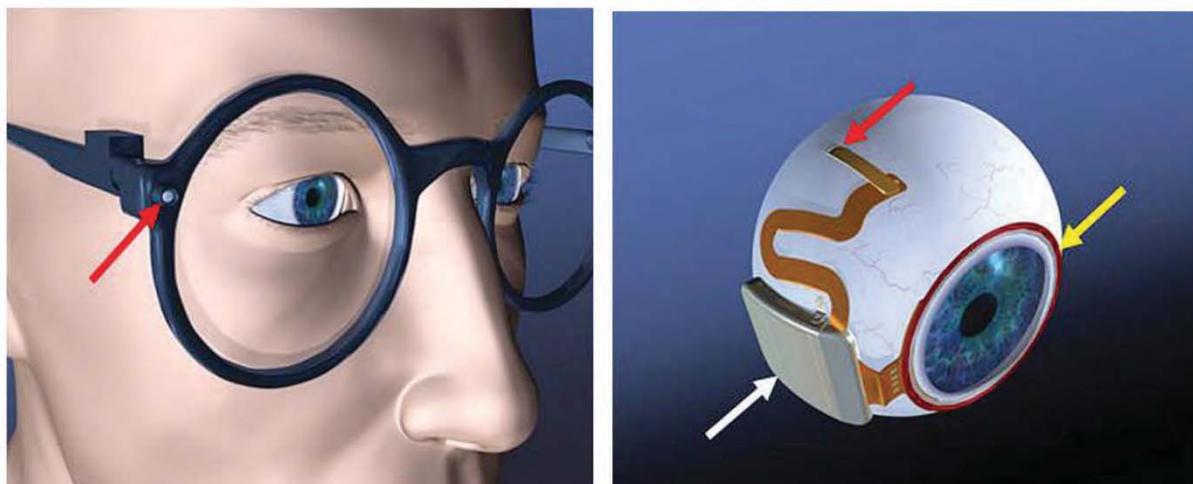


Figure 22 : Images de l'implant développé dans le projet Boston Retinal Implant Project. À gauche : les lunettes prévues pour l'implant, une caméra est située sur la branche des lunettes (flèche rouge). À droite : Le transfert d'information se fait via une bobine située sur la sclère (flèche jaune). Le dispositif de traitement de l'information est situé dans un boîtier en titane contre l'œil (flèche blanche). Seule la matrice d'électrodes est implantée à l'intérieur de l'œil (flèche rouge). Source : [Rizzo 2011]

B.3.d. Intelligent Medical Implant et Pixium Vision

En octobre 2003, l'entreprise allemande Intelligent Medical Implant a lancé des tests cliniques sur une vingtaine de personnes. L'implant épi-rétinien contenait quarante-neuf électrodes et est visible sur la Figure 23. Les patients pouvaient gérer eux-mêmes les paramètres de stimulation et ont rapporté voir des phosphènes de différentes formes et différentes couleurs, les plus fréquentes étant le blanc, le jaune et le bleu [Hornig et al. 2007].

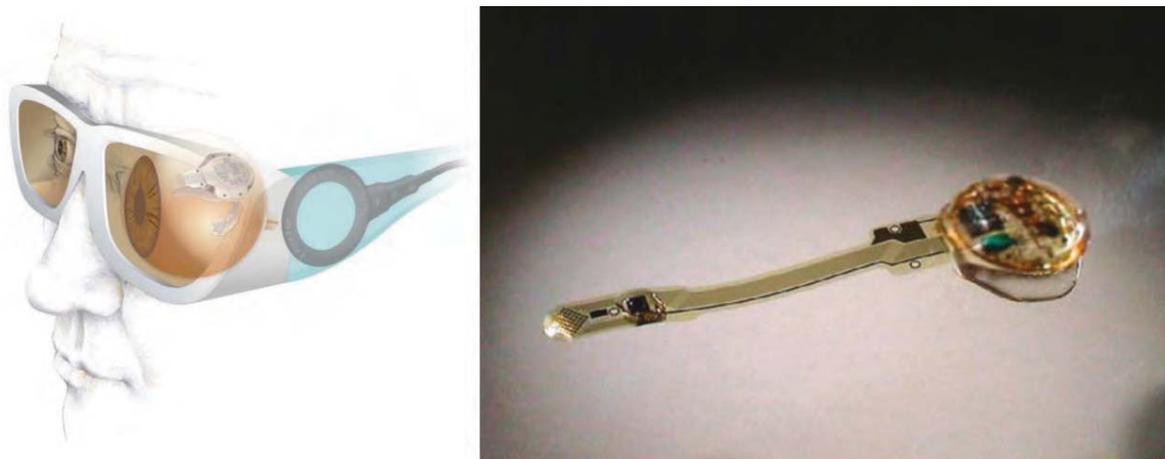


Figure 23 : À droite : Schéma de fonctionnement de l'implant IMI. À gauche : Implant IMI, la matrice d'électrodes est visible sur l'extrémité gauche du dispositif. Source : [Hornig et al. 2007]

En 2012, la société française Pixium Vision racheta IMI, et continua l'exploitation de cet implant désormais appelé IRIS[®]. De nouveaux essais cliniques sont en cours dont les résultats devraient être publiés en 2015. À la fin du mois de juillet 2015, la société a déclaré que cet implant devrait obtenir son marquage CE à la fin 2015 pour pouvoir être commercialisé en 2016¹¹. En parallèle, l'entreprise travaille sur l'implant IRIS[®]II qui devrait contenir cent cinquante électrodes.

Un autre implant est actuellement en développement chez Pixium Vision, il s'agit de l'implant sous-rétinien PRIMA utilisant des cellules photoréceptrices. Des expérimentations ont déjà été effectuées chez le rat [Mandel et al. 2013 ; Mathieson et al. 2012] et sont encore en cours pour évaluer les risques liés à l'implant. Si ces évaluations sont positives, les essais cliniques chez l'homme devront démarrer en 2016. À l'heure actuelle, Pixium Vision est le principal concurrent de la société Second Sight, ces deux entreprises ayant les implants les plus avancés au niveau de leurs évaluations cliniques et de leurs commercialisations.

B.3.e. Bionic Vision Australia

Le consortium australien Bionic Vision Australia dirigé par le professeur Anthony Burkitt se compose de deux programmes :

- ◆ Wide-View device : Le professeur Gregg Suaning est à la tête du développement de cet implant. En 2012 un premier essai clinique a été effectué avec l'implantation de trois patients avec un premier prototype suprachoroidal. Cet implant était composé de trente-trois électrodes de stimulation et de trois électrodes de retour dont une située derrière l'oreille. Seules vingt électrodes pouvaient être activées indépendamment les unes des autres [Ayton et al. 2014]. Des essais sont en cours chez le chat pour valider l'innocuité de l'implant dans le cas d'implantations à long terme [Villalobos et al. 2013; Villalobos et al. 2014]. Un nouveau prototype contenant quarante-quatre électrodes est en développement, des essais cliniques devraient démarrer sous peu.
- ◆ High-Acuity device : Ce projet est mené par Hamish Meffin et a pour objet de réaliser un prototype d'implant de haute résolution. Ce prototype contient deux cent cinquante-six électrodes en diamant pour un meilleur interfaçage avec le tissu organique (le diamant possède une très bonne biocompatibilité). Malheureusement, ces développements risquent d'être arrêtés faute de financements¹².

B.3.f. Artificial Silicon Retina (ASR)

Les frères Alan et Vincent Chow ont travaillé sur la piste de l'implant rétinien dès la fin des années 1990 [Chow & Chow 1997]. Au sein de la société Optobionics, ils ont mis au point un dispositif qu'ils ont appelé Artificial Silicon Retina. Cet implant se compose d'une micro-puce de 2 mm de diamètre et 25 μm d'épaisseur contenant approximativement cinq mille pixels. Chacun de ces pixels est isolé de ses voisins par un canal isolant de 5 μm . Chaque pixel

¹¹ <http://www.businesswire.com/news/home/20150728006610/fr/> (Consulté le 17/08/2015)

¹² http://bionicvision.org.au/news/story/vision_for_high-performance_bionic_eye_jeopardised_by_lack_of_funds

contient une micro-photodiode d'une surface de $400 \mu\text{m}^2$. La lumière incidente sur cette photodiode est transmise sous forme d'énergie à une micro-électrode de stimulation de $81 \mu\text{m}^2$ [Chow et al. 2004]. Une étude pilote portant sur six patients a permis de montrer que l'implant n'induisait pas de complications post-chirurgie et l'ajout de quatre patients supplémentaires a montré une amélioration de l'acuité visuelle chez certains d'entre eux. Par la suite, lors d'une deuxième phase du projet, quarante-deux patients ont été implantés, dont six d'entre eux pour une période de huit ans [Chow 2013]. Dans cet article, il est mentionné qu'après implantation, quelques sujets parviennent à retrouver une vision partielle. Certains d'entre eux parvenant à distinguer la luminosité du jour et de la nuit alors qu'ils n'en étaient pas capables auparavant et ceci sans l'aide de l'implant. Cette amélioration signifie que l'implant permettrait également de rétablir un fonctionnement partiel de la rétine. Des résultats similaires ont été obtenus avec l'Alpha-IMS [Stingl et al. 2015] : les différences de performance avec l'implant en fonctionnement ou à l'arrêt s'estompant avec le temps. Différentes pistes sont évoquées par ces deux équipes mais pour le moment aucune étude n'a porté spécifiquement sur ce sujet. La société Optobionics a été liquidée en 2007 mais Alan Chow a repris à son nom la société ainsi que les brevets liés à l'implant pour poursuivre ce projet¹³.

B.3.g. EpiRet GmbH

L'implant épirétinal Epi-Ret 3 a été conçu en Allemagne et a comme particularité d'être entièrement contenu à l'intérieur de l'œil du patient. Une bobine est située vers l'avant de l'œil pour recevoir les informations ainsi que l'énergie nécessaire aux stimulations. La matrice de stimulation est composée de vingt-cinq électrodes en relief ayant un diamètre de $100 \mu\text{m}$ et une hauteur de $25 \mu\text{m}$ [Mokwa et al. 2008]. La distance entre les centres de deux électrodes est de $500 \mu\text{m}$. Un schéma du dispositif est présenté dans la Figure 24.

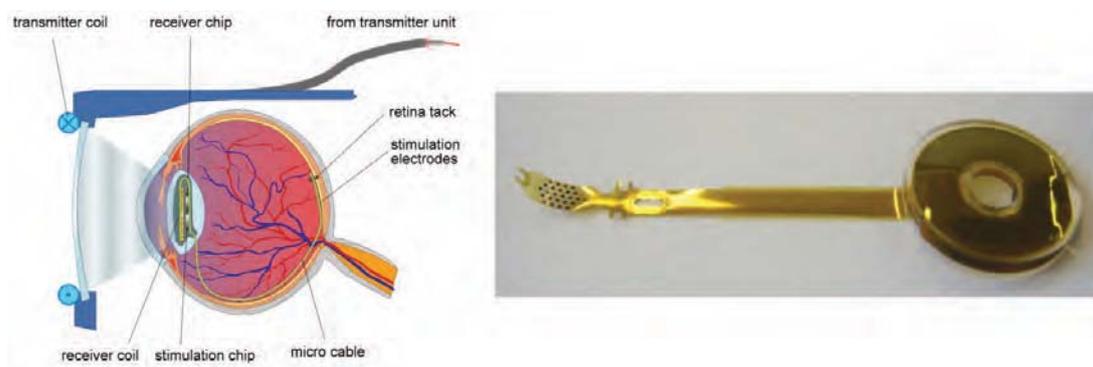


Figure 24 : Epi-Ret 3, À gauche : Schéma de l'implant, l'information et l'énergie sont transmises à l'implant via une bobine située à l'intérieur de l'œil. À droite : Photographie de l'implant, la bobine est reliée aux électrodes de stimulation visibles à gauche et placées contre la rétine. Source : [Mokwa et al. 2008]

Six patients ont été implantés pendant vingt-huit jours et des essais ont été effectués à trois reprises durant cette période [Klauke et al. 2011][Roessler et al. 2009]. Les patients ont rapporté percevoir des phosphènes et ont distingué certains motifs de stimulation comme un cercle ou

des lignes. Deux ans après les implantations les sujets ont été de nouveau sollicités et il a été évalué que l'implantation et l'explantation avaient été bien supportées [Menzel-Severing et al. 2012].

Ainsi trois principaux sites d'implantations des neuroprothèses visuelles sont envisagés : le cortex visuel, le nerf optique et la rétine. Les implants Argus II de Second Sight et IRIS de l'entreprise Pixium Vision sont les implants les plus avancés en termes de développement. Mais ils ne semblent pas parvenir au minimum requis pour permettre une restauration de la navigation, les sujets ne parvenant que difficilement à suivre une ligne blanche au sol ou à retrouver une porte noire de grande taille sur un mur blanc. Pour savoir quels sont les minimums requis pour aider à la navigation une possibilité est de simuler la vision prothétique et de faire des expériences en mettant à contribution des sujets voyants. De telles simulations sont utilisées depuis plus d'une vingtaine d'années pour étudier les apports fonctionnels des rendus prothétiques. Ces expérimentations sont l'objet du chapitre suivant.

Chapitre III

Simulation de la vision prothétique

A. Principe

Les nombreux essais cliniques liés aux neuroprothèses visuelles s'effectuent sur de longues durées et dans des protocoles très fermés. Les sujets implantés ne sont donc pas disponibles pour faire d'autres expérimentations et tester d'autres rendus potentiels des implants. Pour expérimenter sur le rendu des implants, il est cependant possible de simuler la vision prothétique. Cette simulation se base sur les descriptions faites par les sujets des phosphènes qu'ils ont perçus. En général un simulateur de vision prothétique se compose d'une caméra dont le flux vidéo est traité par un programme en temps réel, pour par exemple augmenter les contrastes. Une fois le traitement vidéo effectué, ce flux est envoyé à un simulateur de vision prothétique. Ce simulateur contient les caractéristiques de l'implant que l'on veut simuler (dimension, disposition des électrodes) ainsi que la forme des phosphènes voulus. Le simulateur de vision prothétique a comme fonction d'afficher un rendu qui correspond au plus proche possible à l'implant que l'on souhaite simuler, on parle alors de Vision Prothétique Simulée (VPS). Ce rendu est en général affiché dans un visiocasque qui est utilisé lors des expérimentations.

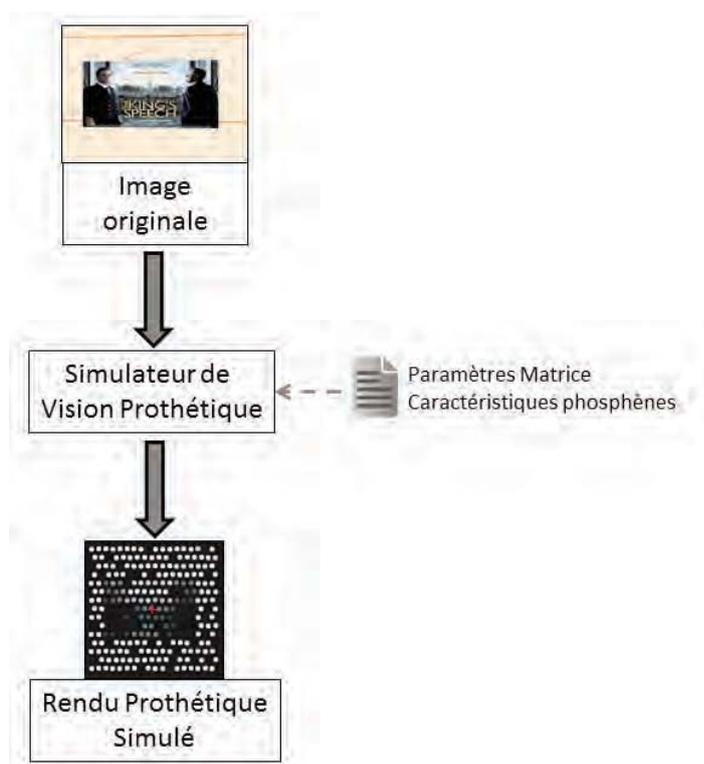


Figure 25 : Principe de la Vision Prothétique Simulée. L'image originale est traitée par un Simulateur de Vision Prothétique qui dispose des Paramètres de la Matrice (taille, disposition, *dropout*) ainsi que des caractéristiques des phosphènes (forme, nombre de niveaux de gris). Le simulateur crée alors un rendu qui peut être affiché ensuite pour les sujets via un moniteur ou un visiocasque.

Ce dispositif de simulation de vision prothétique est entièrement configurable et assure donc une grande souplesse dans son fonctionnement. Il est aussi aisé d'expérimenter sur des matrices de phosphènes correspondant aux implants actuels comme à des implants futurs. Cette facilité d'utilisation permet par exemple d'évaluer la résolution nécessaire pour lire à une vitesse acceptable ou reconnaître des visages. De même, pour une matrice de phosphènes de taille fixe, différents rendus peuvent être comparés en changeant les paramètres de la simulation.

La première expérimentation sur la simulation de vision prothétique qui a fondé le domaine a été effectuée par une équipe américaine en 1992 [Cha et al. 1992a]. Pour cela cette équipe a conçu un système mobile, contenant une caméra et un écran. La matrice de phosphènes était simulée grâce à une carte perforée placée entre les yeux du sujet et l'écran. Tout d'abord, Cha et son équipe ont évalué l'acuité visuelle obtenue par un implant avec des électrodes corticales profondes. Différentes résolutions ont été testées et leurs résultats indiquent que six cent vingt-cinq électrodes couvrant une aire visuelle de $1,7^\circ$ sont suffisantes pour atteindre une acuité visuelle de vingt trentièmes [Cha et al. 1992a]. La seconde étude portait sur la lecture et cette équipe a montré qu'avec le même nombre d'électrodes il était possible d'avoir un rythme de lecture de plus de 100 mots par minute [Cha et al. 1992c]. Finalement, cette équipe a évalué la mobilité des non-voyants en étudiant une tâche de navigation. Les sujets devaient trouver leur chemin dans un couloir coudé. Dans ce couloir le sol et les murs étaient peints en blanc alors que des obstacles sombres pendaient du plafond. Les sujets devaient parvenir à l'extrémité du couloir en un minimum de temps et en évitant au maximum les contacts avec les obstacles. Leur vitesse moyenne de marche ainsi que le nombre de contacts avec les murs et obstacles ont été utilisés pour évaluer leur performance [Cha et al. 1992b]. Comme pour les deux premières études, Cha et son équipe ont conclu qu'à partir de 625 électrodes et une aire visuelle de $1,7^\circ$ il était plus aisé d'éviter des obstacles et les sujets gagnaient en confiance pour se déplacer dans un environnement connu. Les résultats montrent également une amélioration des performances lorsque l'angle de la caméra est augmenté. Le simulateur ainsi qu'un exemple de trajet parcouru par les sujets sont visibles dans les Figure 26 et Figure 27.



Figure 26 : Simulateur de vision prothétique de Cha. Une caméra couplée à une lentille (C) envoie le signal vidéo directement à un écran qui contient un masque perforé pour simuler la vision phosphénique (M). La caméra et le moniteur sont alimentés par une batterie portée à la ceinture (B). Source [Cha et al. 1992a].

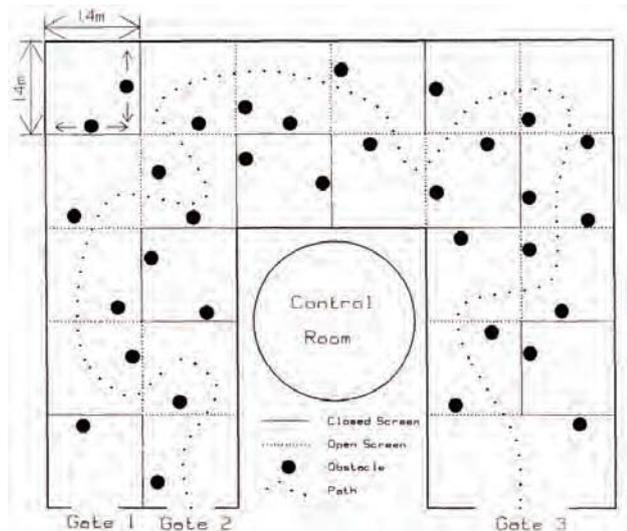


Figure 27 : Dispositif utilisé par Cha dans une expérience de navigation en vision prothétique simulée. Les traits pleins et traits pointillés indiquent la présence ou l'absence de parois en tissu. Les disques noirs sont des obstacles cylindriques suspendus à des rails. Les trajets démarraient à la porte 1 ou 2 pour se terminer à la porte 3. Un exemple de trajet est représenté en pointillé. Source : [Cha et al. 1992b]

Cette expérimentation en trois parties de Cha est importante pour la suite car elle est la première étude utilisant la vision prothétique simulée. Ces résultats ont servi de référence pour les études de vision prothétique simulée portant sur l'acuité visuelle, la lecture ainsi que la mobilité. Après cette première expérience, pratiquement 10 ans se sont écoulés avant que d'autres équipes de recherche ne s'emparent de cette question et réalisent des systèmes de simulation de vision prothétique plus aboutis (rendu des phosphènes plus fidèles, traitement numérique de l'image avant restitution).

La simulation de vision prothétique est utilisée comme outil depuis plus d'une vingtaine d'années pour évaluer les caractéristiques nécessaires des implants afin de permettre aux futures personnes implantées d'accomplir différentes tâches dans leur vie quotidienne. Ainsi cette technologie a permis de déterminer les caractéristiques minimales que doivent avoir les implants pour restaurer diverses fonctions visuelles. Au-delà de l'acuité visuelle, les tâches les plus étudiées et qui sont décrites par la suite sont la lecture, la reconnaissance et localisation de visages, la reconnaissance et saisie d'objets et enfin la navigation.

B. Acuité visuelle

L'acuité visuelle mesurée avec les patients implantés a aussi été évaluée en vision prothétique simulée à de multiples reprises. L'objectif de ces évaluations est principalement de fixer les minima requis pour obtenir une acuité visuelle au-dessus du seuil légal de cécité. En vision prothétique simulée, peu de phosphènes sont en général disponibles. Pour contrer cette

faible résolution, il est possible de bouger la tête et d'utiliser cette dynamique pour nous aider à identifier ce qui est devant nous. Ainsi l'acuité visuelle a été utilisée pour estimer l'impact du mouvement de tête [Chen et al. 2006; Chen et al. 2007]. De même, il est possible de traiter le flux vidéo pour par exemple augmenter les contrastes ou extraire les contours, certaines études portent sur le prétraitement de l'image et son impact sur l'apprentissage de la vision prothétique [Chen et al. 2005; Chen et al. 2004]. Comme la matrice elle-même est simulée, il est possible de jouer sur différents paramètres et mesurer leurs impacts sur l'acuité visuelle [Cai et al. 2005; Hayes et al. 2003; Rheede et al. 2010]. Les principales études portant sur l'acuité visuelle sont reportées dans le Tableau 1.

Tableau 1 : Principales études portant sur l'évaluation de l'acuité visuelle en vision prothétique simulée.

Publications	Tâches évaluées	Taille de la matrice	Forme des phosphènes	Objet de l'étude
[Cha et al. 1992a]	Acuité visuelle	10x10, 16x16, 25x25, 32x32	Points	Première étude en SPV
[Chen et al. 2004]	Acuité visuelle	10x10	Ronds (profil gaussien) 8 niveaux de gris	Comparaison Matrice hexagonale / matrice rectangulaire
[Chen et al. 2005]	Acuité visuelle Apprentissage	10x10	Ronds (profil gaussien)	Dispositions rectangulaire et hexagonale
[Cai et al. 2005]	Acuité visuelle	10x10	Ronds 8 niveaux de gris	Algorithme adaptatif pour le downsampling
[Chen et al. 2006]	Acuité visuelle	10x10	Ronds (profil gaussien)	Évaluer l'intérêt du balayage visuel, influence du mouvement de la tête
[Chen et al. 2007]	Acuité visuelle	10x10	Ronds (profil gaussien)	Influence du mouvement de la tête
[Rheede et al. 2010]	Acuité visuelle, Évaluation d'expression faciale, mobilité	30x30	Ronds, 8 niveaux de gris, fusion	Évaluation de différentes distorsions : Full Field, Region of Interest, Fisheye

C. Lecture

La lecture est une des tâches les plus étudiée en vision prothétique simulée, reflétant par là son importance pour les personnes implantées. Les personnes implantées jusqu'à maintenant ont toujours une cécité acquise; elles savaient donc lire avant la perte de la vision. Certaines études sur la lecture se sont portées sur la reconnaissance de caractère, de l'alphabet latin [Kiral-Kornek et al. 2014] ou des idéogrammes chinois [Zhao et al. 2011; Chai et al. 2007]. Le taux d'erreurs a été étudié pour déterminer qu'une matrice d'au moins 300 électrodes était nécessaire pour limiter le nombre d'erreurs sur des mots isolés [Sommerhalder et al. 2003]. Le même nombre d'électrodes est nécessaire pour permettre la lecture de paragraphe [Dagnelie et al. 2006a]. Pour la lecture en pleine page et en tenant compte de l'apprentissage, une matrice de 600 électrodes semble nécessaire pour arriver à une vitesse de lecture s'approchant d'une vingtaine de mots à la minute. Cette résolution permet également aux sujets de faire seulement 15 % d'erreurs sur les mots lus ce qui amène à une bonne compréhension des textes [Sommerhalder et al. 2004]. Le traitement de l'image a également été étudié pour optimiser la lecture [Pérez Fornos et al. 2005]. Pour des implants de faible résolution, une équipe a montré qu'il était possible de lire à une vitesse de 15 mots par minute avec seulement une quarantaine de phosphènes [Fu et al. 2006]. Cependant ce résultat a été obtenu en reconnaissant le texte au préalable pour afficher les lettres de manière optimale. L'équipe de Jörg Sommerhalder en Suisse a simulé le rendu obtenu avec l'implant Argus II de Second Sight, actuellement commercialisé et contenant 60 électrodes. Leur étude a montré qu'avec un tel implant il était théoriquement possible d'atteindre une vitesse de lecture de 34 mots par minute [Pérez Fornos et al. 2011], ce qui semble très loin des performances effectivement obtenues par les sujets implantés qui parviennent à reconnaître une lettre dans un temps voisin de la minute [da Cruz et al. 2013]. Cette différence de performance peut s'expliquer par le fait que la VPS est un rendu théorique et qu'il correspond toujours à un cas idéal, même si du bruit est ajouté. Les perceptions réelles des sujets sont plus complexes à interpréter et soumises à des variations bien plus importantes que les simulations complètement contrôlées.

Tableau 2 : Principales études portant sur l'évaluation de la lecture VPS.

Publications	Tâches évaluées	Taille de la matrice	Forme des phosphènes	Objet de l'étude
[Cha et al. 1992c]	Lecture	10x10, 16x16, 25x25, 32x32	Points	Première étude en SPV
[Dagnelie et al. 2001]	Reconnaissance visage / Lecture / Manipulation objets (en réel et virtuel)	10x10, 16x16, 25x25, 32x32	Ronds 2, 4, 6, 8 niveaux de gris.	Comparaison de rendus.
[Sommerhalder et al. 2003]	Lecture	83, 140, 286 et 875 pixels	Pixels	Lecture sur un écran.
[Sommerhalder et al. 2004]	Lecture	286 pixels	Pixels	Étude longitudinale, lecture de texte sur un écran.
[Pérez Fornos et al. 2005]	Lecture	28000, 1750, 572, 280 et 166 pixels.	Ronds (profil gaussien)	Comparaison d'algorithmes de redimensionnement.
[Dagnelie et al. 2006a]	Lecture et apprentissage, Coordination oculomotrice	6x10	Ronds (profil gaussien)	Lecture de paragraphes, Utilisation d'un jeu de dames
[Fu et al. 2006]	Lecture	6x6, 8x8, 10x10, 15x15, 20x20, 32x32	2 niveaux de gris.	Influence de la densité des phosphènes, de l'angle caméra
[Chai et al. 2007]	Reconnaissance d'idéogrammes chinois	6x6, 8x8, 10x10, 12x12	2 niveaux de gris.	Trouver la taille nécessaire pour reconnaître un idéogramme
[Dai et al. 2010]	Reconnaissance d'idéogrammes chinois.	12x12	Ronds	Étude de l'impact du bruit sur la position des phosphènes sur la lecture.
[Pérez Fornos et al. 2011]	Lecture	6x10	Pixelisation	Lecture pleine page, évaluation du nombre de pixels par caractère
[Zhao et al. 2011]	Reconnaissance idéogramme chinois	12x12	Ronds, 2, 4, 6, 8 niveaux de gris.	Étude de la distorsion, <i>dropout</i> , variation de taille, nombre de niveaux de gris.
[Kiral-Kornek et al. 2014]	Perception de lettres	16x16	Ronds (profil gaussien)	Étude de la distorsion, du <i>dropout</i> , d'un bruit sur la position des phosphènes. Utilisation d'algorithme de traitement d'images.

D. Reconnaissance et localisation de visages

Un autre domaine très exploré en VPS concerne la reconnaissance ou la localisation des visages. Les expressions du visage sont très importantes pour les interactions sociales [Ekman 1993], la communication contenant une grande part de communication non-verbale [Knapp et al. 2013]. La restauration de la perception des visages est donc primordiale pour les non-voyants. La première expérience de VPS portant en partie sur les visages date de 2000 et se concentrait sur la discrimination entre visages [Dagnelie et al. 2000; Dagnelie et al. 2001]. Cette première expérience a permis de déterminer la résolution minimale nécessaire à la reconnaissance de visages. D'après cette étude, les difficultés pour reconnaître des visages commencent lorsque la matrice d'électrodes contient moins de 100 électrodes, lorsque plus de la moitié des électrodes sont défectueuses, lorsque le nombre de niveaux de gris est inférieur à 4 ou lorsque les phosphènes sont trop gros ou trop espacés. D'autres expériences portent sur l'utilisation d'algorithmes de traitement d'images pour modifier les contrastes d'un visage afin de pouvoir le reconnaître plus facilement [Chang et al. 2012; Thompson et al. 2003; Wang et al. 2014b]. Pouvoir localiser les visages permet par exemple d'indiquer combien de personnes sont présentes à un non-voyant lorsqu'il entre dans un lieu. Le fait de détecter des visages grâce à des algorithmes de vision par machine et d'indiquer uniquement leurs positions via la vision prothétique a été expérimenté à plusieurs reprises [Denis et al. 2013a; Lui et al. 2012]. Yang et collaborateurs ont enregistré les potentiels évoqués par différentes images (dont des visages) en vision prothétique et ont étudié les spécificités de la vision prothétique [Yang et al. 2009]. Enfin en 2010 une équipe basée à Oxford a évalué la reconnaissance d'expressions faciales avec une matrice de 30x30 phosphènes [Rheede et al. 2010]. Les principales études portant sur la reconnaissance de visages sont reportées dans le Tableau 3.

Tableau 3 : Principales études portant sur la reconnaissance et la localisation de visages en vision prothétique simulée.

Publication	Tâches / Évaluation	Taille de la matrice	Forme des phosphènes	Objet de l'étude
[Dagnelie et al. 2001]	Reconnaissance visage / Lecture / Manipulation objets (en réel et virtuel)	10x10, 16x16, 25x25, 32x32	Ronds 2, 4, 6, 8 niveaux de gris.	Comparaison des rendus.
[Boyle et al. 2002]	Reconnaissance objets, visages et scènes	10x10, 16x16, 25x25	Ronds 2, 3 niveaux de gris	Algorithmes de traitement d'image.
[Thompson et al. 2003]	Reconnaissance visage	10x10, 16x16, 25x25, 32x32	Ronds 2, 4, 6, 8 niveaux de gris	<i>Dropout</i> : 10 %, 30%, 50%, 70%.

[Yang et al. 2009]	Reconnaissance de visages	27x32	Ronds	Neurophysiologie. Étude du déclenchement de la N170.
[Guo et al. 2010]	Reconnaissance de visages	14x16, 21x25, 27x32	Ronds	Étude de l'impact de la taille de la matrice. Sur la N170.
[Rheede et al. 2010]	Acuité visuelle, Évaluation d'expression faciale, mobilité	30x30	Ronds, 8 niveaux de gris, fusion	Évaluation de différentes distorsions : Full Field, Region of Interest, Fisheye
[Lui et al. 2012]	Concept de <i>Transformative Reality</i>	25x25	Ronds, binaires	Description d'une approche incluant différents capteurs, une modélisation du monde et une représentation symbolique.
[Chang et al. 2012]	Reconnaissance de visages	8x8, 12x12, 16x16	Ronds (profil gaussien), 10 niveaux de gris.	Algorithme de traitement d'images.
[Denis et al. 2013b]	Détection et localisation de visages	6x10	Ronds (profil gaussien), 8 niveaux de gris	Approche par localisation pour indiquer les visages
[Wang et al. 2014b]	Reconnaissance de visages.	24x24, 32x32	Points (profil gaussien), 8 niveaux de gris.	Différents algorithmes et stratégies de rendu.

E. Reconnaissance et saisie d'objets

Les études précédemment décrites portent sur les différents moyens d'acquérir le plus efficacement possible de l'information sur le monde en VPS. Le flux d'information va ainsi dans un seul sens, du système vers l'utilisateur. Un autre domaine est la manière dont une personne implantée peut interagir avec le monde qui l'entoure. Dès 2001, Dagnelie et son équipe s'intéressent à la coordination visuo-motrice en vision prothétique [Dagnelie et al. 2001]. La plupart des tâches étudiées étaient simples et dans un environnement très contrasté. Dans une étude de 2003 [Hayes et al. 2003], deux tâches de saisie d'objets ont été effectuées par les sujets. La première consistait à verser des friandises d'une coupe vers une autre. Ensuite les sujets devaient découper soigneusement un carré noir imprimé sur une feuille de papier. Pour la première tâche, les auteurs ont conclu qu'une résolution de 6x10 phosphènes était suffisante. Pour le découpage, une résolution de 16x16 phosphènes semblait nécessaire. Angelica Pérez Fornos a proposé en 2008 une autre tâche visuo-motrice [Pérez Fornos et al. 2008]. Durant cette expérience, il était demandé aux sujets de prendre en main un jeton avec une forme simple dessinée dessus. Ils devaient ensuite trouver le dessin correspondant à ce jeton sur un plateau situé devant eux et poser le jeton à cet endroit.

En 2006, Dagnelie et collaborateurs publient un article portant sur le placement de pions sur un jeu de dames, en vision prothétique, tête libre et avec un asservissement de la zone du champ visuel stimulé aux mouvements des yeux [Dagnelie et al. 2006b]. Une nouvelle expérience reprenant le même protocole, mais avec une simulation d'implant cortical a été ef-

fectuée en 2009 par cette même équipe [Srivastava et al. 2009]. La matrice était composée de six cent cinquante phosphènes avec un taux d'électrodes défectueuses variable. La détection et la coordination visuo-motrice étaient évaluées. Les sujets comptaient le nombre de cases blanches d'un échiquier partiellement masqué puis plaçaient des jetons sur ces cases.



Figure 28 : Tâche visuo-motrice en vision prothétique. Les sujets doivent regarder la forme imprimée sur l'objet qu'ils ont en main puis déposer l'objet sur la forme en question présente sur le tableau devant eux. Source : [Pérez Fornos et al. 2008]

En 2010, une équipe à l'université d'Oxford a publié un autre article sur la reconnaissance et la manipulation d'objets en vision prothétique. Dans cette étude, ils explorent également l'acuité visuelle et la navigation [Rheede et al. 2010]. Douze sujets ont participé au protocole et le simulateur de vision prothétique contenait également un oculomètre pour asservir la position de la matrice au regard comme montré dans la Figure 29. Pour toute cette expérience, la résolution de la taille de la matrice a été fixée à neuf cents pixels (30x30).

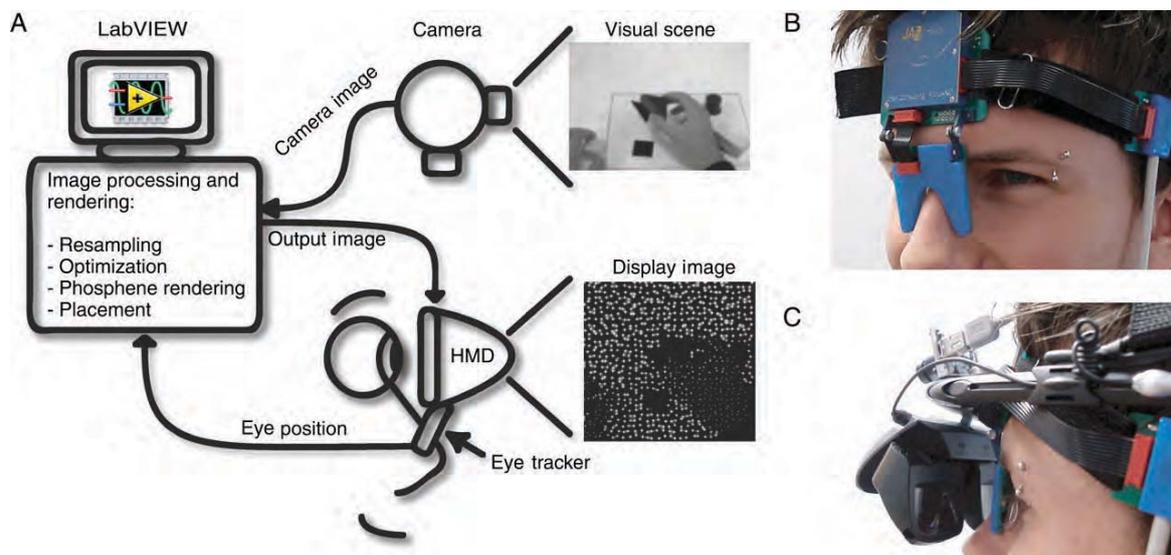


Figure 29 : Le système mis en place par Rheede et ses collaborateurs comporte un oculomètre ainsi qu'un visiocasque. Des tissus sombres étaient utilisés en plus pour masquer la vision périphérique des sujets. Source : [Rheede et al. 2010]

La tâche de manipulation s'inspire de celle utilisée par Fornos et al., [Pérez Fornos et al. 2008], et consistait à faire correspondre des pièces de bois taillées de différentes manières sur les formes imprimées correspondantes disposées sur une table. Trois rendus prothétiques ont été comparés dans cette étude. Le premier, nommé « Plein Champ » correspond aux rendus classiques, la résolution de l'image de la caméra est réduite pour correspondre à celle de la matrice de phosphène. Ainsi avec ce rendu, si le champ de la caméra couvre soixante degrés, la vision prothétique affichera ce champ de soixante degrés dans la zone de champ visuel couvert par la matrice, qui est ici de 8°. Le deuxième rendu prothétique nommé « Région d'Intérêt » est construit en sélectionnant dans le champ caméra de soixante degrés, un espace de huit degrés afin d'avoir une correspondance 1:1 avec la zone de champ visuel couvert par la matrice d'électrodes. Ainsi un sous-échantillon du champ visuel disponible est sélectionné pour être représenté avec une plus haute résolution par rapport à la condition « Plein Champ ». Enfin le dernier rendu appelé « Fisheye » cherche à combiner la large zone affichée dans le rendu « Plein Champ » avec la haute résolution obtenue dans le rendu « Région d'Intérêt ». Pour cela, les phosphènes situés au bord de la matrice affichent les bords du champ caméra, et un facteur de grandissement progressif est utilisé en se rapprochant du centre de la matrice pour arriver jusqu'à la résolution du rendu « Région d'Intérêt ». Pour l'étude de manipulation, il n'y a pas eu de différences significatives entre les rendus au niveau du pourcentage des objets bien placés. En revanche, les sujets avec le rendu « Plein Champ » ont obtenu une meilleure précision (de vingt-cinq pour cent) qu'avec le rendu « Région d'Intérêt » ainsi qu'un meilleur temps d'accomplissement de la tâche (de vingt-huit pour cent) qu'avec le rendu « Fisheye ».

En 2013, une autre étude a porté sur la localisation et saisie d'objets [Parikh et al. 2013]. Sept sujets ont participé à cette expérience et durent trouver un, deux ou trois objets situés sur une table devant eux. Deux rendus ont été évalués. Le premier était un rendu prothétique simple avec soixante phosphènes (matrice 6x10). Dans le second rendu, huit phosphènes supplémentaires étaient présents sur les bords de la matrice de phosphènes afin d'indiquer la direction des objets recherchés. Dans cette tâche, la présence d'indices a permis aux sujets de trouver les objets en effectuant deux fois moins de mouvements de tête et en étant deux fois plus rapides.

Un autre rendu de localisation a été évalué dans notre équipe [Denis et al. 2012]. Douze sujets devaient chercher un objet précis placé parmi neuf autres sur une table. Trois tailles de matrices (6x10, 15x18 ou 32x38 phosphènes) ont été utilisées avec un rendu classique. Pour le quatrième rendu, un algorithme de reconnaissance visuelle bio-inspiré [Dramas et al. 2010] a été utilisé pour localiser dans le champ caméra l'objet recherché. Une fois l'objet localisé, sa position était indiquée dans la matrice de phosphènes par un phosphène de luminance maximale, tous les autres phosphènes étant éteints. Ce rendu était utilisé avec une matrice de phosphènes de faible taille (6x10 phosphènes). Cette approche a permis d'avoir de meilleures performances que le rendu classique avec une matrice de résolution quatre fois supérieure (15x18). Il n'y avait pas de différences significatives entre les résultats avec le rendu par localisation en

6x10 et les résultats du rendu classique avec une matrice de taille 32x38. Ce rendu a été utilisé dans une autre expérimentation [Macé et al. 2015] avec quatorze sujets. Son utilisation a permis d'obtenir des performances équivalentes entre une matrice en croix contenant uniquement neuf phosphènes fonctionnels et une matrice contenant cent phosphènes.

Enfin plus récemment, Wang et collaborateurs ont utilisé un algorithme permettant d'isoler l'arrière-plan fixe d'un objet mobile au premier plan dans le flux d'une caméra vidéo. Ils ont ensuite utilisé cet algorithme pour faire reconnaître à des sujets des véhicules en mouvements, ou des actions effectuées par des personnes [Wang et al. 2014a].

Tableau 4 : Principales études portant sur la reconnaissance et la saisie d'objets en vision prothétique simulée.

Publications	Tâches évaluées	Taille de la matrice	Forme des phosphènes	Objet de l'étude
[Dagnelie et al. 2001]	Reconnaissance visage / Lecture / Manipulation objets (en réel et virtuel)	10x10, 16x16, 25x25, 32x32	Ronds 2, 4, 6, 8 niveaux de gris.	Comparaison des rendus.
[Boyle et al. 2002]	Reconnaissance objets, visages et scènes	10x10, 16x16, 25x25	Ronds 2, 3 niveaux de gris	Différents rendus comparés pour l'identification d'items fixes.
[Hayes et al. 2003]	Reconnaissance orientation, symboles, objets	4x4, 6x10, 16x16	Pixels, ronds (profil gaussien) 8 niveaux de gris	Études de l'impact de la taille de la matrice et du <i>dropout</i> sur l'acuité et la reconnaissance d'objets.
[Dagnelie et al. 2006b]	Lecture et apprentissage, Coordination oculomotrice	6x10	Ronds (profil gaussien)	Lecture de paragraphes, Utilisation d'un jeu de dames
[Pérez Fornos et al. 2008]	Reconnaître des symboles, manipuler des objets.	17920, 1991, 498, 221 et 124 pixels	Pixels	Influence du champ caméra sur des tâches de manipulation
[Zhao et al. 2010]	Reconnaissance d'objets	8x8, 16x16, 24x24, 32x32, 48x48, 64x64	Carrés, Ronds	Évaluation de différents algorithmes de traitement d'image
[Li et al. 2011]	Reconnaissance d'objets	8x8, 16x16, 24x24, 32x32, 48x48, 64x64	Carrés, Ronds	Étude de l'impact de la taille de la matrice.

[Denis et al. 2012]	Saisie d'objets	6x10, 15x18, 32x38	Ronds (profil gaussien), 8 niveaux de gris.	Approche par localisation pour indiquer l'emplacement d'un objet recherché.
[Hu et al. 2013]	Discrimination d'objets	8x8, 16x16, 24x24, 32x32	Ronds (profil gaussien), 2, 4, 6 et 8 niveaux de gris	Influence de la résolution, des niveaux de gris, d'une distorsion, du <i>dropout</i> .
[Lu et al. 2013]	Identification d'objets	16x16, 24x24, 32x32	Ronds (profil gaussien), 2, 4, 6 et 8 niveaux de gris.	Influence de la résolution, distorsion, <i>dropout</i> , nombre de valeurs de gris.
[Parikh et al. 2013]	Mobilité, saisie d'objet, identification de cibles.	6x10	Ronds (profil gaussien), 8 niveaux de gris	Rendu indiquant la présence d'obstacles. Utilisation de la localisation.
[Wang et al. 2014b]	Identification de scènes	24x24, 32x32	Ronds (profil gaussien), 8 niveaux de gris	Comparaison de différents algorithmes de traitement vidéo pour isoler les éléments en mouvement de l'arrière-plan.
[Macé et al. 2015]	Saisie d'objets	9 (en croix), 10x10	Ronds, 2 niveaux de gris.	Utilisabilité de l'approche par localisation pour saisir des objets. Comparaison entre deux tailles et deux positionnements de matrices (central ou latéral).

F. Mobilité et navigation

La mobilité est une composante importante de l'autonomie des non-voyants et des études de VPS ont porté sur cette question, plus particulièrement sur la détection et l'évitement d'obstacles. Une des premières études de Cha décrites plus haut portait déjà sur la mobilité [Cha et al. 1992b]. La thèse de Jason Dowling à l'université de Queensland en Australie a porté sur la mobilité en vision prothétique [Dowling 2007]. Dans une expérience de compréhension de scènes fixes, huit images montrant des scènes que l'on peut rencontrer en mobilité étaient présentées à des sujets en vision prothétique. Au préalable, les contours des objets de ces images avaient été mis en surbrillance. Dowling utilisa quatre types de traitement différents et présenta les images en vision prothétique avec une « haute » définition contenant deux mille cinq cents phosphènes (matrice 50x50). Le traitement consistant en une simple désaturation de l'image fut celui qui permit le mieux au sujet d'identifier ce qui leur était présenté [Dowling et al. 2004]. D'importants écarts existaient entre les différentes images avec des performances d'identification allant de quarante pour cent à quatre-vingt-quinze pour cent suivant l'image de départ, ce qui amène à penser qu'il existe sans doute différents rendus optimaux pour différentes situations. Dans une autre expérimentation les sujets disposaient d'un écran devant les yeux et devaient parcourir un trajet à l'intérieur du laboratoire [Dowling et al. 2005]. Trois rendus prothétiques furent testés, tous de la même résolution (32x24 phosphènes). Deux cent cinquante-six niveaux de gris étaient utilisés dans un rendu tandis que les deux autres disposaient de huit niveaux de gris. Pour l'un de ces derniers, un alerteur visuel

était intégré dans le système pour indiquer la présence d'un obstacle, en cas de détection, les phosphènes correspondants à l'obstacle devenaient roses. Ni l'alerteur ni les rendus n'ont eu d'impact significatif sur les performances des sujets qui étaient calculées à partir de la vitesse moyenne sur une dizaine de mètres.



Figure 30 : Casque utilisé pour la tâche de mobilité dans l'expérience de Dowling. Source : [Dowling et al. 2005]

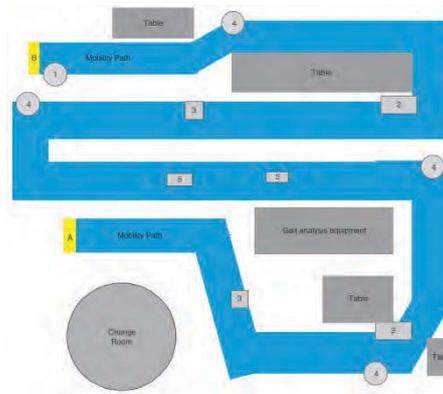


Figure 31 : Carte du trajet à parcourir dans l'expérience de Dowling. Les sujets devaient aller du point A au point B en effectuant différentes actions sur le trajet. Source : [Dowling et al. 2005]

Une dernière expérience de Jason Dowling et ses collaborateurs portait sur la fréquence de rafraîchissement ainsi que la résolution de la matrice de phosphènes. Dowling a comparé différentes résolutions et fréquences de rafraîchissement [Dowling et al. 2006]. Comme pour l'expérience précédente, il a utilisé comme mesure de performance la vitesse moyenne de marche sur dix mètres et a décompté le nombre de contacts avec des obstacles. Ses conclusions sont qu'une haute résolution est plus utile qu'un rafraîchissement de l'image élevé pour éviter les obstacles et suivre un trajet. Cependant une fréquence de rafraîchissement élevée permet aux utilisateurs de gagner en vitesse. Ainsi pour repérer des dangers éventuels, il vaut mieux avoir une bonne résolution, alors qu'il est plus efficace d'avoir une fréquence de rafraîchissement élevée lorsqu'il s'agit de se déplacer dans un environnement dégagé.

L'équipe de Gislin Dagnelie, située à Baltimore a publié en 2007 une étude sur la mobilité en vision prothétique [Dagnelie et al. 2007]. De façon analogue aux études de Dowling, la première partie de cette expérimentation portait sur la navigation dans un bâtiment inconnu des sujets. L'objectif principal était de déterminer le rendu visuel minimum requis pour parcourir un trajet simple, retenir des points de repère et éviter des obstacles au sol. Les sujets suivaient un parcours prédéterminé à l'aide d'instructions orales. Trois tailles de matrices de phosphènes ont été utilisées (4x4, 6x10 et 16x16). Ces trois tailles correspondant aux résolutions des implants développés ou en développement à l'époque par l'entreprise Second Sight. Cette première étude a conclu que deux cent cinquante-six phosphènes sont nécessaires pour naviguer efficacement (matrice 16x16).

La deuxième expérimentation reportée dans cet article portait sur la navigation dans un environnement virtuel. Les sujets se déplaçaient dans l'environnement à l'aide d'une manette de

jeu et visualisaient cet environnement dans un visiocasque. La tâche consistait à traverser une série de pièces. Les murs étaient blancs, tandis que les sols et plafonds étaient noirs. Chaque pièce comportait un numéro inscrit en noir sur un de ses murs. Les sujets devaient aller de la salle 1 à la salle 10 en un temps minimal. La matrice de phosphènes avait toujours la même résolution (10x6) avec une taille apparente de vingt-sept degrés (horizontale) par seize degrés (verticale). Les facteurs évalués lors de cette expérience étaient la présence de bruit dynamique sur la luminance des phosphènes ainsi que le taux d'électrodes défectueuses compris entre zéro et trente pour cent. Le temps de parcours des sujets a été enregistré ainsi que le nombre d'erreurs de navigation. L'ajout de bruit et le taux d'électrodes défectueuses de trente pour cent n'ont pas eu d'impact sur les performances des sujets. Les auteurs notent cependant une grande variabilité entre les sujets, ce qui pourrait expliquer l'absence d'effet des facteurs. Contrairement aux précédentes expériences, on peut considérer que celle-ci évaluait plus l'orientation que la mobilité. En effet, aucun obstacle à éviter n'était présent dans les environnements et les sujets devaient s'orienter pour trouver leur chemin et naviguer d'une salle à la suivante.

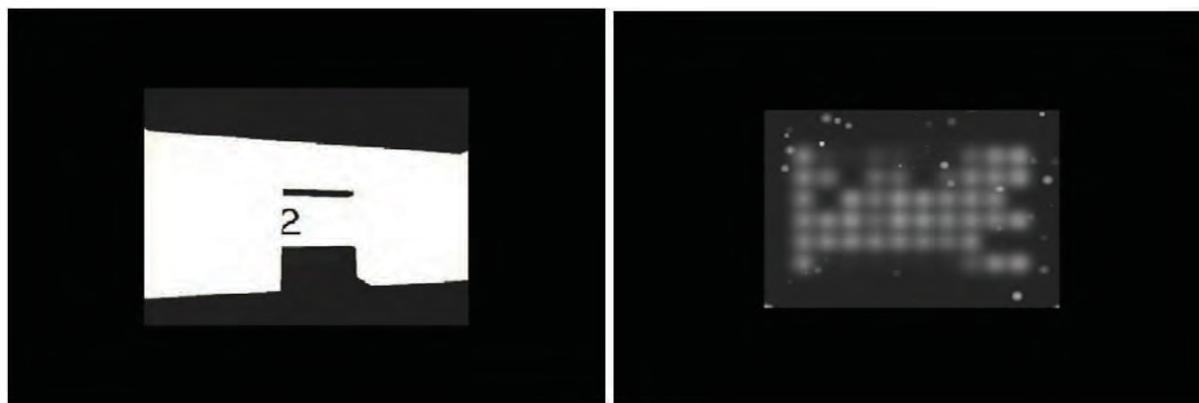


Figure 32 : À gauche : L'environnement virtuel utilisé par Dagnelie et ses collaborateurs avec le rendu en vision prothétique simulée à droite. Source : [Dagnelie et al. 2007]

Suite à cette expérience, une autre étude en environnement virtuel fut menée [Wang et al. 2008]. Dans cette nouvelle étude, les conditions différaient sur trois facteurs : les contrastes, le bruit de fond ainsi que le dysfonctionnement de 30 % des électrodes. Le système de vision prothétique fut amélioré par rapport à l'expérience précédente, avec l'ajout d'un asservissement de la position des phosphènes à la direction du regard, pour éviter une exploration visuelle de la matrice de phosphènes par les sujets. Quatre sujets devaient trouver leur chemin à travers une séquence de salles en allant de la salle 1 à la salle 10. Les conclusions de cette étude sont que les sujets ont pu s'adapter au verrouillage de la matrice sur la position de l'œil au bout d'une vingtaine d'essais. Ils ont noté également un impact du nombre d'électrodes défectueuses, le bruit de fond en revanche ne semble pas avoir eu d'impact sur les performances des sujets.

L'année suivante, une autre étude faite en collaboration avec Gislin Dagnelie portait sur la simulation d'un implant cortical plutôt qu'un implant rétinien [Srivastava et al. 2009]. La résolution de la matrice de phosphènes était plus importante : 650 phosphènes au lieu de 60. La matrice s'affichait à l'intérieur d'un cercle de 36° de rayon visuel. Trois différents taux de dysfonctionnement des électrodes furent testés : 0 %, 25 % et 50 %. Les environnements virtuels

testés étaient constitués de salles dont les murs étaient blancs, le sol et le plafond noirs. Dans chaque salle, un mur contenait l'inscription d'une lettre noire pour indiquer aux sujets leur avancement dans la tâche. Le contour des passages entre les salles et les angles entre les murs étaient affichés en gris pour faciliter l'orientation des sujets. Les résultats ont montré qu'avec un apprentissage suffisant, les sujets étaient en capacité d'accomplir la tâche même avec un taux de dysfonctionnement des électrodes important (50 %). Dans cette tâche de navigation, l'étude révèle à nouveau de grandes différences interindividuelles.

L'équipe de Rheede a également publié une étude portant sur la navigation [Rheede et al. 2010]. Trois rendus différents ont été comparés en utilisant une matrice de phosphènes de grande taille (30x30). Ces trois rendus dénommés « Plein Champ », « Région D'Intérêt » et « Fisheye » sont ceux déjà décrits ci-dessus (section II.C.5 Reconnaissance et saisie d'objets). Pour effectuer cette tâche de navigation, les sujets devaient parcourir deux trajets qui leur étaient indiqués auparavant par des instructions verbales. Cette navigation se faisait dans un environnement virtuel et des obstacles, que les sujets devaient contourner, étaient également présents. Le temps pour effectuer la tâche ainsi que la distance parcourue étaient mesurés pour évaluer les performances de navigation. Le rendu le plus efficace pour naviguer, tant en termes de temps que d'optimisation de trajet était le rendu « Plein Champ ». Les deux autres rendus ne se démarquaient pas l'un de l'autre. Nous pouvons interpréter cela en inférant que pour la navigation, la vision périphérique est importante mais il faut également que les perspectives ne soient pas déformées afin de mieux comprendre l'environnement et de s'y déplacer. Une autre raison possible est qu'avec le rendu « Fisheye » les sujets n'ont pas eu suffisamment d'apprentissage pour l'utiliser à leur avantage.

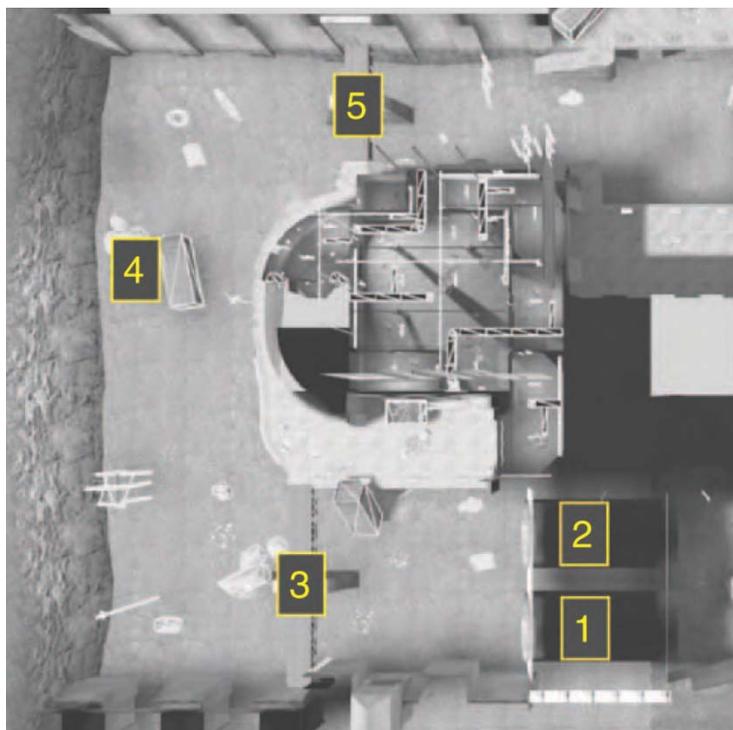


Figure 33 : Environnement virtuel utilisé par Rheede et ses collaborateurs vu du dessus. Les sujets devaient parcourir deux trajets dans cet environnement. Pour le premier ils devaient traverser le tunnel 1 (1), passer sous un pont (3) puis retourner sous le tunnel 2 (2). Pour le deuxième ils devaient traverser le tunnel 2 (2), passer sous un pont (3), tourner à droite, suivre la structure et passer à côté d'un conteneur (4) et enfin tourner de nouveau à droite pour passer sous un autre pont (5). Source : [Rheede et al. 2010]

Une autre technique de rendu de l'environnement a été proposée en traduisant la distance à un élément la scène visuelle en intensité lumineuse. Un phosphène ne représente alors plus la luminosité d'une certaine zone mais la distance moyenne de cette zone à l'utilisateur [Tatur 2011]. En 2011 ce rendu fut expérimenté pour la mobilité en vision prothétique [Lieby et al. 2011; Barnes et al. 2011]. Cette expérience a été faite en environnement réel : le traitement de l'information et la simulation des phosphènes étaient réalisés en temps réel. La matrice comportait 35x30 phosphènes et affichait 256 niveaux de gris. Les sujets devaient parcourir un couloir composé de plusieurs segments en un minimum de temps. Les rendus de distance et de luminance apparente ont été comparés. L'encodage de la distance en fonction de la luminance a permis aux sujets d'avoir de meilleures performances de mobilité qu'avec le rendu classique lorsque des obstacles devaient être évités. C'est le résultat inverse pour un trajet sans obstacle, ce qui montre à quel point il est important d'adapter le rendu utilisé à la tâche et au contexte. Un algorithme de traitement d'image a été proposé par la même équipe afin de mieux détecter les obstacles situés au sol [McCarthy et al. 2011]. Cet algorithme permet de créer un rendu de distance augmenté dans lequel les obstacles, même s'ils ont la même luminance que le fond, sont présentés avec une luminance différente pour être mieux perçus. Le calcul de la distance des objets dans l'image est combiné avec un algorithme de segmentation d'image pour arriver à ce résultat. McCarthy et Barnes proposèrent en 2012 en autre rendu pour essayer de dépasser les rendus basés sur la luminance des objets ou leur distance dans l'image [McCarthy & Barnes 2012]. La luminance des phosphènes indiquait le temps restant avant que l'objet n'entre en contact avec le sujet.

Un autre rendu augmenté a été proposé par une équipe californienne [Parikh et al. 2013]. Le rendu proposé indiquait les objets saillants présents dans le champ visuel de la caméra. À ce champ visuel représentant une carte de saillance, les auteurs ont proposé d'ajouter un indice indiquant la direction de l'objet le plus proche. La matrice de phosphènes utilisée correspondait à l'Argus II (6x10 phosphènes). Un taux de dysfonctionnement des électrodes était présent et fixé à 30 % et huit niveaux de gris ont été utilisés dans cette expérience. Les sujets effectuaient la traversée d'un couloir long d'une dizaine de mètres en évitant des obstacles. Le temps d'accomplissement de la tâche était plus long avec l'indice que sans indice. Les auteurs ont supposé que cela proviendrait d'une surcharge cognitive due à l'interprétation de l'indice.

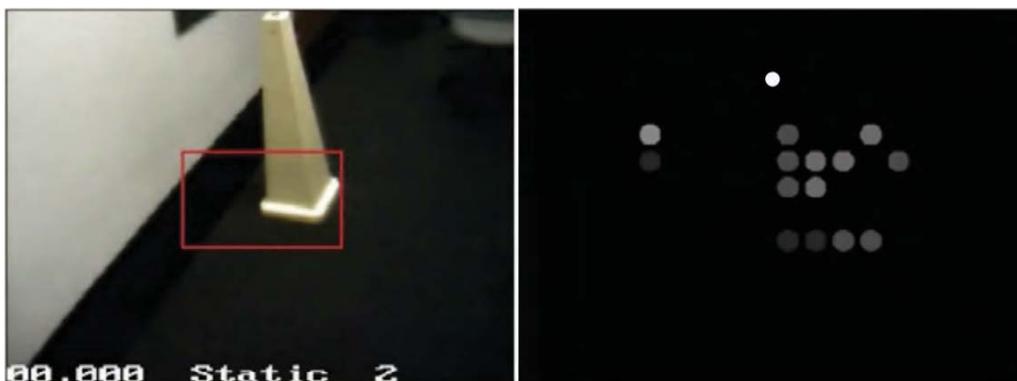


Figure 34 : Un obstacle et sa représentation en vision prothétique simulée par Parikh et collaborateurs. Les bords de la matrice de phosphène servent à indiquer la direction des obstacles immédiats. Dans cet exemple, le cône est détecté et sa direction est indiquée par un phosphène blanc situé dans la partie supérieure de la matrice. Source : [Parikh et al. 2013]

McCarthy et collègues expérimentèrent également un rendu de « distance augmentée » en 2014 [McCarthy et al. 2014]. En utilisant un nouvel algorithme discriminant finement les différentes surfaces. En plus d'un encodage de la distance, ce rendu augmenté ajoutait une mise en valeur des obstacles détectés. Huit sujets ont été invités à traverser un couloir contenant des obstacles au sol le plus rapidement possible. Les obstacles étaient assez bas (environ 15cm de hauteur) et de la même couleur que le sol. La matrice utilisée pour cette expérimentation avait une résolution très faible (5x4 phosphènes) et s'étendait sur 15° d'angle visuel (diagonale). Quatre conditions différentes ont été comparées : Distance Augmentée, Distance, Intensité et Aléatoire. Le rendu aléatoire servait ici de condition contrôle. Les résultats montrent que le rendu de distance augmentée (voir Figure 35) limite le nombre de collisions par rapport aux autres rendus.

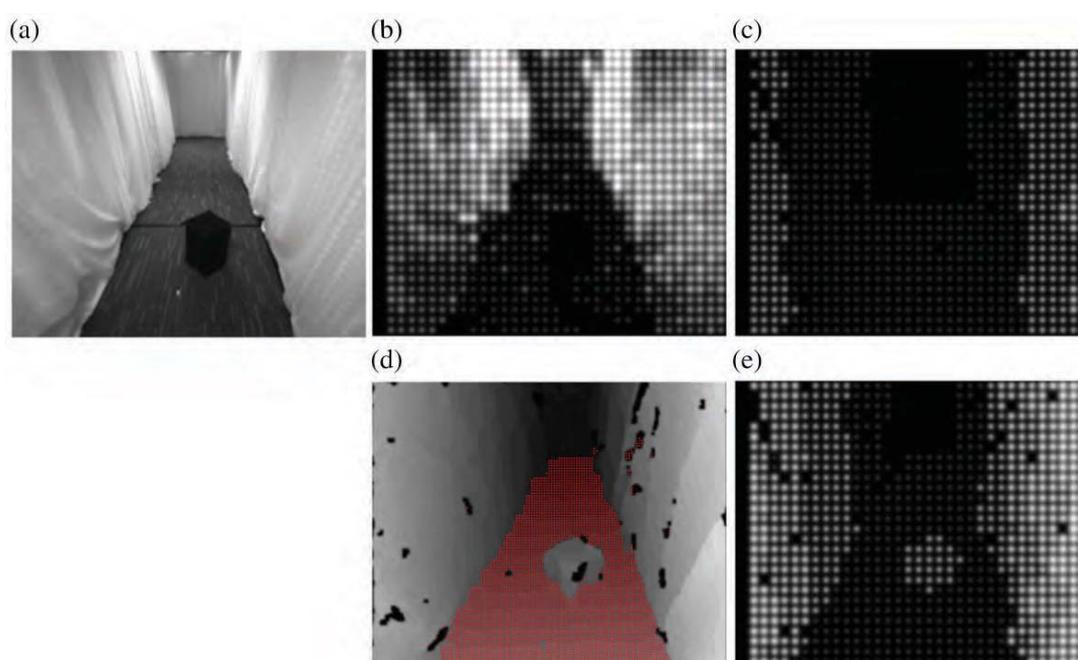


Figure 35 : Rendus développés par McCarthy. (a) Photo de départ en nuances de gris. Le faible contraste entre le sol et l'obstacle le rend difficile à voir sans traitement (b). En appliquant seulement un rendu de distance, l'obstacle n'est pas visible car trop près du sol (c). L'algorithme qu'ils ont développé leur permet d'isoler l'objet en le distinguant du sol (d) ce qui permet de le mettre en surbrillance dans le rendu de distance augmentée qu'ils ont développé (e). Source : [McCarthy et al. 2014]

Marc P. Zapf de l'équipe de Nigel Lovel a expérimenté lors de ses travaux de thèses d'autres tâches en vision prothétique simulée [Zapf et al. 2014]. La première tâche consistait à éviter des obstacles dans un couloir virtuel. La seconde consistait à suivre une ligne blanche au sol afin de passer entre des voitures fixes. Deux facteurs ont été étudiés : d'une part la résolution de la matrice de phosphène, la première étant de cent phosphènes (10x10) et la deuxième de quatre cents phosphènes (20x20). D'autre part, la technique d'interaction utilisée : les sujets se déplaçaient dans l'environnement virtuel à l'aide du couple souris/clavier, d'une manette ou de la Kinect (les sujets marchaient sur place et s'orientaient en tournant les épaules). Un effet d'apprentissage a été obtenu, quelle que soit la technique d'interaction, la vitesse moyenne des sujets augmentait au fur et à mesure des essais. Dans cette expérience, il n'y a pas de différence

notable entre les résolutions de matrices ou les techniques d'interaction. Les environnements utilisés ainsi que les rendus phosphéniques sont visibles dans la Figure 36.

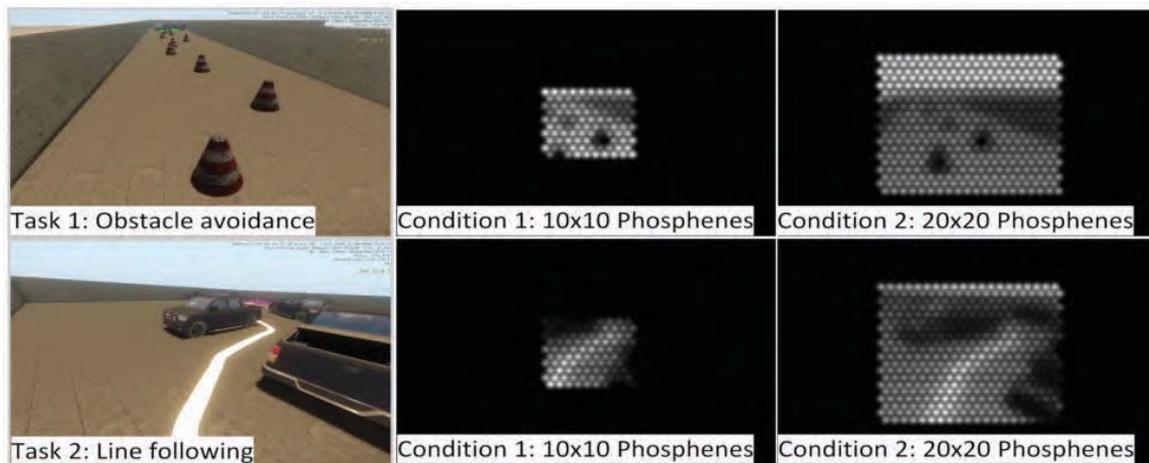


Figure 36 : Rendus utilisés par Zapf lors de ses travaux de thèse, deux tâches de mobilités ont été effectuées avec deux tailles de matrice différentes (10x10 phosphènes ou 20x20 phosphènes). Source : [Zapf et al. 2014]

Zapf a mené ensuite une expérience sur la possibilité pour des personnes ayant une vision tubulaire d'accéder à des informations périphériques à l'aide d'un implant rétinien. Cette expérimentation couplait donc une VPS à un rendu de vision tubulaire centrale [Zapf et al. 2015]. Les performances des sujets ont été évaluées sur quatre tâches de mobilité. Pour cela, les deux tâches de l'expérimentation précédente ont été utilisées, les sujets devaient également se déplacer en évitant des piétons statiques et en mouvement (voir Figure 37 pour une représentation de ces tâches). Pour chacune de ces tâches en vision tubulaire, l'ajout d'information de périphérie grâce aux phosphènes a permis d'améliorer les performances des sujets.

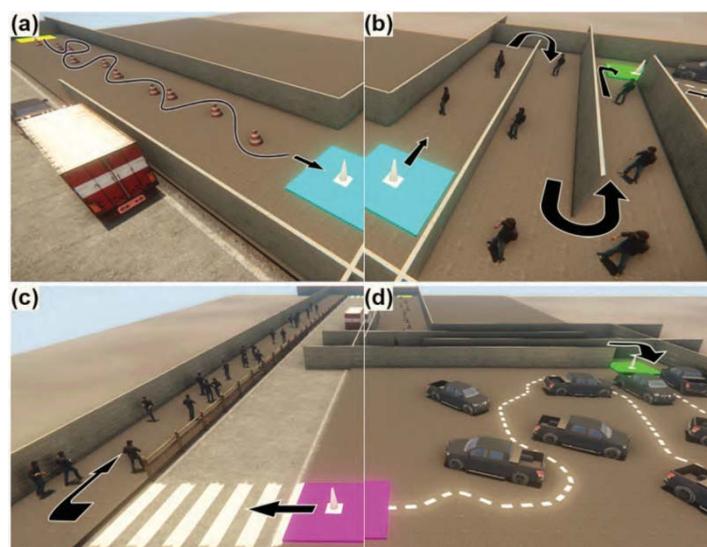


Figure 37 : Tâches effectuées par les sujets dans l'expérimentation de Zapf. Les sujets devaient d'abord traverser un couloir en évitant des plots (a) puis éviter des piétons statiques (b) avant de suivre une ligne blanche au sol (d) pour finalement longer un trottoir en évitant des piétons en mouvement (c). Source : [Zapf et al. 2015]

Le tableau suivant récapitule les différentes expérimentations effectuées avec des sujets en vision prothétique simulée et indique les tâches évaluées ainsi que les propriétés des matrices de phosphènes utilisées.

Tableau 5 : Études portant sur la mobilité et la navigation en vision prothétique simulée.

Publications	Tâches évaluées	Taille de la matrice	Forme des phosphènes	Objet de l'étude
[Cha et al. 1992b]	Mobilité	10x10, 16x16, 25x25, 32x32	Points	Première étude en SPV
[Dowling et al. 2004]	Identification de scènes liées à la mobilité	50x50	2 niveaux de gris	Comparaison de différents algorithmes de traitement d'image.
[Dowling et al. 2005]	Mobilité	24x32	8, 256 niveaux de gris.	Test d'une alerte visuelle
[Dowling et al. 2006]	Mobilité	24x32, 12x16.	8 niveaux de gris.	Influence du rafraîchissement sur la performance.
[Wang et al. 2008]	Mobilité	6x10	Ronds (profil gaussien), 256 niveaux de gris.	Navigation dans un environnement virtuel
[Rheede et al. 2010]	Acuité visuelle, Évaluation d'expression faciale, mobilité	30x30	Ronds, 8 niveaux de gris, fusion	Évaluation de différentes distorsions : Full Field, Region of Interest, Fisheye
[Barnes et al. 2011]	Mobilité, évitement d'obstacle	30x30	Non Renseigné	Comparaison rendu vs. profondeur et rendu vs. intensité
[Lieby et al. 2011]	Mobilité	30x35	Ronds (profil gaussien), 256 niveaux de gris.	Présentation du système. Rendus d'intensité et de profondeur.
[Parikh et al. 2013]	Mobilité, saisie d'objet, identification de cibles.	6x10	Ronds, 8 niveaux de gris	Rendu indiquant la présence d'obstacles. Utilisation de la localisation.
[McCarthy et al. 2014]	Mobilité, évitement d'obstacles.	30x35	Ronds (profil gaussien), 8 niveaux de gris.	Traitement du flux vidéo pour mettre en surbrillance les obstacles.
[Zapf et al. 2014]	Mobilité, Évitement d'obstacle, suivre une ligne.	10x10, 20x20	256 niveaux de gris	Comparaisons entre techniques d'interaction.
[Zapf et al. 2015]	Mobilité, évitement d'obstacles et de piétons, suivre une ligne	10x10	256 niveaux de gris.	Vision tubulaire couplée avec vision prothétique périphérique.

Comme nous l'avons vu, la quasi-totalité des études en VPS portent sur la mobilité. Dans ces expériences, les sujets doivent avancer en ligne droite en évitant des obstacles ou suivre des instructions verbales. Les expérimentations portant sur des tâches les plus proches de la navigation sont celles effectuées par Dagnelie et Wang [Wang et al. 2008 ; Dagnelie et al. 2007]. En effet, dans ces expérimentations, les sujets devaient s'orienter en permanence pour trouver leur chemin vers la salle suivante. Cependant, un seul trajet était possible pour accomplir la tâche et aucun choix ne se présentait réellement aux sujets. Jusqu'à présent, aucune étude en VPS n'a étudié l'importance de la taille de l'implant et du rendu sur la capacité à trouver son chemin dans un environnement. De même la mémorisation de l'espace dans lequel les sujets naviguent n'a jamais été abordée en VPS. Le chapitre suivant détaille les modèles de cognition spatiale humaine développés en psychologie, notamment dans le cadre de la navigation piétonne.

Chapitre IV

Modèles de cognition spatiale

A. Introduction et définition

La cognition spatiale regroupe les facultés et les processus qui permettent à un être de percevoir, traiter et manipuler des informations spatiales. L'étude de la cognition spatiale s'est développée à partir de l'apprentissage de l'espace chez l'enfant [Piaget & Inhelder 1948; Piaget 1937]. Des facultés très variées rentrent dans ce domaine. Par exemple lors d'un voyage touristique une personne peut décider d'explorer la ville, elle va intégrer ses déplacements pour se repérer dans ce nouvel environnement à grande échelle et y naviguer sans se perdre. Cette faculté à s'orienter entre dans le cadre de la cognition spatiale. Lors de son parcours, elle peut avoir besoin de sortir une carte de son sac et de la déplier afin de la lire. Cet acte nécessite de faire appel à la faculté de manipulation de petits objets qui est apprise dès l'enfance. La carte est une représentation à petite échelle de la ville, lorsque cette personne explore la carte et interprète cette représentation de la ville puis fait un lien entre cette représentation à petite échelle (la carte imprimée) et sa représentation à grande échelle (la ville elle-même). Elle a alors utilisé une représentation à petite échelle pour nourrir une représentation à grande échelle. Ainsi ces capacités recouvrent aussi bien la manipulation de petits objets, l'interprétation de représentation de petites tailles (capacités à petite échelle) et la navigation elle-même (capacités à grande échelle).

Peu de différences théoriques étaient présentes au départ entre la cognition spatiale à grande échelle et celle à petite échelle. Pour évaluer la capacité des individus à s'orienter dans l'espace, des tests papiers-crayon étaient utilisés.

Trois tests papiers-crayons sont principalement utilisés pour évaluer la cognition spatiale :

- ♦ Le test de rotation mentale de Vandenberg et Kuse (Vandenberg Mental Rotation Test) évalue la capacité à effectuer mentalement la rotation d'un objet [Shepard & Metzler 1971] [Vandenberg & Kuse 1978][Peters et al. 1995]. Ce test présente aux sujets le dessin d'un objet 3D et demande de choisir parmi quatre autres dessins celui qui représente l'objet de départ sous un autre angle.

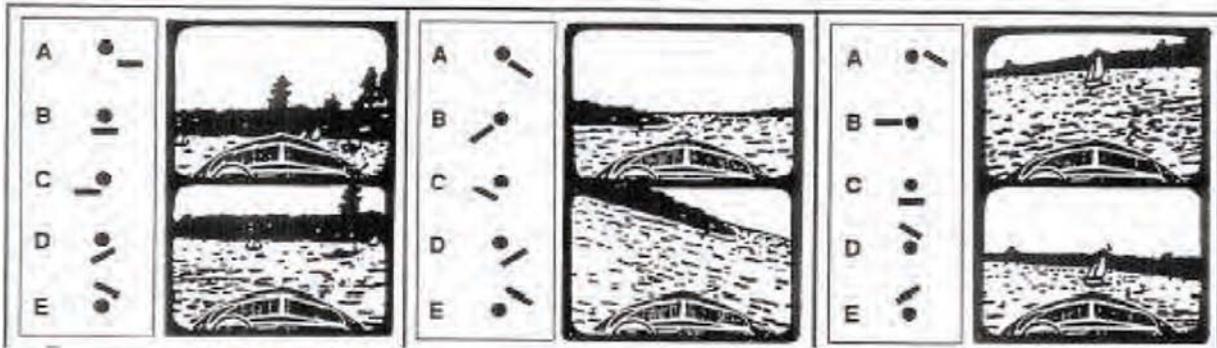


Figure 40 : Exemples d'items du test de Guilford et Zimmerman. Pour chacun des trois items, les sujets doivent choisir parmi les cinq propositions celle qui correspond au mouvement du bateau. Source : [Dejos 2012].

Les expérimentations ont semblé confirmer pendant un certain temps que la cognition spatiale à petite échelle et celle à grande échelle était un seul et même ensemble de processus. Des résultats obtenus sur des tests à petites échelles étaient alors généralisés à des capacités de cognition spatiale à grande échelle. Cette tendance a évolué depuis, les théories et expérimentations se focalisant de plus en plus spécifiquement sur la cognition spatiale à grande échelle [Kuipers 1977; Allen 2004]. En 2006, Mary Hegarty et ses collaborateurs publient un article présentant les différentes relations possibles entre les capacités spatiales à petite et grande échelles [Hegarty et al. 2006]. Différents modèles sont évoqués, le premier prévoit que ces deux groupes sont les mêmes (*unitary model*), le deuxième qu'il y a une dissociation totale entre les deux groupes (*total dissociation model*), le troisième traite d'une dissociation partielle (*partial dissociation model*), enfin le dernier modèle, appelé modèle de médiation (*mediation model*), implique que d'autres processus externes à ceux des groupes serviraient d'intermédiaires. Les résultats de cette expérience favorisent le modèle de dissociation partielle : les capacités à petite et grande échelle partageraient un ensemble de processus communs ; d'autres processus leur étant propres par ailleurs.

Comme indiqué en introduction de ce chapitre, différents espaces sont concernés par la cognition spatiale. Il est important d'indiquer précisément de quels espaces il est question lorsque l'on parle de cognition spatiale. Pour être plus précis, nous allons utiliser la classification des espaces mise au point par Montello [Montello 1993; Montello 1998]. Daniel R Montello indique que l'on peut catégoriser les espaces dans quatre groupes différents. Le premier contenant les espaces de petite échelle et les trois derniers de grande échelle.

- ◆ *Figural space* : Ce niveau contient les espaces pouvant être appréhendés sans mouvement du corps. Cet espace contient les images et les petits objets. Certains points de repère distants de taille limitée peuvent être également contenus dans cet espace. Ce niveau peut être subdivisé en deux niveaux supplémentaires que sont *pictorial space* et *object space*. Le *pictorial space* est l'espace des images et des objets en 2D, l'*object space* concerne quant à lui les objets en 3 dimensions.
- ◆ *Vista space* : La taille de ces espaces est analogue à la taille du corps ou plus grande. L'idée étant qu'un *vista space* peut être appréhendé dans son entièreté depuis un point de vue,

sans déplacement nécessaire. Les exemples de *vista spaces* donnés par Montello sont les salles, les parcs de taille réduite, les petites vallées et les horizons.

- ◆ *Environmental spaces* : il s'agit ici des espaces de taille plus importante. Contrairement aux *vista spaces*, il est nécessaire de se déplacer dans ces espaces pour pouvoir les percevoir entièrement et donc en construire une représentation complète. Montello donne comme exemples d'*environmental spaces*, les bâtiments, les quartiers ainsi que les villes. Même si un temps important peut être nécessaire pour parcourir un *environmental space*, il est possible de les parcourir et de les appréhender entièrement grâce à l'expérience directe.
- ◆ *Geographical spaces*. Ces espaces ne peuvent pas être appréhendés uniquement par l'expérience directe et nécessitent l'utilisation de représentations de plus petites tailles comme des cartes pour pouvoir être compris et appris. Montello rappelle bien que les cartes, tout en représentant des *environmental spaces* ou *geographical spaces* sont elles-mêmes des *pictorial spaces* et donc des *figural spaces*. Les états, les continents et le système solaire sont des exemples de *geographical spaces*. Au contraire la surface de l'eau vue depuis un avion en vol est un *vista space* car appréhendée directement sans déplacement nécessaire.

Dans le cadre de cette thèse, nous utilisons la définition plus restrictive donnée par Thorndyke de la **cognition spatiale** : « *the acquisition and use of knowledge about large-scale space* » ou en français : « l'acquisition et l'utilisation de connaissances portant sur les grands espaces » [Thorndyke & Hayes-Roth 1982]. Les grands espaces en question sont donc les *vista spaces*, les *environmental spaces* et les *geographical spaces*.

De même nous utilisons comme définition de la **navigation** le fait de se déplacer et de s'orienter dans un environnement. Dans le cadre de ces travaux, la navigation s'effectue surtout à pied mais la définition n'exclut pas d'autres moyens de locomotion (voitures, bus etc.). Ce concept concerne donc à la fois la mobilité (locomotion à petite échelle, évitement d'obstacles) et l'orientation (ne pas se perdre, anticiper son trajet et le réaliser à grande échelle). De plus, la navigation permet d'acquérir des connaissances sur un environnement pour s'y déplacer ainsi que de mobiliser ces connaissances afin de parcourir un trajet. Cette définition de la navigation inclut donc la notion de *wayfinding* [Lynch 1960] qui consiste à trouver son chemin pour rejoindre un lieu précis [Golledge 1992]. Les connaissances acquises grâce à la navigation ont été étudiées pendant toute la deuxième partie du xx^e siècle et différents modèles ont été élaborés pendant cette période.

B. Représentations acquises lors de la navigation

B.1. L'intuition de Tolman

En 1948, Edward T. Tolman publie une revue d'articles portant sur la navigation chez le rat [Tolman 1948]. Il fait le point sur deux écoles de pensées en psychologie animale. La première considère que les rats et leur déplacement dans les labyrinthes fonctionnent uniquement sous un schéma stimulus-réponse. Selon cette école, les bonnes performances obtenues après un

apprentissage seraient dues à un renforcement positif des connexions permettant aux rats d'automatiser des choix. La deuxième école de pensée à laquelle Tolman dit adhérer soutient le fait que les rats disposeraient d'une représentation analogue à une carte de l'environnement appelée dans l'article (et pour la première fois) « carte cognitive ».

L'objectif de cet article est d'appuyer l'idée de la carte cognitive chez le rat. Plus précisément, Tolman s'appuie sur des expériences passées pour formuler deux hypothèses :

- ◆ Les rats disposent d'une carte cognitive
- ◆ Cette carte cognitive peut être pauvre et rigide ou large et efficace.

La différence qui est faite entre des cartes pauvres et des cartes larges tient essentiellement dans leur efficacité. Après ces analyses chez le rat, Tolman indique que d'autres études l'amènent à penser que les cartes cognitives ne sont pas spécifiques aux rats mais que les hommes auraient un fonctionnement similaire. Il indique également quatre facteurs qui généreraient la création de cartes cognitives efficaces :

- ◆ Lésions cérébrales
- ◆ Indices environnementaux présentés de manière inadéquate
- ◆ Répétition excessive du trajet d'entraînement original
- ◆ Facteurs de motivation ou de frustration trop forts

Cet article présente pour la première fois la notion de carte cognitive qui s'apparente à une représentation mentale de l'environnement acquise lors de la navigation.

B.2. Le cadre dominant : modèle L-R-S

Alexander Siegel et Sheldon White proposèrent en 1975 un modèle de cognition spatiale devenu un modèle de référence aujourd'hui [Siegel & White 1975]. Ce modèle intitulé modèle « L-R-S » définit 3 niveaux de connaissances spatiales.

Landmark : Le premier niveau est constitué de repères appelés *Landmarks*. La définition que donne Lynch [Lynch 1960] est la suivante : tout élément qui peut potentiellement servir de point de repère. Dans la suite du manuscrit, les *landmarks* seront appelés **points de repère**. Les points de repère servent à reconnaître un lieu, à s'orienter lors d'un trajet ou même à s'orienter plus globalement à l'intérieur d'une ville par exemple. Même si la définition de Lynch peut regrouper un grand nombre d'éléments, un objet présent dans un environnement doit vérifier différentes propriétés pour être qualifié de point de repère. Tout d'abord, un point de repère se doit d'être fixe et persistant, un point de repère découvert en un lieu précis à un temps précis doit y persister pour pouvoir le retrouver plus tard. Ensuite il est important qu'un point de repère soit perceptivement saillant, il doit se détacher de l'environnement pour pouvoir être mémorisé et reconnu. Enfin un point de repère se doit d'être informatif, il doit fournir à la personne qui l'utilise des informations pertinentes pour sa navigation [Siegel & White 1975]. Lors de la navigation, les points de repère peuvent donc être utilisés pour identifier le début et la fin d'un trajet mais également comme point de confirmation ou de décision. Les points de repère peuvent être de forme et de nature très diverses : intersections de rues, grands bâti-

ments, arrêts de bus etc. Certains points de repère peuvent être visibles seulement depuis un lieu qu'ils permettent d'identifier alors que d'autres peuvent être visibles depuis une longue distance et servir de repère en de multiples lieux comme la tour Eiffel. Deux sous-catégories de points de repère peuvent alors être définies : les **points de repère locaux** permettant d'identifier seulement un lieu et visible uniquement depuis ce lieu et les **points de repère distants** pouvant servir de points de repères entre différents lieux [Presson & Montello 1988; Steck & Mallot 2000]. À noter que si uniquement ce niveau de connaissances est disponible alors les points de repère sont reconnaissables mais les trajets pour aller de l'un à l'autre ou leur agencement dans l'espace sont complètement inconnus.

Route : Les éléments du deuxième niveau de connaissance spatiale sont appelés *Routes*. Ces représentations contiennent les différentes informations pour aller d'un point de repère à un autre. Il s'agit donc d'une représentation séquentielle détaillant les étapes nécessaires pour se déplacer entre deux points de repère. Pour Siegel et White, alors que les points de repère sont essentiellement visuels, les *routes* sont plutôt de nature sensorimotrice. Les niveaux de connaissance *Landmark* et *Route* sont suffisants pour naviguer en ville. En effet, en utilisant uniquement ces deux niveaux de représentations, il est possible de mémoriser des trajets en retenant la séquence menant d'un point à un autre. En revanche, ces connaissances ne permettent pas d'évaluer la distance entre deux lieux distincts. Les liens entre points de repère se font par une séquence mémorisée par rapport au corps de la personne qui parcourt le trajet. Il s'agit donc d'une représentation égocentrée, les séquences d'action à mener sont toujours présentes entre deux points de repère éloignés.

Survey : Le dernier niveau introduit par les auteurs est une réelle nouveauté par rapport aux modèles préexistants provenant de la géographie et de l'étude des paysages urbains [Lynch 1960]. Les représentations de ce niveau sont appelées des *configurations*. Ce niveau permet d'intégrer les deux premiers dans une représentation unifiée. Les liens entre points de repère ne sont alors plus nécessairement des *routes* et peuvent se faire directement. Avec ce niveau de connaissance, il est possible depuis un lieu de pointer vers un point de repère lointain non visible et d'évaluer la distance entre deux points de repère même si ces derniers ne sont pas visibles simultanément. La représentation de l'environnement ne repose donc plus sur le corps (ou véhicule) de la personne qui navigue mais sur les éléments mêmes qui peuplent cette représentation. Nous parlons alors de vision **allocentrée** : il est possible de prendre de la hauteur et d'imaginer l'environnement vu de haut. Pour l'explicitier autrement, le repère principal n'est plus le navigateur en mouvement mais les points cardinaux ou un point de repère suffisamment marqué pour l'utiliser en permanence dans sa navigation même s'il n'est pas visible. Nous pouvons citer par exemple la place du Capitole à Toulouse. Pour que ce niveau de représentation soit fiable, il est nécessaire d'avoir une idée des distances entre points de repère et des différents angles que peuvent prendre leurs configurations. Ainsi une métrique est présente dans ce niveau de représentation, ce qui n'est pas nécessairement le cas pour les niveaux des *Route* et *Landmark*.

Pour résumer, les trois niveaux définis par ce modèle ont des fonctions bien spécifiques. Le premier sert à poser des ancres appelées points de repère qui jettent les bases de la représentation. Le deuxième niveau crée des liens séquentiels entre ces points de repère appelés routes. Enfin le dernier niveau intègre dans une représentation de plus grande dimension les deux premiers niveaux. Ce niveau « Survey » permet de faire la jointure entre par exemple notre expérience de navigation dans une ville que nous découvrons et l'emplacement de cette ville à plus grande échelle dans le territoire national. Par rapport à la catégorisation des espaces de Montello, les éléments du premier niveau sont inclus dans des *vista spaces*, alors que ceux du deuxième sont dans des *environmental spaces*. Les éléments du troisième niveau peuvent appartenir aux *environmental spaces* ou aux *geographical spaces*.

Une fois publié, ce modèle est devenu tellement influent par la suite que Montello [Montello 1998] l'a qualifié de cadre dominant (*dominant framework*). En effet ce cadre a été très utilisé en psychologie mais également en géographie [Golledge 1992] ou en robotique [Choset & Nagatani 2001].

B.3. Dépasser le cadre dominant : apports de Montello

Plus récemment, des évolutions ont permis de sortir peu à peu de ce cadre. Par exemple Ishikawa et Montello [Ishikawa & Montello 2006] proposent que des informations de distance soient présentes dès le début de la mémorisation en lien avec la navigation.

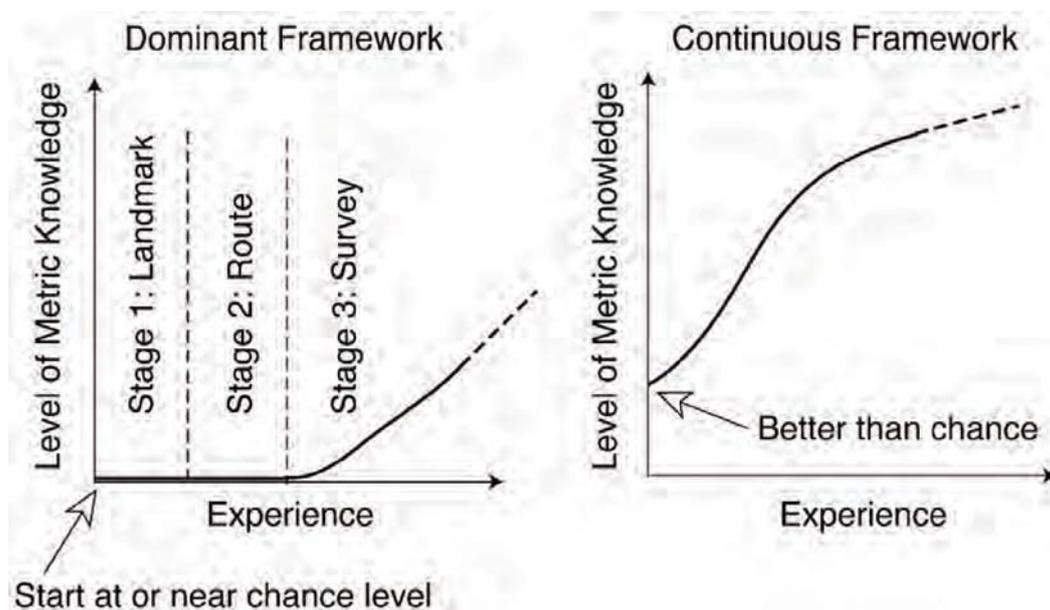


Figure 41 : À gauche le modèle de Siegel et White considère que des informations de distance ne sont mémorisées qu'à partir du niveau de représentation Survey. À droite : Ishikawa et Montello proposent que l'information métrique soit mémorisée dès le début du processus et que les trois niveaux de représentations ne soient plus successifs dans le temps. Source : [Ishikawa & Montello 2006]

Montello s'est aussi intéressé à la navigation et à l'exploration des *environmental spaces*. Ainsi dans un article de 1998, il propose cinq principes pour construire un modèle de l'exploration d'*environmental spaces* par des adultes [Montello 1998]. Ces principes ne s'appliquent donc pas aux enfants ou à la navigation à l'aide de carte et se déclinent comme suit :

- ◆ L'information métrique est présente dès le début de la construction d'une représentation mentale d'un lieu. Pour Siegel et White l'information métrique est présente uniquement dans le dernier niveau de représentation *Survey*. Montello fait ici l'hypothèse que dès le début de l'exploration d'un lieu inconnu, des informations métriques sont stockées et utilisées. Cependant ces données métriques ne sont pas acquises automatiquement.
- ◆ Tant que la familiarité d'un lieu ainsi que le temps d'exposition à un lieu augmentent, il y a une augmentation continue de la quantité et la précision de la connaissance spatiale. De même cette connaissance est de plus en plus complète.
- ◆ L'intégration de connaissances obtenues sur des lieux appris séparément pour en obtenir une connaissance plus intégrée et organisée de manière hiérarchique est un processus sophistiqué et complexe. Ce processus amène le seul changement qualitatif lors de l'apprentissage d'un *environmental space*.
- ◆ Des personnes avec un même niveau d'exposition à un lieu peuvent obtenir des connaissances différentes. De fortes différences interindividuelles sont à prévoir et dues au degré d'intégration spatiale que les personnes arrivent à effectuer.
- ◆ Il existe une connaissance spatiale non-métrique en plus de la connaissance métrique. Cette connaissance vient en plus de la connaissance métrique et n'en est pas forcément un précurseur ou une partie intrinsèque.

Ces principes permettent de poser les bases pour un nouveau modèle qui s'écarte du cadre défini par Siegel et White en intégrant plus finement l'acquisition de connaissances métriques sur un environnement et en distinguant plus d'espaces dans lesquels une navigation peut s'effectuer. En s'appuyant fortement sur les travaux de Daniel Montello, Tobias Meilinger a proposé un modèle de cognition spatiale faisant la distinction entre les différents espaces que Montello a explicités et en respectant les cinq principes vus précédemment.

B.4. Network of Reference Frame theory

Le modèle de Meilinger décrit la structure des connaissances utilisées et sollicitées lors de la navigation dans un *environmental space*. Pour construire ce modèle, Meilinger introduit une nouvelle notion de cadre de référence [Meilinger 2008a]. À chaque *vista space* est attribué un cadre de référence local à partir duquel sont mémorisés les éléments du *vista space*, ce référentiel se dénomme *reference frame* et possède une origine et une orientation. Ainsi les points de repère présents dans un *vista space* ont leurs coordonnées exprimées dans ce *reference frame*. Ces *reference frames* constituent les nœuds d'un graphe dont les liens sont les transitions pour aller d'un *vista space* à un autre et s'appellent des *perspective shifts*. Les *perspective shifts* ont un rôle analogue aux routes du modèle de Siegel et White. Ce graphe est le modèle proposé par la théorie de Meilinger qu'il intitule « *Network of Reference Frame Theory* » et qu'il a développé durant ses travaux de thèses en psychologie [Meilinger 2008b].

Pour comprendre ce modèle, il faut le comparer à la façon dont le modèle LRS de Siegel et White intègre les informations métriques. Dans le modèle LRS, le premier niveau (*Landmark*)

et le deuxième niveau (*Route*) ne contiennent aucune information métrique. Les points de repère sont uniquement reliés entre eux par des routes analogues à des séquences d'instructions. Ces deux niveaux sont donc analogues à un graphe topologique. L'information métrique est présente uniquement dans le troisième niveau (*Survey*) ce qui permet d'unifier la représentation en une carte cognitive cohérente. Dans une représentation de niveau *Survey* les différents points de repère sont situés précisément les uns par rapport aux autres dans une seule carte cognitive unifiée. Le passage du niveau *Route* au niveau *Survey* amène donc un saut qualitatif important. Ces deux types de représentations sont visibles dans la Figure 42.

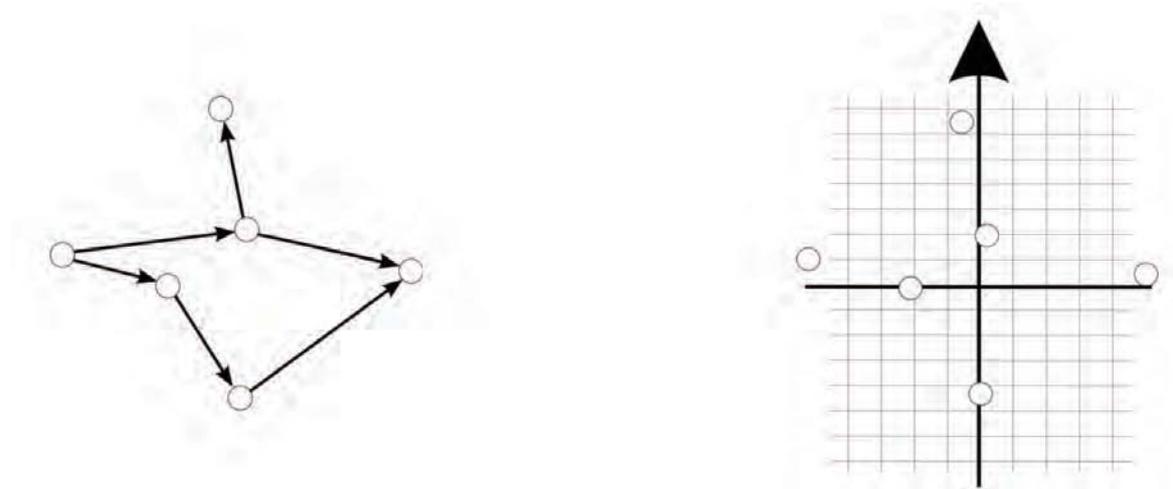


Figure 42 : À gauche : Représentation en graphe des connaissances spatiales, ce qui correspond au niveau route de Siegel et White, les points de repère sont reliés entre eux par des routes. À droite : Représentation des connaissances spatiales sous forme d'une carte cognitive intégrée (*Survey*). Source : [Meilinger 2008a]

Dans le modèle de Meilinger, il n'y a pas de référentiel global à proprement parler, chaque élément présent en mémoire est situé par rapport à un *reference frame* local qui est lui-même rattaché à un *vista space*. Cependant l'information métrique est présente dès l'apparition des *vista spaces* ce qui rejoint les principes posés par Montello. D'après Meilinger, chaque *vista space* et les éléments (points de repère et structure) qui le composent sont situés par rapport au *reference frame* local du *vista space* en question. Ainsi dès qu'un lieu est découvert, certaines informations métriques sont mémorisées. Par exemple lorsqu'une personne découvre la place du Capitole à Toulouse, elle mémorise déjà une approximation de la taille de la place et des bâtiments environnants comme l'hôtel de ville.

Les *perspective shifts* qui connectent les *reference frames* entre eux possèdent une composante de rotation et une composante de translation permettant de passer d'un *reference frame* à un autre. Lors de la découverte d'un lieu inconnu, ces composantes sont très imprécises. Les *perspective shifts* possèdent également une force (*strength*) qui peut découler de la familiarité du trajet. Pour un *perspective shift* particulier, plus le trajet en question sera effectué, plus le *perspective shift* sera fort et précis. Il est à noter que les *perspective shifts* vont d'un *reference frame* à un autre dans un sens donné et sont asymétriques. Ce réseau de *reference frames* est visible dans la Figure 43.

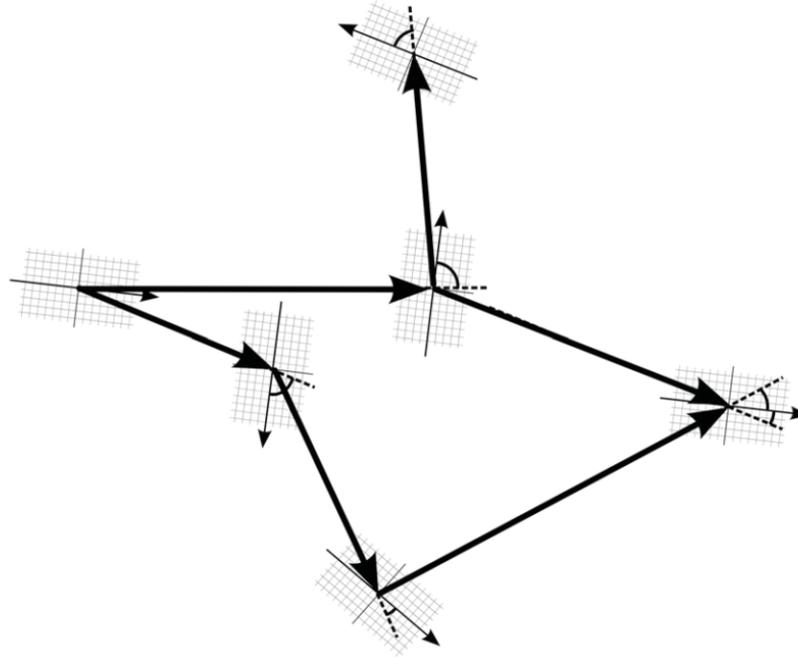


Figure 43 : Modèle de Meilinger, les *vista spaces* contiennent chacun un référentiel et sont reliés entre eux par des *perspective shifts*. Source : [Meilinger 2008a]

L'orientation des *reference frames* décidée par leur axe principal ne se détermine pas par le hasard. Lorsque quelqu'un découvre pour la première fois un *vista space* qu'il ne connaissait pas, le *reference frame* se crée au même moment. Le *reference frame* peut se fixer une orientation dépendante de la structure du *vista space*, par exemple une place rectangulaire nous amènera à favoriser un de ces deux axes pour s'orienter. De même, si le *vista space* possède un point de repère fort comme une fontaine ou un hôtel de ville, il est possible que l'axe de référence passe par ce point de repère. Enfin, cette orientation peut également dépendre du point d'entrée dans le *vista space*. Ainsi lors de la découverte d'une place sans structure forte ni élément particulier pour l'orienter, il est possible d'avoir comme axe de référence le point de vue adopté lors de l'entrée dans le *vista space*. Ce processus de création et d'orientation de *reference frame* est appelé par Meilinger *Encoding*. Trois autres processus sont proposés dans son modèle. Le premier est le fait de reconnaître un *vista space* grâce à un point de repère ou à sa structure. Cette reconnaissance permet d'identifier le lieu où l'on est ainsi que de retrouver l'orientation du *reference frame*. Ce processus est appelé par Meilinger *Self-localization by recognition*. Lors de la navigation elle-même, pour anticiper un trajet le modèle de Meilinger prévoit une propagation de l'activation qu'il appelle *Route navigation by activation spread*. Cette propagation se produit lorsqu'une recherche d'itinéraire est effectuée. Lors de la navigation, les *reference frames* concernés sont pré-activés ce qui permet de les reconnaître plus facilement. Les *reference frames* que l'on traverse se désactivent une fois qu'ils sont passés. Une différence importante du modèle de Meilinger par rapport au modèle dominant de Siegel et White est donc l'absence d'une manipulation mentale d'un environnement dans son ensemble.

Il est possible d'effectuer une navigation par l'imagination en se projetant vers un *reference frame* adjacent puis un autre pour finalement arriver à évaluer une direction vers un

autre *reference frame* non visible depuis la position de départ. Ce processus utilisé pour pointer vers un lieu non-visible ou trouver des raccourcis est appelé *Survey navigation by imagination* par Meilinger.

Les *perspective shifts* peuvent se créer et évoluer de différentes façons. Tout d'abord si des éléments d'un *vista space* sont visibles depuis un *vista space* voisin, dans ce cas le *perspective shift* peut se déduire de la scène visuelle elle-même. Ensuite le déplacement lui-même effectué entre deux *vista spaces*, favorise une mise à jour de cette transition. Enfin une aide peut être apportée par un point de repère lointain et visible depuis les deux *vista spaces* en question.

Avec le modèle de Meilinger, plus le réseau de *reference frames* contient de nœuds et d'arêtes, plus la difficulté est grande pour se déplacer à l'intérieur de ce réseau. Ceci implique que les nombres de *vista spaces* et de *perspective shifts* sont des indices de la difficulté à naviguer dans un environnement. Ainsi un environnement contenant des espaces de petite taille avec de nombreuses ramifications va amener l'utilisateur à créer de nombreux *vista spaces* et *perspective shifts* en s'y déplaçant, ce qui aura comme conséquence de compliquer sa navigation. À l'inverse, si l'utilisateur navigue dans un environnement contenant de grands espaces avec un faible nombre de liens entre eux, les *vista spaces* seront alors peu nombreux et de grande taille, ce qui permettra de naviguer plus facilement.

B.5. Cognition spatiale des non-voyants

Pour les voyants, l'orientation est rarement critique en environnement urbain. Si une personne voyante se perd lors de l'un de ses trajets, il est toujours possible de demander une indication et de la suivre facilement. La généralisation des smartphones a également amené tout un chacun à avoir en permanence à portée de main une carte des environs disponible avec un système de guidage intégré. Les personnes voyantes sont donc très rarement en situation de danger ou de stress lorsqu'elles sont perdues et parviennent à retrouver leur chemin en autonomie. Pour les non-voyants, le fait de se perdre est beaucoup plus anxiogène car il est alors quasiment impossible de retrouver son chemin sans une aide externe importante. Se représenter correctement son espace et son trajet au préalable est donc crucial pour se déplacer et naviguer en autonomie. Pour cela, les personnes non-voyantes apprennent un trajet jusqu'à le connaître par cœur et n'effectuent en général pas seules des trajets inconnus. La cognition spatiale chez les non-voyants a été déjà très étudiée dès le milieu du $xx^{\text{ème}}$ siècle et les différents modèles peuvent se ranger selon une classification des théories qui les sous-tendent.

La cognition spatiale chez les non-voyants se décline en 3 théories principales identifiées par [Fletcher 1980] :

- ♦ La **théorie du déficit** (*Deficiency Theory*) proposait que le développement de concepts liés à l'espace était impossible chez les aveugles congénitaux. D'après cette théorie, un minimum d'expérience visuelle était requis pour développer une compréhension même minimale de l'espace. Depuis cette théorie a été complètement abandonnée.

- ♦ La **théorie de l'inefficacité** (*Inefficiency Theory*) considère que les aveugles peuvent développer des représentations mentales de l'espace, mais que ces représentations sont fonctionnellement inférieures à celles développées par les voyants ou les non-voyants tardifs. Selon cette théorie, les personnes aveugles ne peuvent pas se faire une idée globale de la forme d'objets trop complexes.
- ♦ D'après la dernière théorie intitulée **théorie de la différence** (*Difference Theory*), les non-voyants pourraient construire des représentations spatiales fonctionnellement équivalentes à celle des voyants. Cependant ces représentations seraient plus longues à construire et à manipuler car l'information non visuelle à traiter et à intégrer est de nature différente.

Actuellement, la dernière théorie est la plus suivie, des études récentes ont montré que la différence entre non-voyants et voyants seraient plus quantitative que qualitative [Iachini & Ruggiero 2010]. Dans une étude portant sur le volume de l'hippocampe (structure cérébrale impliquée dans la navigation) et la performance de navigation, le groupe de non-voyants a eu de meilleures performances que le groupe de voyants avec les yeux bandés dans l'apprentissage et la restitution d'une route par la navigation [Fortin et al. 2008]. Dans le même sens, dans une même tâche d'imagerie mentale et de manipulation d'information 2D non-visuelle les voyants et non-voyants peuvent utiliser des stratégies différentes tout en obtenant des performances en imagerie mentale qui soient similaires [Vanlierde & Wanet-Defalque 2004]. Pour se représenter certains objets de la vie quotidienne, les non-voyants congénitaux semblent en difficulté lorsqu'il s'agit de se les imaginer à certaines distances dans le sens où pour eux les objets ont toujours la même taille quelle que soit la distance depuis laquelle ils sont perçus [Vanlierde & Wanet-Defalque 2005].

Un facteur très étudié dans l'étude de la cognition spatiale à grande échelle chez les non-voyants est l'existence d'une expérience visuelle préalable. Dans le cadre de la navigation, il est possible qu'une certaine expérience visuelle soit indispensable pour se représenter le monde d'un point de vue allocentré et non plus égocentré [Thinus-Blanc & Gaunet 1997]. Ainsi, la plupart des études portant sur ce sujet comparent des sujets non-voyants congénitaux, non-voyants tardifs et des voyants en tant que groupe contrôle. Les aveugles congénitaux semblent avoir plus de difficultés pour indiquer la distance et la direction de points de repère que les non-voyants tardifs ou des sujets voyants avec les yeux bandés [Veraart & Wanet-Defalque 1987]. En revanche, une série d'expérimentations menées récemment montre que des sujets aveugles (de naissance ou non) peuvent faire des erreurs moindres que des sujets voyants les yeux bandés sur des tâches nécessitant une représentation de type *Survey* [Tinti et al. 2006]. La méthode de recrutement des sujets peut biaiser les résultats (dans un sens ou dans l'autre), si par exemple les sujets proviennent tous d'un centre dédié aux non-voyants où ils ont pu apprendre à naviguer. Les résultats des études ne précisant pas l'expérience des sujets non-voyants sont donc à prendre avec précautions [Loomis et al. 1993].

Pour un voyant, les informations perçues lors de la navigation sont très précises, nombreuses et sont traitées simultanément. À l'inverse, un non-voyant qui navigue perçoit des

informations parcellaires traitées séquentiellement. De plus, il utilise son système vestibulaire ainsi que sa proprioception à la place de la vision [Berthoz & Viaud-Delmon 1999]. Cette différence dans le flux de l'information semble avoir un impact conséquent sur la navigation et sur les performances obtenues si l'on limite le nombre d'informations disponibles simultanément chez des sujets voyants [Ruotolo et al. 2012]. De plus, des différences peuvent apparaître sur les stratégies de mémorisation utilisées : les non-voyants étant moins performants que des voyants s'ils optent pour une stratégie d'imagerie mentale de l'espace alors qu'ils obtiennent des performances similaires en utilisant une stratégie verbale [Cornoldi et al. 2009].

Les expérimentations menées dans le cadre de ce doctorat sont décrites dans la partie qui suit. Les connaissances sur les neuroprothèses visuelles, leur simulation ainsi que sur la tâche précise de navigation ont été mobilisées pour mettre au point les protocoles expérimentaux et interpréter les résultats. En particulier la théorie de Meilinger de *Network of Reference Frames* a été utilisée pour produire certains rendus prothétiques et les analyser. Ces contributions sont constituées de quatre chapitres : le premier évalue l'application d'un rendu déjà utilisé pour des tâches de localisation d'objets à la navigation, le deuxième décrit une expérimentation se concentrant sur un rendu permettant de mettre en valeur la structure de l'environnement, le troisième compare différentes stratégies pour rendre au mieux la structure ainsi obtenue et compare certains de ces rendus à leurs analogues en vision non prothétique. Le quatrième chapitre de contribution traite d'un questionnaire qui a été rempli par des sujets voyants et non-voyants afin de comprendre quelles sont les représentations qu'ils ont d'un trajet précis dans le centre de Toulouse. Le dernier chapitre reprend les différents résultats des expérimentations pour les discuter et en extraire quelques lignes directrices sur les rendus prothétiques destinés à la navigation.

PARTIE 2. CONTRIBUTIONS

Un grand nombre des expériences en vision prothétique simulée (VPS) et portant sur la navigation se sont concentrées sur la mobilité; les sujets devant parcourir un trajet rectiligne dans un couloir tout en évitant des obstacles [McCarthy et al. 2014]. L'orientation a été abordée avec des trajets plus complexes mais une contrainte très forte de ces expérimentations était le fait que le parcours était imposé aux sujets [Rheede et al. 2010]. La navigation libre dans un environnement autorisant différents trajets pour effectuer une même tâche n'a ainsi jamais été évaluée en VPS. De plus, pour évaluer les performances de navigation des sujets, le nombre de collisions avec des obstacles ou le temps de parcours sont régulièrement utilisés [Wang et al. 2008], en revanche les représentations mentales que se construisent les sujets durant la navigation [Coluccia et al. 2007] ne sont jamais évaluées. Ainsi même lorsque la tâche expérimentale se rapprochait d'une tâche de navigation avec orientation (*wayfinding*), les évaluations menées ne permettaient pas de savoir ce que les sujets avaient mémorisé de leur environnement ou du trajet effectué.

La méthode la plus immédiate pour produire un rendu prothétique consiste à réduire le nombre de pixels de l'image à rendre pour l'amener au nombre de phosphènes voulus. Le rendu obtenu par cette méthode qui n'utilise pas d'algorithmes très poussés de traitement d'image est souvent dénommé « *scoreboard* » et a été le premier rendu utilisé dans une tâche de navigation [Cha et al. 1992a]. Pour la suite de ce manuscrit nous l'appellerons le rendu « brut ». Pour que ce rendu soit efficace dans une tâche de navigation, il a été évalué que le nombre de phosphènes nécessaires devait être égal à au moins six cent [Cha et al. 1992b]. Cependant les implants actuellement commercialisés contiennent moins de cent électrodes indépendantes (ARGUS II de la compagnie Second Sight Inc.). Les implants en développement contiendront vraisemblablement moins de trois cents électrodes (ARGUS III de Second Sight, ou IRIS II de Pixium Vision), ce qui reste insuffisant pour naviguer avec le rendu brut. L'évaluation de rendus prothétiques pertinents pour aider à la navigation des personnes implantées avec un nombre d'électrodes inférieur à trois cent constitue le cœur des travaux menés dans cette thèse. Les expérimentations détaillées dans ce chapitre portent sur la navigation et son évaluation en VPS. Différents rendus sont créés et comparés au rendu brut afin de voir s'ils facilitent la navigation et la construction d'une représentation mentale des environnements parcourus. Lors de la première étude exploratoire nous avons utilisé un rendu qui s'est avéré performant dans des tâches de localisation d'objets, de textes ou de personnes [Denis et al. 2012], afin de l'évaluer dans une tâche de navigation. La deuxième expérimentation compare le rendu classique brut à un rendu permettant d'augmenter la perception des contrastes entre le sol et les murs de l'environnement. La troisième expérimentation poursuit ce travail en comparant différentes stratégies pour augmenter la structure de l'environnement. La dernière contribution enfin est un questionnaire qui a été soumis à des personnes voyantes et non-voyantes sur un trajet qu'ils connaissaient dans le centre-ville de Toulouse. Les résultats de ce questionnaire permettent d'alimenter la discussion sur les résultats des expériences précédentes et de tirer des lignes directrices sur le design de rendus prothétiques destinés à la navigation et comportant un nombre limité de phosphènes.

Chapitre V. Approche par localisation appliquée à la navigation

Dans la vie quotidienne, il est important de connaître précisément la position de certains objets ou personnes présentes autour de soi, sans forcément en percevoir la taille ou la forme. Pour rendre la position des objets, le rendu par localisation a été créé et évalué. Ce rendu consiste à reconnaître dans le champ caméra des éléments préalablement mémorisés afin de restituer précisément leurs positions.

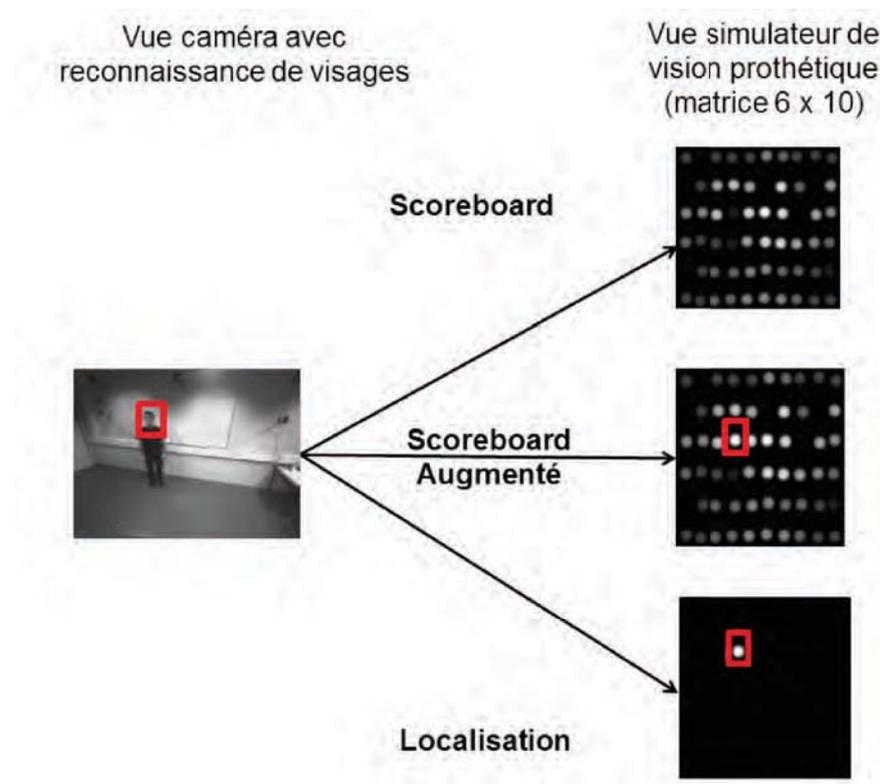


Figure 1 : Exemple de rendu par localisation testé dans notre équipe.
Le traitement de l'image permet de détecter les visages pour n'en présenter que la position aux sujets dans l'approche par localisation (Extrait de [Denis 2014]).

En VPS, ce principe a déjà été évalué pour compter le nombre de personnes présentes dans une pièce avec une matrice contenant un faible nombre d'électrodes [Denis et al. 2013]. Dans cette expérimentation, un algorithme de détection de visages est utilisé pour localiser les visages présents dans le champ de la caméra portée par le sujet. Les positions des différents visages sont alors représentées par un simple phosphène. Les résultats de cette expérience montrent l'efficacité du rendu par localisation avec des matrices de phosphènes de taille réduite (6x10 phosphènes). De même, ce rendu est efficace pour localiser et saisir un objet situé

devant soit avec seulement neuf phosphènes utilisés [Macé et al. 2015]. La même idée peut être appliquée pour la détection de blocs de texte autour de l'utilisateur en s'appuyant sur des algorithmes spécifiques. Un tel système pourrait donc aussi être utilisé pour alerter les utilisateurs de la présence de texte [Denis 2014].

Ce rendu n'a jamais été évalué pour une tâche de navigation. D'après le modèle de référence de Siegel et White [Siegel & White 1975], les premiers éléments qui permettent de commencer à structurer une représentation de l'environnement pour la navigation sont les points de repère (niveau *Landmark*). Ainsi avec un rendu par localisation affichant les points de repère il peut être possible de naviguer efficacement. Les sujets peuvent se diriger d'un point de repère à un autre pour effectuer un trajet déjà connu ou décrit étape par étape. Dans cette première étude exploratoire nous avons évalué l'approche par localisation dans le cadre de la navigation. Dans cette expérience, le rendu par localisation affichait les différents points de repère présents dans le champ caméra grâce à un code visuel. Les sujets devaient suivre des instructions et parcourir un trajet précis dans un environnement virtuel. Le moteur 3D qui a été utilisé pour la navigation dans les environnements 3D était le moteur Irrlicht¹. Ce rendu en VPS par localisation a été comparé avec le rendu brut et le rendu fourni directement par le moteur 3D Irrlicht. Les études faites en mobilité montrent qu'avec des implants de faible résolution, la navigation est difficile voire impossible [Cha et al. 1992b]. Même dans des environnements dont les murs sont très clairs et le sol très sombre, il a été montré qu'un long apprentissage peut être nécessaire [Wang et al. 2008]. Notre hypothèse sur cette expérimentation était que les sujets allaient obtenir de meilleures performances avec le rendu par localisation qu'avec le rendu brut.

A. Matériel et méthodes

A.1. Rendus

Cette première expérimentation se propose de comparer deux rendus en VPS différents, un troisième rendu fourni par le moteur 3D Irrlicht est utilisé en tant que contrôle.

Pour les rendus en VPS, nous avons choisi une matrice contenant deux cent soixante-dix phosphènes (matrice 15x18). Ce choix a été fait pour deux raisons principales :

- ◆ Les personnes implantées avec une matrice de soixante phosphènes ne sont pas capables d'effectuer réellement des tâches de navigation. Elles rencontrent déjà des difficultés importantes pour suivre une ligne blanche au sol ou repérer une porte sombre sur un mur blanc [Humayun et al. 2012]. Soixante phosphènes paraissent trop peu pour naviguer efficacement.
- ◆ Nous ne voulions pas utiliser de matrice de trop grande résolution pour rester proches des implants en phase de commercialisation ou en développement. La résolution de 15x18 correspond à la taille de l'implant Argus III de Second Sight actuellement en développement.

¹ <http://irrlicht.sourceforge.net/> (Dernier accès le 29/01/2016).

Ainsi la taille de matrice de 15x18 phosphènes est un bon compromis entre les implants déjà commercialisés et visiblement de trop faible résolution pour naviguer et les implants haute résolution qui ne sont pas encore en développement.

Les matrices de phosphènes affichées couvraient une partie du champ visuel de neuf degrés sur neuf degrés. Pour que les rendus de VPS soient proches des perceptions rapportées par les patients implantés les phosphènes étaient blanchâtres et de forme ronde avec un profil gaussien [Chen et al. 2009]. Certains patients implantés avec des électrodes rétiniennes parviennent à distinguer entre six à dix niveaux de gris pour chaque phosphène [Humayun et al. 2003]. Ainsi pour cette simulation le nombre de niveaux de gris utilisables était limité à huit, ce qui signifie que chaque phosphène ne pouvait prendre que huit valeurs de luminosité différentes. De plus, après implantation, certaines électrodes ne sont pas utilisables ou cessent de fonctionner. Pour cette expérimentation nous avons simulé ce dysfonctionnement des électrodes (ou phénomène de *dropout*) en éteignant de manière aléatoire dix pour cent des électrodes, ce qui correspond aux résultats de l'Argus II [Humayun et al. 2012].

Le premier rendu en VPS appelé « Brut » consistait en une réduction de la résolution de l'image fournie par le moteur 3D pour l'aligner sur la résolution de la matrice d'électrodes (15x18). Pour chaque portion de l'image correspondant à un phosphène, la moyenne de la luminosité était calculée en utilisant des outils inclus dans le moteur 3D Irrlicht. Le niveau de gris attribué à chaque phosphène était calculé par un système de seuils. Le phosphène prenait le niveau de gris situé immédiatement en dessous de la luminosité moyenne des pixels concernés.

Le second rendu en VPS appelé « Localisation » se basait sur l'approche par reconnaissance et localisation d'objets déjà utilisée dans des tâches de saisie d'objets ou de localisation de visages. L'environnement utilisé était une représentation simplifiée de couloirs et de bureaux. Les objets dont la reconnaissance et la localisation étaient simulées dans cette expérience étaient les points de repère dispersés dans l'environnement et correspondaient aux distributeurs d'eau, aux posters présents sur les murs et aux portes situées à l'entrée de chaque pièce. Ces éléments étaient les seuls éléments présents dans les environnements. Les posters étaient affichés sur les murs, les distributeurs d'eau étaient des objets physiques pouvant gêner la navigation et les portes indiquaient le passage d'une pièce à une autre. Il semblait important de pouvoir distinguer correctement ces éléments. La première distinction a été d'attribuer à chacun de ces types de repères un niveau de gris particulier. Ainsi les posters étaient les plus sombres en gris foncés, les distributeurs d'eau étaient en gris clair et les portes étaient représentées par des phosphènes encore plus clairs. Une deuxième distinction a été décidée pour distinguer encore plus précisément ces éléments. En plus d'être représentés par un niveau de gris particulier, un nombre de phosphènes précis leur a été attribué. Les posters étaient représentés par quatre phosphènes identifiant leurs coins. Trois phosphènes étaient utilisés pour représenter les portes, deux phosphènes correspondant aux coins supérieurs et un dernier indiquant la base de la porte. Les distributeurs d'eau étaient affichés par deux phosphènes correspondant à leur sommet et leur base.

Le dernier rendu correspondait à l'image générée par le moteur 3D Irrlicht; ce n'était donc pas une simulation de vision prothétique. Ce rendu appelé ici rendu « Irrlicht » a servi de contrôle dans le protocole expérimental et couvrait une surface du champ visuel de quarante degrés par trente degrés.

Ces trois rendus sont visibles dans la Figure 2.

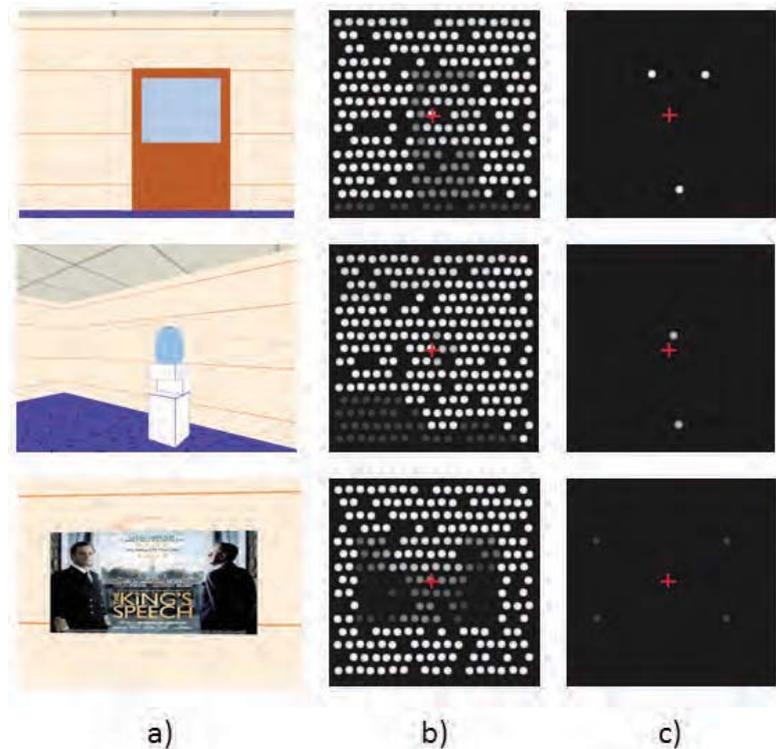


Figure 2 : Conditions et points de repère de l'expérience de navigation par localisation. Chaque ligne correspond à un point de repère donné. Dans l'ordre : une porte, un distributeur d'eau et un poster. Chaque colonne correspond à une condition : a) Irrlicht, b) Brut et c) Localisation. Dans l'approche par localisation, les portes sont rendues par trois points blancs, les distributeurs d'eau par deux points gris clair et les posters par quatre points gris foncé. Pour aider les sujets à fixer leur regard, une croix rouge fixe était présente au centre de l'écran.

Chez les patients implantés, les neuroprothèses visuelles sont fixées directement sur la rétine, le nerf optique ou l'aire corticale V1. Lors de l'utilisation de l'implant, la matrice de phosphènes est donc perçue à peu près au même endroit dans le champ visuel et suit le regard des sujets. Pour se rapprocher de cette perception, il était donc nécessaire d'asservir la matrice de phosphène au regard des sujets. Pour cette expérience, un oculomètre captait le regard des sujets afin de déplacer la matrice pour qu'elle soit asservie à la direction du regard. Pour aider les sujets à stabiliser au maximum leurs percepts, une croix rouge fixe est présente au centre de l'affichage. Accessoirement, cette croix rouge nous permettait aussi de détecter les pertes de calibrages. Ce principe est illustré dans la Figure 3.

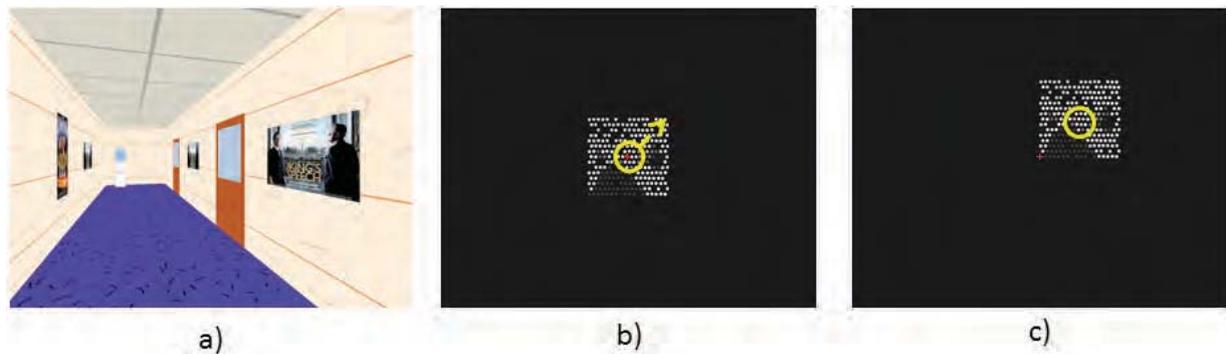


Figure 3 : Asservissement des percepts. a) Image de l'environnement générée par le moteur 3D Irrlicht. b) Quand le sujet dirige son regard (cercle jaune) sur la croix de fixation, la matrice de phosphènes est placée au centre de l'écran. La flèche jaune indique un futur déplacement du regard vers le haut et la droite. c) La position de la matrice de phosphènes est modifiée pour correspondre au déplacement du regard. Le sujet ne peut donc pas explorer la matrice de phosphènes. La croix rouge encourage le sujet à garder le regard droit devant.

A.2. Matériel

Les rendus prothétiques étaient affichés dans un visiocasque Nvisor SX-60 de la marque NVIS (Reston, Virginie). La résolution de ce casque était de 1280x1024 pixels pour couvrir le champ visuel vertical sur trente-quatre degrés et le champ visuel horizontal sur quarante-quatre degrés. Un oculomètre ViewPoint PC-60 de 60Hz et de la marque Arrington Research (Scottsdale, Arizona) était fixé sur le casque sous l'ocilleton destiné à l'œil gauche. Cet oculomètre permettait de capter la position de l'œil des sujets pour asservir la position de la matrice aux déplacements des yeux.

Pour cette expérience, différents programmes ont été codés et il était nécessaire de les faire fonctionner simultanément. Pour assurer la communication entre ces programmes, un bus logiciel de communication dénommé IVY² a été utilisé. Chacun des programmes pouvait envoyer des messages sur ce bus et recevoir ceux qui lui étaient destinés pour les lire. Les différents programmes sont présentés dans la Figure 4.

2 <http://www.eei.cena.fr/products/ivy/> (Dernier accès le 20/07/2015)

Schéma Fonctionnel du Dispositif d'Expérimentation

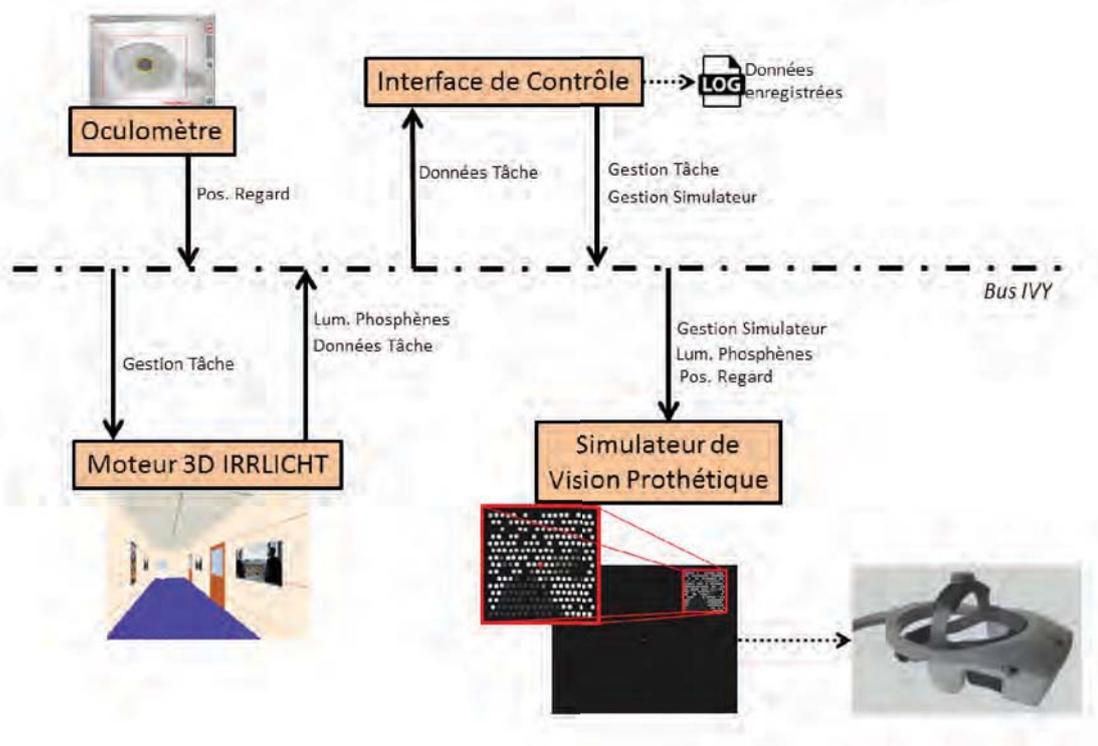


Figure 4 : Schéma fonctionnel du dispositif d'expérimentation utilisé. Quatre agents sont en fonctionnement simultanément. Tous sont connectés à un bus logiciel afin d'échanger des informations. Les flèches représentent les différents messages IVY réceptionnés par les agents ou envoyés au bus.

Quatre programmes différents étaient utilisés pour cette expérience :

- ◆ Agent gérant l'oculomètre (C++). Cet agent fournissait la position du regard au Simulateur de Vision Prothétique.
- ◆ Programme de gestion de la tâche utilisant le moteur 3D Irrlicht (C++, librairie Irrlicht). Ce programme recevait les instructions de l'Interface de Contrôle pour gérer la tâche et renvoyait à cette dernière les données liées à la tâche (position, orientation, passage de portes, etc.). Le moteur 3D Irrlicht affichait l'environnement pour l'expérimentateur et gérait la navigation. La navigation dans l'environnement virtuel devait prendre en compte les événements provenant du clavier de l'utilisateur ainsi que les collisions éventuelles au sein de l'environnement virtuel. Irrlicht générait aussi différents sons lors de la navigation : des bruits de pas lors des déplacements, des sons caractéristiques pour le début et la fin de la tâche ainsi que le franchissement des portes intermédiaires, enfin un son mat prévenait l'utilisateur lors d'une collision. Le calcul du niveau de gris pour chaque phosphène était effectué par ce programme qui en communiquait les résultats au Simulateur de Vision Prothétique.
- ◆ Simulateur de Vision Prothétique (C++, OpenGL). Ce programme réceptionnait la position de l'œil et la couleur des phosphènes pour afficher les rendus prothétiques simulés. Les caractéristiques de la matrice (profil des phosphènes, nombre de phosphènes, po-

sitions des phosphènes, nombre de phosphènes non fonctionnels, nombre de niveaux de gris) étaient chargées par le simulateur de vision prothétique à partir d'un fichier de configuration avant la tâche.

- ♦ Interface de Contrôle pour l'expérimentateur (JAVA). Cette interface permettait que l'expérimentateur contrôle le déroulement de la tâche. C'est également ce programme qui s'occupait d'enregistrer les différentes données récoltées pendant la tâche.

Un premier ordinateur (Windows 7, processeur Intel Core i5-2400) contenait l'exécutable du moteur 3D Irrlicht ainsi que celui gérant l'oculomètre. Un deuxième ordinateur (Windows 7, processeur Intel Core 2 Quad) était utilisé pour lancer l'interface de contrôle et le simulateur de vision prothétique.

B. Protocole

B.1. Sujets

Six sujets (quatre hommes et deux femmes) ayant entre 22 et 46 ans ont participé à l'expérience. Tous les sujets étaient des utilisateurs d'ordinateurs réguliers. Quatre sujets avaient déjà utilisé un visiocasque avant l'expérience, deux sujets n'avaient jamais porté de tels casques. Deux sujets étaient déjà familiers avec la vision prothétique simulée mais n'avaient jamais réalisé une tâche de navigation avec le simulateur.

B.2. Tâche

Les sujets étaient invités à s'asseoir face à un clavier d'ordinateur et étaient équipés du visiocasque. Avant de commencer la navigation, une minute était nécessaire pour calibrer l'oculomètre. Dans cette première expérimentation, la tâche effectuée était très proche de tâches déjà évaluées en VPS [Rheede et al. 2010]. Ce choix a été fait pour ne pas trop s'éloigner de la littérature en VPS. Les sujets avaient différents parcours à effectuer dans un environnement et devaient suivre des instructions verbales données au fur et à mesure par l'expérimentateur. Chaque trajet était balisé par sept portes à franchir. Au début de la tâche, les instructions pour atteindre la première porte étaient indiquées aux sujets. Les points de repères présents dans l'environnement étaient utilisés dans les instructions pour atteindre la porte suivante. Au franchissement d'une porte, les instructions nécessaires pour atteindre la porte suivante étaient données. Les sujets se déplaçaient dans l'environnement en utilisant les flèches directionnelles du clavier, ils devaient effectuer un appui long sur la flèche directionnelle avant pour avancer avec une vitesse constante. Les éventuelles collisions étaient indiquées par un son. Si un sujet se perdait ou exprimait explicitement des difficultés, une aide lui était fournie. L'aide pouvait prendre la forme d'instructions supplémentaires pour rapprocher les sujets du trajet souhaité ou d'instructions beaucoup plus précises pour ramener le sujet à la dernière porte franchie s'il s'était trop éloigné du trajet prévu. Le nombre d'aides fournies a été compté pour chaque trajet. L'oculomètre étant fixé sur le visiocasque, il arrivait régulièrement que la

tête des sujets bouge légèrement à l'intérieur du casque ce qui entraînait un décalage dans la mesure de position des yeux. Les sujets pouvaient détecter cela en observant que la matrice n'était pas au centre lorsqu'ils portaient leur regard sur la croix rouge centrale. Dans ce cas, la tâche était mise en pause et une dizaine de secondes était nécessaire pour calibrer de nouveau l'oculomètre. Ensuite la tâche reprenait normalement. Pour chaque condition, chaque sujet effectuait une fois l'itinéraire. Chaque sujet effectuait un parcours différent pour chaque condition. La répartition des combinaisons itinéraire-rendu était contre-balançée entre les sujets pour limiter l'impact possible des différences entre itinéraires. De même, l'impact de l'ordre des conditions a été limité par un contre-balancement lors des passations.

B.3. Variables dépendantes

Les performances étaient évaluées en utilisant le temps mis par les sujets entre le franchissement de la première porte et celui de la dernière porte. Le nombre de collisions avec les murs était également enregistré. Pour évaluer les connaissances spatiales acquises par les sujets lors de la tâche, ils dessinaient sur un fond de carte de l'environnement le trajet qu'ils pensaient avoir effectué. Il n'était pas demandé au sujet de restituer l'ensemble de l'environnement, mais seulement le trajet effectué dans cet environnement. La position de départ était indiquée sur les fonds de cartes fournis aux sujets. Une erreur était décomptée chaque fois que le dessin du trajet s'éloignait du trajet effectué (mauvaise direction à une intersection). Pour l'enchaînement de deux intersections, si une erreur était faite sur la première mais pas sur la deuxième, seulement une erreur était décomptée.

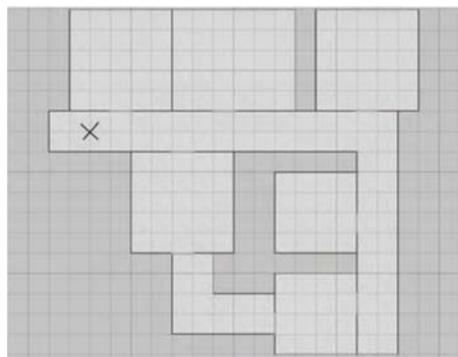


Figure 5 : Fond de carte utilisé par les sujets pour dessiner les trajets effectués.

Une fois toutes les conditions effectuées, les sujets répondaient à un questionnaire portant sur leur utilisation de l'ordinateur et leur expérience avec les systèmes de réalité virtuelle. Il leur était demandé d'évaluer la difficulté, l'immersion et la fatigue ressentie pour chacune des conditions sur une échelle de Likert à 5 niveaux.

B.4. Hypothèses

Notre hypothèse générale était que la navigation et la construction d'une représentation mentale des trajets effectués étaient facilitées en VPS par l'approche par localisation.

Nous supposons également que le rendu Irrlicht restait plus efficace pour naviguer que les deux rendus en VPS. Dans le cadre de notre expérimentation les hypothèses opérationnelles étaient donc les suivantes :

- ♦ Le temps mis par les sujets pour accomplir les trajets demandés serait plus court en rendu par localisation qu'avec le rendu brut.
- ♦ Les sujets feraient plus d'erreurs dans leurs dessins en rendu brut qu'en rendu par localisation.
- ♦ Le temps mis par les sujets pour effectuer les trajets serait plus court en rendu Irrlicht qu'avec les deux rendus en VPS.
- ♦ En rendu Irrlicht, les sujets feraient moins d'erreurs dans leurs dessins qu'avec les rendus VPS.

C. Résultats

C.1. Résultats quantitatifs

Les moyennes fournies dans ces résultats le sont avec la déviation standard. Le temps moyen pour effectuer les trajets avec le rendu par localisation (M : 551,6 s; DS : 135,0 s) est plus long que le temps moyen mis avec le rendu brut (M : 324,5 s; DS : 135,5 s). Ces deux conditions présentent des moyennes de temps plus longues que la condition contrôle (M : 97,5; DS : 18,5 secs). Avec la localisation, le nombre moyen d'aides est égal à 10,7 (DS : 6.8), avec le rendu brut ce nombre descend à 1 (DS : 2,5) et aucune aide ne fut nécessaire avec le rendu contrôle. Un test de Wilcoxon a montré que la différence dans le nombre d'aides entre le rendu brut et le rendu d'approche par localisation est significative ($Z = 2,023$; $p < 0,05$).

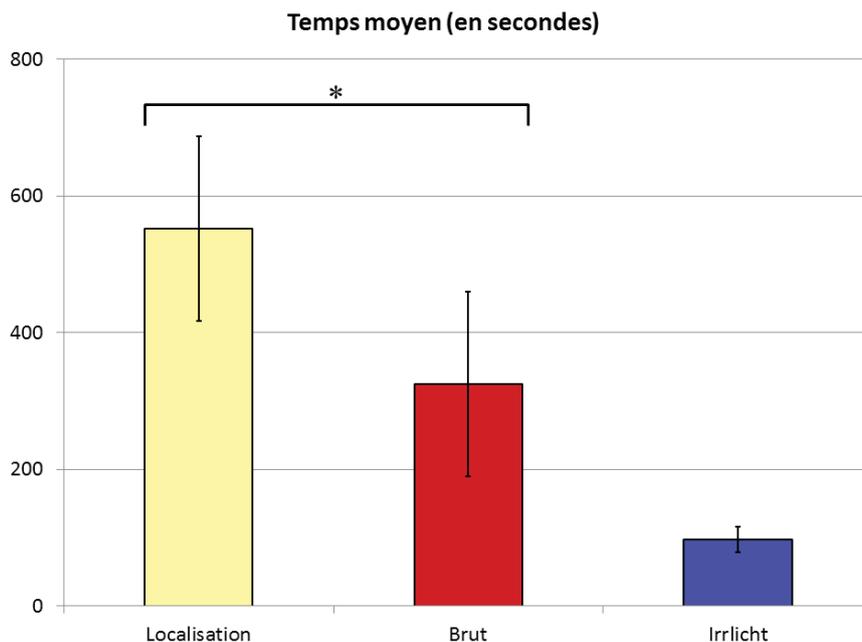


Figure 6 : Temps moyen mis par les sujets pour accomplir la tâche. Les barres d'erreurs sont les intervalles de confiance à 95 %.

En rendu brut aucune collision avec un mur ne fut enregistrée, alors qu'en localisation les sujets ont heurté au minimum 23 fois les murs.

Le nombre d'erreurs sur les dessins des trajets produits par les sujets est compris entre zéro et quatre pour le rendu par localisation (M : 1,8; DS : 1,5) et entre zéro et trois pour le rendu brut (M : 1,7; DS : 1,5). Ces données ne permettaient pas de montrer une distinction entre les deux rendus de vision prothétique. En revanche avec le rendu contrôle, les sujets n'ont fait aucune erreur. Une différence significative était présente entre le rendu VPS par localisation et le rendu contrôle ($p < 0,05$).

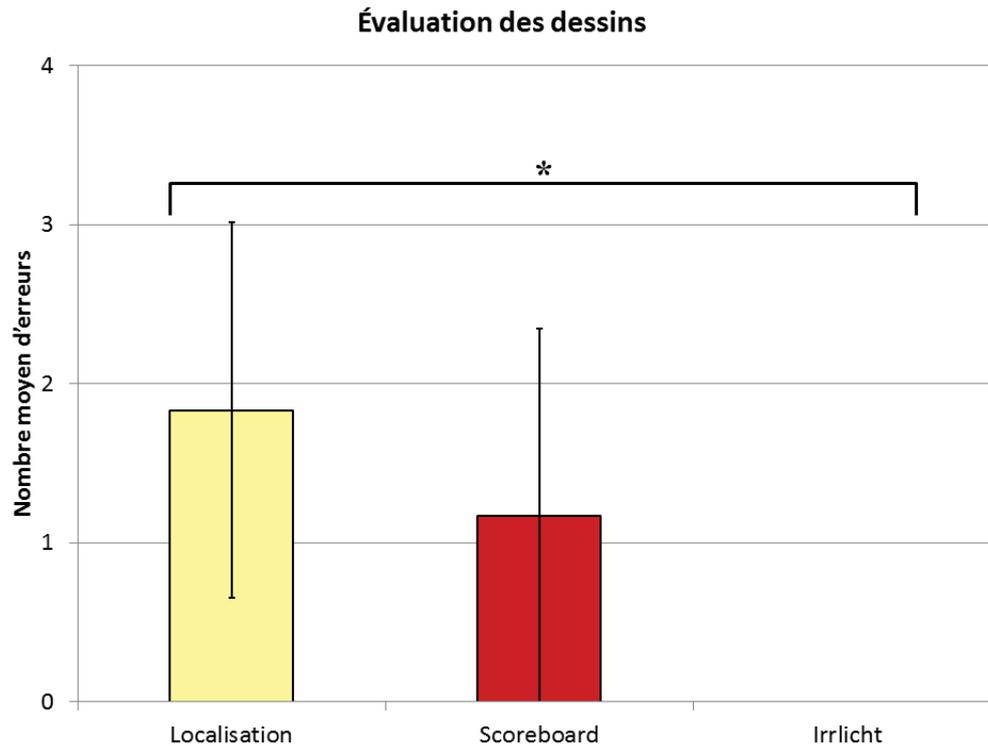


Figure 7 : Nombre moyen d'erreurs présentes dans les dessins pour chaque condition. En condition contrôle Irrlicht aucun sujet n'a fait d'erreur.

C.2. Résultats qualitatifs

Quatre sujets sur six ont évalué l'approche par localisation plus fatigante que le rendu brut. La difficulté de la tâche en rendu localisation a été évaluée à 4,7 sur 5 (DS : 0,5). La difficulté du rendu brut a été évaluée à 3,1 (DS : 0,2). En rendu Irrlicht cette difficulté a été évaluée à 1 (DS : 0). Quatre sujets sur six ont évalué le rendu brut comme plus facile que le rendu de localisation. Les sujets ont clairement signalé la difficulté à comprendre l'environnement en vision par localisation ainsi que la frustration générée par cette incompréhension. Cette incompréhension a également soulevé un stress pour certains sujets qui ont éprouvé de grandes difficultés pour mémoriser les éléments présents dans les trajets. Enfin certains sujets ont aussi expliqué leur fatigue par le poids du casque.

D. Discussion

L'objectif de cette expérimentation était d'évaluer l'approche par localisation dans le cadre de la navigation en rendu VPS. Il s'avère qu'au vu des performances pour effectuer une tâche de navigation, l'approche par localisation est plus difficile à utiliser (M : 551,6; DS : 154,1 secs) que l'approche du rendu brut classique (M : 324,5; DS : 154,6 secs). Ainsi, non seulement notre hypothèse n'a pas été vérifiée mais l'effet découvert va à contre-sens de l'hypothèse de départ. Pour cette première expérience exploratoire, nous avons suivi le modèle de Siegel et White [Siegel & White 1975] pour créer des rendus et analyser les performances que les sujets obtenaient en les utilisant. L'approche par localisation, déjà utilisée avec succès en vision prothétique [Denis 2014] fut testée ici en supposant que le fait de localiser les points de repère pouvait suffire pour naviguer en suivant des instructions. Les difficultés rencontrées par les sujets en approche par localisation indiquent que les points de repère seuls ne sont pas suffisants pour naviguer. Plusieurs causes peuvent être à l'origine de ces difficultés.

La principale difficulté en VPS avec l'approche par localisation peut être l'absence de contexte. En effet, en rendu brut, dans un environnement faiblement contrasté, l'évaluation de sa propre vitesse est possible grâce à l'arrière-plan, même si les éléments présents ne sont pas identifiables. Cet effet est d'autant plus vif que l'instruction de déplacement à l'aide du clavier ne permettait pas aux sujets d'utiliser d'indices kinesthésiques liés à la marche [Berthoz & Viaud-Delmon 1999]. L'absence complète de contexte et d'arrière-plan dans notre rendu par localisation empêchait cette évaluation entraînant des difficultés pour les sujets à s'orienter et à naviguer convenablement. À l'inverse, l'environnement que nous avons utilisé présentait un contraste fort entre le sol et les murs, il était ainsi aisé d'évaluer les rotations effectuées dans l'environnement avec le rendu brut.

Une autre limite de l'approche par localisation en navigation est la multiplication des items visibles. En effet lors de l'expérience les sujets pouvaient parfois percevoir différents points de repère simultanément. Les représentations phosphéniques de ces différents points de repère pouvaient alors se fondre, perturbant la compréhension de la scène pour les sujets. Pour désambiguïser le percept visuel, les sujets devaient se déplacer dans l'environnement jusqu'à ce que les différents éléments se distinguent clairement les uns des autres. Cet effet est visible dans la Figure 8. Le fait que des codes visuels précis soient utilisés plutôt que la surface apparente des points de repère a pu gêner les sujets qui ont mal compris certains phénomènes d'occlusions. Avec l'approche par localisation, les volumes des pièces de l'environnement dans lesquelles les sujets naviguaient n'étaient pas représentés, les angles entre les différents murs étaient absents, et il était difficile pour les sujets de détecter la fin d'un couloir ou une intersection.

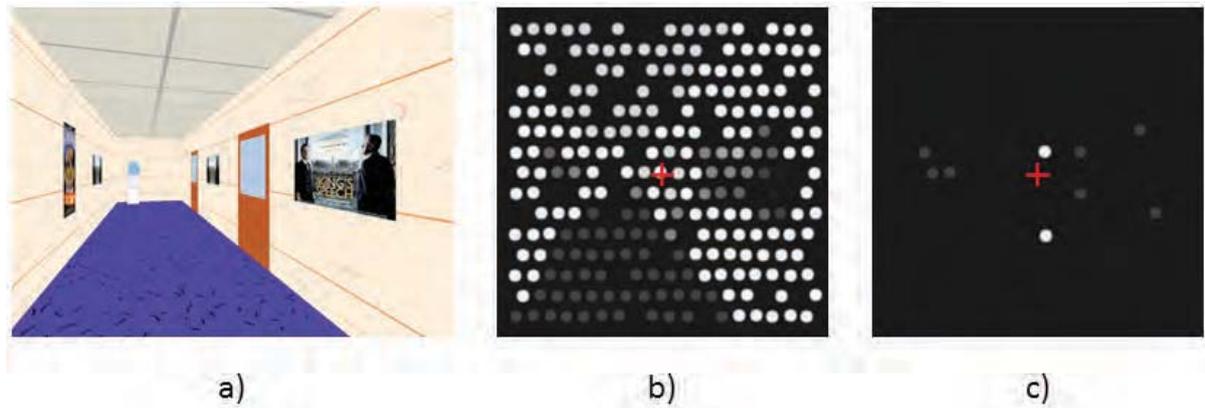


Figure 8 : Effet de superposition des items en VPS par localisation. a) Capture d'un couloir rendu par le moteur 3D Irrlicht. b) Rendu Brut. c) Rendu par localisation : seul le poster de droite est reconnaissable, la porte étant trop écrasée, elle peut facilement être confondue avec un distributeur d'eau. Le code correspondant au poster de gauche n'est pas entièrement visible non plus car l'angle supérieur droit se confond avec un des deux autres angles.

Par ailleurs, les sujets ont exprimé leur difficulté à utiliser les rendus de VPS et en particulier le fait que la matrice de phosphènes était difficile à stabiliser au centre du champ visuel, même avec l'aide de la croix de fixation. L'oculomètre était fixé à la structure du casque, ainsi lorsque la tête des sujets bougeait à l'intérieur du casque, même très légèrement, ceci avait un impact sur le calibrage qui était alors à recommencer. Le casque étant assez lourd à porter, même avec un système de contrepoids, certains sujets bougeaient involontairement la tête à l'intérieur du casque en voulant par exemple le soutenir à l'aide de leur main. Dans la condition de rendu brut, il était assez aisé de repérer alors le décalage de la matrice par rapport au regard. En revanche en vision par localisation, les phosphènes n'étant pas tous activés simultanément les sujets ne percevaient pas les bords de la matrice. Sans percevoir les bords, il pouvait être difficile de se rendre compte que la matrice était décalée par rapport au regard, ce qui augmentait d'autant plus la difficulté de la tâche.

Enfin l'évaluation des dessins faits après la tâche avec les deux rendus prothétiques montre une absence de différences entre les rendus VPS, alors que les performances des sujets en temps sont significativement différentes. Plusieurs hypothèses peuvent expliquer cet écart. La première est que les rendus ne présentaient pas de différence dans la compréhension de l'environnement. Cette hypothèse est cependant limitée par les différences de performances importantes entre les deux rendus phosphéniques. Une deuxième hypothèse est que la tâche elle-même incitait les sujets à avoir une représentation des lieux équivalente quel que soit le rendu prothétique utilisé. Lors de la tâche, les sujets étaient invités à suivre une série d'instructions pour aller d'une porte à une autre. Lors de la restitution du trajet sur un fond de carte, certains sujets ont pu se remémorer les instructions pour essayer de retrouver le trajet en question et ce pour les deux rendus prothétiques. Une troisième hypothèse est que l'évaluation faite sur les dessins ne permet pas de mesurer la représentation que les sujets ont de l'environnement. Lors de la navigation, les sujets ne disposaient que de la représentation des points de repère pour se repérer et il leur était demandé de restituer le trajet sur un fond de carte. Ainsi avec le rendu **Localisation** ils devaient restituer leur parcours dans l'environnement

uniquement à partir de sa structure qui n'était pas directement affichée durant la tâche et des points de repère mémorisés qui étaient absents du fond de carte. Ceci signifie que pour mémoriser leur trajet ils devaient inférer la structure à partir des points de repère perçus.

Lors de cette première expérience exploration, nous avons étudié deux rendus prothétiques sur de la navigation guidée. Cette expérience nous a permis d'approcher une limite du rendu par localisation qui est la tâche de navigation. Nous avons montré que ce rendu très efficace pour saisir des objets ou localiser des visages [Denis 2014] n'est pas fonctionnel pour une tâche de navigation même guidée. De plus, les résultats de cette expérience indiquent qu'il est nécessaire de fournir le contexte et des éléments de structure pour permettre une navigation et une mémorisation des trajets par les sujets. Dans une démarche incrémentale, nous avons amélioré ce protocole pour évaluer plus finement les performances de navigation et les représentations mentales construites par les sujets. Cette expérimentation nous a servis pour imaginer des rendus prothétiques plus adaptés pour la navigation et pour les évaluer dans une nouvelle tâche de navigation, ce qui est l'objet de l'expérimentation suivante.

Chapitre VI.

Contextualiser pour mieux naviguer en vision prothétique augmentée

L'expérience précédente a montré que le rendu par localisation des points de repère en VPS ne permet pas d'avoir des performances satisfaisantes. Si le nombre de points de repère visibles simultanément est trop faible, il est difficile de comprendre ses déplacements et si ces points de repère sont trop nombreux il est alors difficile de les distinguer les uns des autres. Les études de l'équipe de Dagnelie sur la navigation en VPS utilisent des rendus que l'on peut qualifier de fortement contrastés : le sol est affiché en sombre alors que les murs le sont en blanc [Wang et al. 2008]. De plus, ces environnements ne contiennent comme points de repère que des nombres affichés sur les murs des pièces. En réalité virtuelle, il est assez aisé de procéder ainsi mais des algorithmes de traitement d'images utilisant les images de caméra classiques ou de profondeur sont déjà capables d'isoler les différentes surfaces de l'environnement en temps réel [McCarthy et al. 2011]. En utilisant de tels algorithmes il serait possible de faire apparaître un percept plus contrasté entre le sol et les murs en environnement réel pour les personnes implantées.

La tâche et l'évaluation utilisées dans l'expérience précédente étaient très proches d'expériences déjà menées en VPS se focalisant sur la mobilité [Dowling et al. 2005]. Pour la suite des travaux nous souhaitons que la tâche de navigation donne plus d'importance à l'orientation, mais aussi pouvoir évaluer les représentations mentales de l'environnement construites lors de l'exploration et ce en lui laissant le choix de son trajet. Des différences dans la navigation peuvent apparaître entre des sujets utilisant des environnements réels et ceux utilisant des environnements virtuels [Zacharias 2006]. Afin de nous rapprocher des études en environnement réel et de permettre à l'utilisateur d'utiliser en partie son système vestibulaire, nous souhaitons que l'orientation de la caméra soit asservie à l'orientation de la tête [Riecke et al. 2010].

En rendu brut, si les éléments capturés par la caméra ne se distinguent pas les uns des autres par un fort contraste lumineux, il peut être difficile de faire la distinction entre différentes parois ou de repérer des points de repère. Dans cette deuxième expérience, nous avons donc cherché à rendre plus perceptible la structure de l'environnement. Nous avons également tenté de dépasser deux limitations présentes dans l'expérience précédente que sont la tâche effectuée et le mode de restitution. La tâche ne consistait plus à suivre des instructions pour parcourir un trajet mais à explorer un environnement inconnu. L'équipe de Sanchez [Sanchez et al. 2009] située au Chili a développé des dispositifs de réalité virtuelle pour les non-voyants. En particulier, un des environnements virtuels créés représentait le bâtiment du laboratoire et permettait à de jeunes aveugles d'apprendre la disposition des différents étages plus effica-

ement qu'en utilisant une technique classique de description des lieux. La tâche effectuée par les sujets dans ce bâtiment virtuel consistait à collecter un maximum de bijoux avant de sortir du bâtiment tout en évitant un ennemi présent qui pouvait prendre un de leurs bijoux pour le déposer ailleurs dans l'environnement. Cette tâche de navigation donnait un objectif aux sujets et fournissait une variable dépendante avec le nombre de bijoux récupérés par unité de temps pour évaluer la performance. L'expérimentation que nous avons mise en œuvre dans cette seconde étude s'inspirait très fortement de cette tâche et l'exploitait en VPS. Les sujets ont donc été amenés à naviguer activement avec la consigne supplémentaire de comprendre et mémoriser l'environnement. La méthode de restitution était également modifiée, les sujets dessinant la carte de l'environnement en partant d'une feuille vierge.

Pour cette expérimentation, trois rendus ont été évalués. Le premier rendu de VPS était le rendu brut. Pour le deuxième rendu, la structure de l'environnement a été augmentée pour produire le rendu brut augmenté. Finalement le rendu du moteur 3D Irrlicht fut également utilisé en tant que contrôle. Nous supposons que les sujets obtiendraient de meilleures performances de navigation avec le rendu brut augmenté qu'avec le rendu brut original. De même, le rendu brut augmenté devait permettre aux sujets de mieux se représenter l'environnement et donc de mieux le restituer qu'avec le rendu brut original. Nous supposons également que le rendu Irrlicht permettrait aux sujets d'avoir de meilleures performances et de produire des restitutions plus fidèles que les deux rendus en VPS.

A. Matériel et Méthodes

A.1. Environnements

Pour cette expérimentation nous voulions créer des environnements virtuels urbains et il a été décidé de mettre au point des environnements urbains simplifiés pour cette expérimentation³. Les trois environnements virtuels conçus représentaient des quartiers urbains simples avec des murs de cinq mètres de hauteur. Chacun d'entre eux contenait entre sept et neuf segments représentant des routes et définissant trois îlots au cœur des environnements. Tous les virages et les intersections présentaient des angles droits. Dans chaque environnement, les routes avaient une longueur maximale de deux cents mètres. Pour se déplacer, les sujets utilisaient les flèches directionnelles d'un clavier mis à disposition et la vitesse de déplacement était de trois mètres par seconde. Pour assurer une uniformité dans la complexité des environnements, les bâtiments du premier environnement ont été déplacés et réorientés pour mettre au point les deuxième et troisième environnements.

³ Le service Géomatique de Toulouse Métropole nous a fourni une archive contenant les informations 3D sur l'ensemble du territoire de Toulouse. Après avoir fait des tests et pré-expérimentation en utilisant ces données dans l'environnement Irrlicht, il s'est avéré que la navigation en utilisant ces données ne permettait pas de contrôler convenablement l'expérience (problèmes d'orientation des façades des bâtiments).



Figure 9 : Les trois environnements utilisés dans le cadre de cette expérimentation. Les joyaux sont représentés en bleu, la base de départ en Jaune. La flèche noire indique l'orientation de départ.

A.2. Rendus

La matrice de phosphènes utilisée pour les rendus prothétiques contenait 15x18 phosphènes ce qui correspond à la résolution de l'implant de la prochaine génération de l'entreprise Second Sight : l'Argus III [Stronks & Dagnelie 2014]. Comme pour l'expérience précédente, la matrice couvrait 9x9 degrés d'angle visuel. Pour être proche du fonctionnement des implants actuels, un dysfonctionnement de 10 % des électrodes était simulé, c'est-à-dire que 10 % des phosphènes étaient inactifs en permanence [Chen et al. 2009]. Chaque phosphène avait l'apparence d'une tache ronde avec un profil gaussien. Nous avons décidé de limiter le nombre de niveaux de gris disponibles par rapport à l'expérimentation précédente pour se rapprocher un peu plus des retours des personnes implantées. Si une personne implantée est parvenue à distinguer dix niveaux de gris, la majorité d'entre eux en distinguent moins de six [Stingl et al. 2015]. Quatre niveaux de gris étaient utilisés; ainsi pour chaque phosphène, l'intensité lumineuse était réduite à une des quatre valeurs possibles.

Le premier rendu en VPS appelé rendu **Brut** était conçu en redimensionnant les images du flux de départ pour les faire correspondre à la résolution de la matrice, comme pour l'expérimentation précédente. Dans ce rendu, il n'y avait pas d'augmentation de contraste entre le sol et les murs, en revanche le ciel était clairement perceptible car très blanc (voir Figure 10-b).

Le deuxième rendu de vision prothétique simulée appelé rendu **Brut Augmenté** s'appuyait sur l'existence d'algorithmes permettant de distinguer les différentes surfaces de l'environnement. Ainsi l'utilisation d'un tel algorithme était simulée en augmentant le contraste perçu entre le sol et les murs. Dans ce rendu, le sol était assombri et se démarquait des murs de l'environnement qui étaient représentés en gris. Les trois rendus sont visibles dans la Figure 10. Le dernier rendu était le rendu **Irrlicht** fourni directement par le moteur 3D Irrlicht.

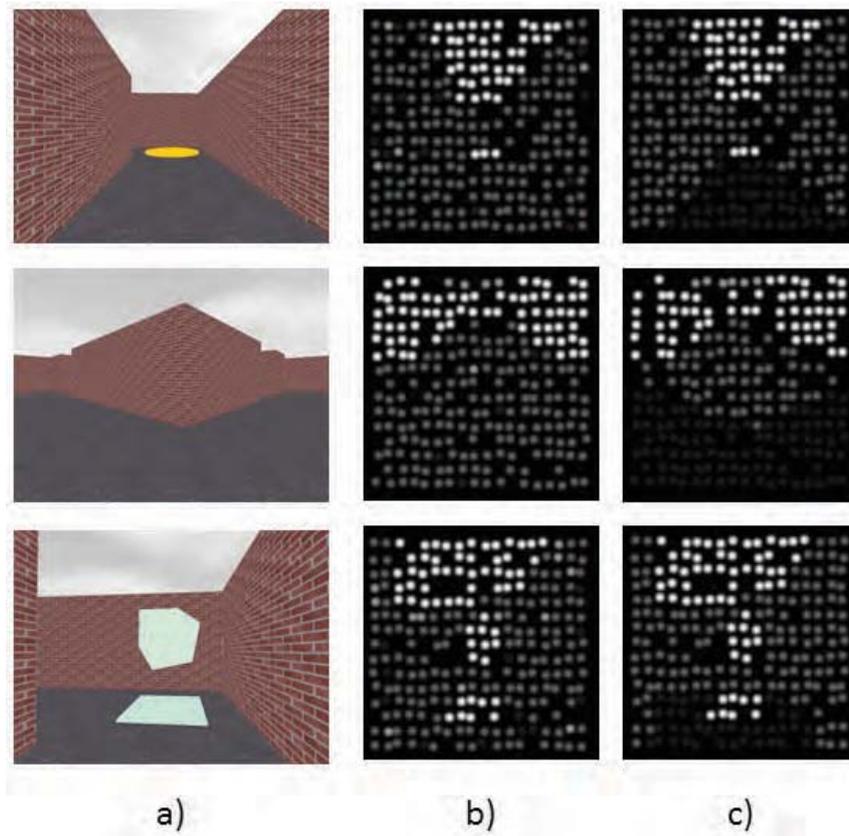


Figure 10 : Présentation des trois rendus de la deuxième expérimentation.
 a) Le rendu Irrlicht utilisé ici comme condition contrôlée. b) Rendu brut classique.
 c) Rendu brut augmenté. La base de départ est affichée sur la première ligne.
 Un angle dans l'environnement est représenté dans la deuxième ligne.
 La troisième ligne montre la perception d'un joyau.

A.3. Tâche

Avant de démarrer la tâche, les sujets s'entraînaient dans un environnement de familiarisation plus petit que les environnements de test. Il leur était expliqué alors les contrôles et le rendu qu'ils allaient utiliser ensuite. Les instructions pour la tâche principale étaient rappelées ainsi que les objectifs (récolter le maximum de joyaux et mémoriser la configuration générale de l'environnement). Ils étaient informés au préalable qu'ils allaient devoir dessiner les environnements explorés. Pour la tâche de familiarisation, les sujets récoltaient seulement trois joyaux. Une fois la familiarisation effectuée, la tâche principale pouvait démarrer. Les sujets disposaient de cinq minutes pour explorer l'environnement. Ils devaient collecter et ramener à la base neuf joyaux dispersés dans l'environnement. Les sujets avaient l'instruction de ramener les joyaux aussi vite que possible et de mémoriser l'environnement. Ces deux objectifs n'étaient pas hiérarchisés. Pour collecter un joyau il suffisait de l'atteindre dans l'environnement virtuel puis de revenir à la base de départ où il était virtuellement déposé. La base de départ était matérialisée dans l'environnement par un socle circulaire très clair. Chaque joyau était constitué d'un cube à hauteur de la caméra ainsi que d'un socle carré, ces deux éléments étaient très contrastés également. Lorsqu'un joyau était ramassé, le cube et le socle correspondant disparaissaient. Pour encourager les sujets à mémoriser l'environnement ils ne

pouvaient transporter que trois bijoux simultanément, ce qui signifie qu'ils devaient d'effectuer au moins trois allers et retours dans l'environnement pour collecter tous les bijoux. Pour effectuer la tâche les sujets disposaient de cinq minutes maximum. Une fois qu'ils avaient ramené tous les bijoux ou que les cinq minutes étaient écoulées, la navigation prenait fin.

Les sujets étaient invités à recommencer la même tâche deux fois de plus dans les mêmes conditions. Après ces trois essais, les sujets dessinaient une carte de l'environnement en vue aérienne, allocentrique. Ils avaient comme instructions de dessiner les couloirs ainsi que l'emplacement des bijoux et de la base de départ. Les sujets étaient informés de ces instructions avant de commencer la tâche. Une fois le dessin effectué, les sujets passaient à la tâche de familiarisation pour la condition suivante et ainsi de suite. Sur l'ensemble de la passation, les sujets effectuaient donc trois fois la familiarisation, neuf fois la tâche principale de navigation et produisaient trois dessins. L'ordre des conditions ainsi que les associations entre les conditions et les environnements étaient contrebalancés entre sujets.

Une fois l'expérimentation proprement dite terminée, les sujets étaient invités à remplir un rapide questionnaire contenant cinq questions portant sur le plaisir et la difficulté qu'ils avaient eue à effectuer la tâche. Ils indiquaient également si les rendus étaient utilisables et fournissaient suffisamment d'informations pour accomplir la tâche. Finalement ils indiquaient la fréquence à laquelle ils jouaient aux jeux vidéo.

A.4. Matériel

L'environnement virtuel en rendus contrôle ou prothétiques était affiché via le même visiocasque NVisor SX-60 HMD que celui utilisé lors de l'expérimentation précédente. L'orientation de la caméra virtuelle était dépendante de l'orientation du casque, un système de capture du mouvement OptiTrack, fourni par la société Natural Point (États-Unis) permettait d'utiliser l'orientation du casque pour diriger la caméra. Les sujets étaient assis et utilisaient les flèches directionnelles d'un clavier pour se déplacer dans l'environnement. Le bruit de leur pas et les chocs avec les murs étaient indiqués par des sons.

Lors de la première expérience, la matrice de phosphènes était asservie au regard des sujets grâce à un oculomètre. Dans cette expérience nous avons voulu favoriser l'immersion des sujets en les autorisant à tourner la tête dans l'environnement. Si nous avons ajouté l'utilisation de l'oculomètre au dispositif nous aurions eu des erreurs importantes sur la mesure de la position de l'œil à cause des mouvements de tête des sujets. Ainsi dans cette expérience nous n'utilisons pas d'oculomètre et la position des matrices de phosphènes n'était pas asservie à la position du regard.

Pour les programmes utilisés, le dispositif recourait à des agents très proches de ceux de l'expérience précédente. Ces différents agents communiquaient via le bus IVY et étaient répartis sur deux ordinateurs, le fonctionnement du système est visible dans la Figure 11.

Les agents utilisés étaient donc les suivants :

- ◆ Agent de Motion Capture (C++)

- ◆ Programme de gestion de la tâche utilisant le moteur 3D Irrlicht (C++, librairie Irrlicht)
- ◆ Simulateur de Vision Prothétique (C++, OpenGL)
- ◆ Interface de Contrôle pour l'expérimentateur (JAVA)

La principale différence par rapport à l'expérience précédente était que le système de capture de mouvement communiquait la position et l'orientation du casque au moteur 3D Irrlicht qui en répercutait les mouvements à la caméra virtuelle. Au niveau de la programmation, l'agent utilisant le moteur 3D Irrlicht a été grandement modifié pour intégrer ces mouvements ainsi que pour gérer la tâche d'exploration et de récolte des bijoux qui était absente de l'expérimentation précédente.

Schéma Fonctionnel du Dispositif d'Expérimentation

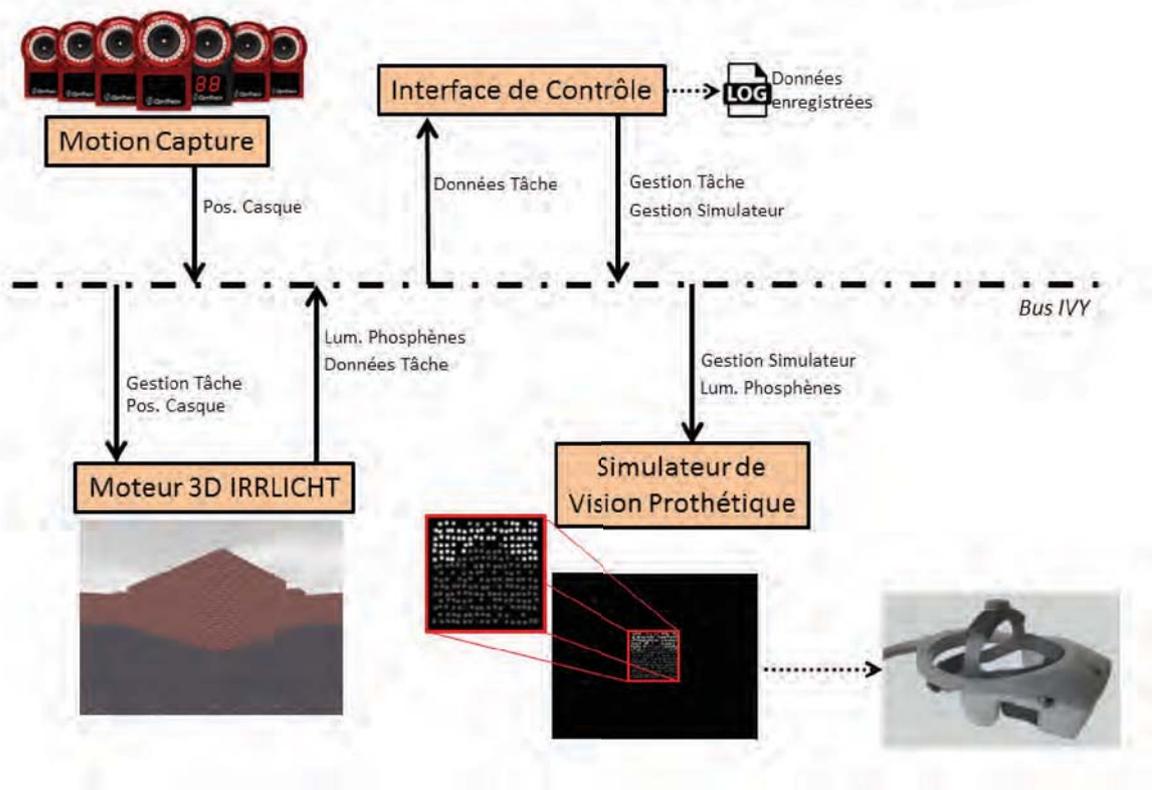


Figure 11 : Schéma fonctionnel des différents programmes utilisés dans la deuxième expérimentation. Les caméras de motion capture permettent de mesurer les déplacements du casque pour les transmettre au moteur 3D Irrlicht.

A.5. Sujets

Douze sujets (six hommes et six femmes ayant entre 23 et 39 ans) ont participé à cette expérience. Chaque participant avait une vision normale ou corrigée. Le protocole de cette expérience a été validé au préalable par le comité d'éthique CLERIT de l'Université de Toulouse. Tous les sujets ont signé un accord de consentement éclairé avant de commencer l'expérimentation.

A.6. Variables dépendantes

Comme décrit précédemment, l'objectif de la tâche était de collecter et ramener neuf bijoux à la base de départ en moins de cinq minutes. Si les neuf bijoux étaient ramenés avec cette contrainte, le temps pour effectuer la tâche était enregistré. À l'inverse, si moins de neuf bijoux étaient ramenés au bout des cinq minutes, le nombre de bijoux ramenés à la base était enregistré. L'indice de performance (PI) utilisé a été calculé de la manière suivante :

$$PI = \frac{J_i - J_{MIN}}{J_{MAX} - J_{MIN}} * 50 + \frac{T_{MAX} - T_i}{T_{MAX} - T_{MIN}} * 50$$

Avec :

J_{MIN} le nombre minimal possible de bijoux collectés (ici $J_{MIN} = 0$),

J_{MAX} le nombre maximal possible de bijoux collectés (ici $J_{MAX} = 9$),

J_i le nombre de bijoux collectés par le sujet,

T_{MIN} le temps minimal théorique calculé sur l'environnement en question,

T_{MAX} le temps maximal autorisé (ici $T_{MAX} = 5 \text{ min}$),

T_i le temps mis par les sujets pour accomplir la tâche.

Si un sujet parvenait à ramener les neuf bijoux en un temps très proche des cinq minutes, son indice de performance était très proche des 50 %. Si un sujet parvenait à ramener les neuf bijoux en moins de cinq minutes, son indice de performance était supérieur à 50 %. Le temps minimal théorique a été calculé en produisant pour chacun des environnements le trajet de longueur minimale pour récupérer les neuf bijoux et le temps nécessaire pour le parcourir.

Trois juges ont évalué la qualité des cartes dessinées par les sujets pour chaque condition. Les juges n'étaient pas impliqués dans l'expérience et ne connaissaient pas les hypothèses de l'expérience. Ils avaient comme instruction d'évaluer la similarité entre les dessins et les vraies cartes des environnements sans tenir compte d'éventuelles approximations sur les distances. Les scores attribués aux dessins pouvaient s'étaler sur une échelle de 0 (dessin et carte complètement différents) à 10 (dessin et carte parfaitement semblables). Chaque dessin produit par un sujet s'est donc vu attribuer trois notes par les juges.

A.7. Hypothèses

Nous supposons que les deux rendus prothétiques étaient moins efficaces que le rendu contrôle Irrlicht. Nous supposons également que le rendu brut augmenté était plus efficace que le rendu brut. Par voie de conséquence les hypothèses opérationnelles étaient les suivantes :

- ◆ L'indice de performance obtenu avec le rendu Irrlicht était plus élevé que l'indice de performance obtenu avec les rendus brut et brut augmenté.
- ◆ De même, la moyenne obtenue par les dessins en condition de rendu Irrlicht était plus élevée que celle obtenue dans les conditions de rendu brut et brut augmenté.

- ◆ L'indice de performance obtenu avec le rendu brut augmenté était plus élevé que l'indice de performance obtenu avec le rendu brut.
- ◆ De même, la moyenne des dessins en condition de rendu brut augmenté était plus élevée que la moyenne des dessins en condition de rendu brut.

Pour effectuer toutes les statistiques, le logiciel R (v3.0.1) a été utilisé. Les données ne suivant pas une loi normale, nous avons utilisé le test de Wilcoxon avec un seuil de signification fixé à $p < 0.05$ et ajusté avec une correction de Bonferroni pour les comparaisons multiples.

B. Résultats

B.1. Résultats quantitatifs

Avec le rendu brut, les sujets ont ramené en moyenne 5,2 bijoux (DS : 2,3) alors qu'en moyenne 7,3 (DS : 2,3) bijoux ont été collectés avec le rendu brut augmenté, cette différence était significative ($Z=4,4$; $p < 0,001$). De même, le temps moyen en seconde pour effectuer la tâche était significativement supérieur ($Z=3,8$; $p < 0,001$) avec le rendu brut ($M : 295,1$; DS : 21,6) qu'avec le rendu brut augmenté ($M : 265,8$; DS : 47,2). Ainsi, l'indice de performance différait significativement entre ces deux rendus ($Z=4,6$; $p < 0,001$). Cet indice était égal à 30,5 (DS : 16,7) en rendu brut simple et 52,1 (DS : 24,2) en rendu brut augmenté. La moyenne des indices de performance en rendu Irrlicht était élevée ($M : 81,5$; DS : 16,5) et significativement supérieure aux deux autres conditions (brut : $Z=5,2$; $p < 0,001$ et brut augmenté : $Z=5,0$; $p < 0,001$). Ces résultats sont visibles dans la Figure 12. À noter qu'il n'y a pas de différence significative entre les trois environnements utilisés dans l'expérience ($p=1$ pour chacune des comparaisons).

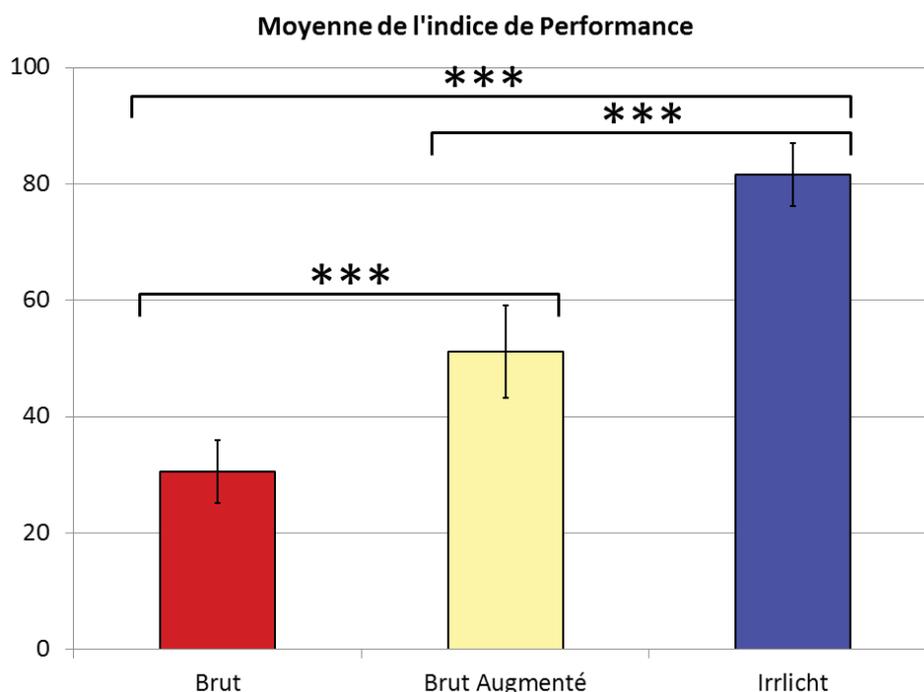


Figure 12 : Indice de performance des trois conditions. Les sujets ont effectué la tâche plus rapidement en condition augmentée qu'en condition brut simple. Les barres d'erreurs sont les intervalles de confiance à 95 %.

La moyenne des dessins pour le rendu de brut simple était égale à 2,69 (DS : 2,14) et était significativement inférieure à la moyenne des notes obtenues avec le rendu de brut augmenté (M : 4,88 DS : 3,00; $Z=3,1$; $p < 0,01$). La moyenne des notes pour la condition contrôle était égale à 7,86 (DS : 1,57) et était significativement supérieure aux moyennes des deux autres rendus (brut : $Z=5,7$; $p < 0,001$ et brut augmenté : $Z=4,5$; $p < 0,001$). Comme pour l'indice de performance, les scores des dessins n'étaient pas impactés par les différents environnements utilisés ($p=1$ pour toutes les comparaisons).

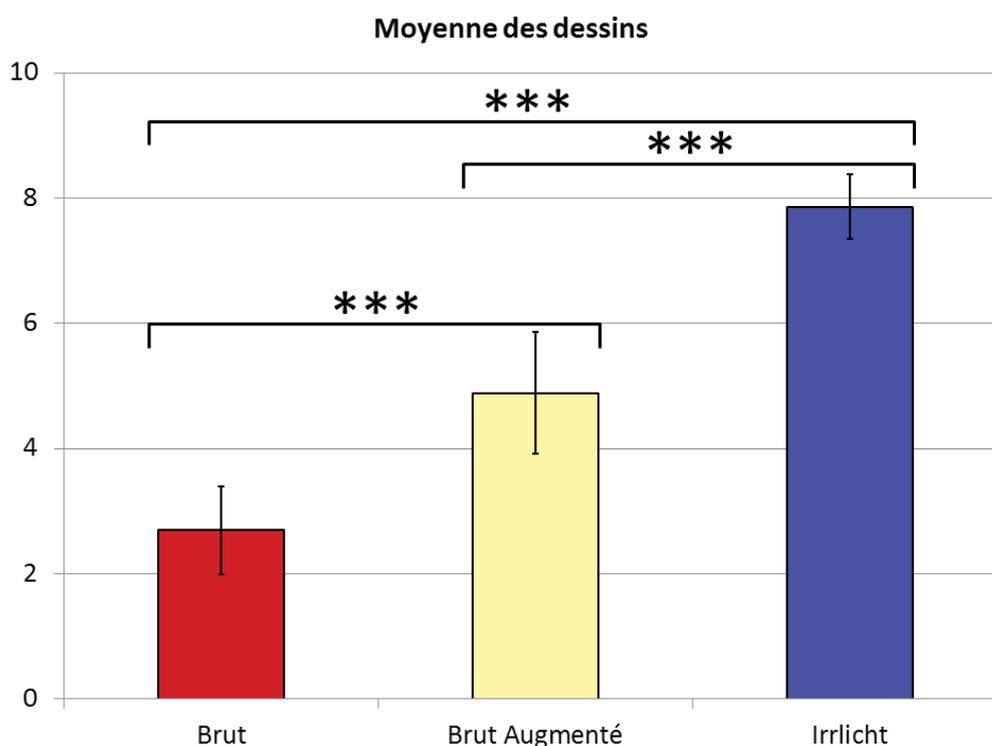


Figure 13 : Moyenne des dessins de carte par condition. Le fait d'améliorer la perception des contrastes entre le sol et les murs améliore l'image mentale des sujets. Les barres d'erreurs sont les intervalles de confiance à 95 %.

B.2. Résultats qualitatifs

Onze sujets sur douze ont évalué la tâche plus facile et plus agréable avec le rendu brut augmenté qu'avec le rendu brut simple. La même proportion de sujets a trouvé le rendu brut augmenté plus utilisable que le rendu brut simple. Dix sujets ont indiqué que le rendu brut augmenté apportait assez d'information pour effectuer la tâche, seulement trois sujets ont rapporté la même chose pour le rendu brut simple.

Les deux sujets ayant les meilleures performances en brut augmenté étaient les seuls à se considérer comme des joueurs passionnés. Les quatre sujets avec les performances les plus faibles avec les rendus prothétiques simulés étaient quant à eux non-joueurs.

C. Discussion

Dans cette expérience, la capacité à naviguer et à élaborer une représentation mentale de l'environnement avec l'aide de rendus prothétiques a été évaluée. Ces rendus correspondent à des implants de basse résolution de la prochaine génération d'implants. Les résultats obtenus avec l'indice de performance et le score des dessins évalués par les juges valident l'ensemble des hypothèses opérationnelles de l'expérimentation. Ainsi, le fait d'augmenter la structure en VPS permet d'effectuer une navigation plus efficace qu'avec un rendu brut classique. De même, l'augmentation de la perception de la structure permet également de construire une représentation spatiale plus juste des environnements dans lesquels la navigation s'effectue. Cette différence a pu s'observer sur les restitutions des environnements dessinées par les sujets.

Lors de l'expérience précédente, la position de la matrice de phosphènes était asservie au regard des sujets. Un oculomètre fixé sur le visiocasque porté par les sujets permettait de capter la position de l'œil gauche des sujets. Si les sujets bougeaient la tête à l'intérieur du casque et détectaient un décalage de la matrice il était alors nécessaire d'interrompre la tâche, de procéder à une nouvelle calibration de l'oculomètre avant de poursuivre la tâche. Il s'est avéré que ce décalage se produisait régulièrement même si les sujets avaient comme instructions d'éviter de bouger la tête lors de la tâche. Pour cette expérimentation, nous souhaitons que la caméra virtuelle soit asservie aux mouvements de la tête effectués par les sujets pour améliorer le sentiment de présence [Sanchez-Vives & Slater 2005]. Si nous avons conservé l'asservissement de la position de la matrice au regard, cela aurait engendré des interruptions régulières de la tâche pour calibrer l'oculomètre ce que nous souhaitons éviter. Pour cette raison, il a été fait le choix de ne pas asservir la position de la matrice au regard des sujets. En conséquence, contrairement aux patients implantés les sujets de cette expérience avaient la possibilité de balayer du regard la matrice de phosphènes. Il leur a été demandé de limiter les balayages et de se concentrer sur le centre de la matrice.

La tâche demandée aux sujets permettait de montrer des différences dans la performance de navigation mais également dans la mémorisation de la structure de l'environnement. Contrairement à l'expérience précédente et aux études antérieures portant sur la mobilité en VPS [Dagnelie et al. 2007; Rheede et al. 2010; Zapf et al. 2015] les sujets sont ici dans une situation de navigation où l'orientation est importante et où des décisions sont à prendre pour optimiser son trajet. Il s'agit ici de la première étude portant sur la navigation dans un environnement inconnu sans indications précises. La restitution par le dessin a permis de montrer des différences importantes entre les différentes conditions de cette expérience. Lors de l'expérience précédente, pour effectuer le dessin d'un trajet sur un fond de carte, il était suffisant de mémoriser une séquence d'actions (représentation de niveau « Route » de l'environnement) et d'indiquer les étapes qui ont été suivies [Siegel & White 1975]. Pour restituer une carte complète à partir d'une feuille blanche, il est nécessaire d'avoir des connaissances de niveau « configuration » (« Survey »).

Plusieurs remarques peuvent être formulées sur cette expérience. En premier lieu, contrairement aux environnements de l'expérience précédente, ceux-ci ne présentaient pas

de points de repère à proprement parler. En effet seule la base de départ peut être considérée comme un point de repère car elle est pérenne; elle ne disparaît pas pendant l'exploration contrairement aux joyaux qui ont été collectés. Or selon la théorie de Siegel et White, les points de repère sont la base des représentations spatiales. Ceci signifie donc que d'autres éléments de l'environnement sont venus structurer la représentation mentale des sujets. La notion de *vista space* [Montello 1993] peut être utilisée ici pour analyser la construction des représentations. Un *vista space* est défini comme un espace observable depuis un seul point en l'absence de déplacement (sans locomotion). Cette notion est à la base du modèle de *Network of Reference Frame* de Meilinger [Meilinger 2008] qui indique que les représentations spatiales se construiraient non autour des points de repère mais plutôt sur la présence de *vista spaces* qui contiennent des points de repère locaux. Les espaces marquants qui pouvaient servir de *vista space* dans nos environnements virtuels semblent être ceux contenant la base de départ (qui est bien un point de repère ici) ainsi que les intersections et les renforcements contenant les joyaux qui sont des points de décision pour les sujets. En particulier les croisements en croix qui sont peu nombreux dans les environnements ont sans doute été très utilisés car ils étaient discernables des autres intersections. Nous supposons qu'avec le rendu brut augmenté, les sujets sont mieux parvenus à mémoriser certains *vista spaces* et à faire des liens entre eux.

Lors de l'expérimentation, certains sujets se sont grandement appuyés sur le fait que le ciel soit plus clair que les murs pour se déplacer dans l'environnement et naviguer relativement efficacement. Ces sujets n'étaient pas très nombreux, mais il est possible que cette aide supplémentaire ait limité les différences entre les rendus. De fait, la présence d'un ciel plus clair permettait de percevoir distinctement le haut des murs et d'en inférer la structure de l'environnement. Lors de l'utilisation quotidienne de ces implants, le ciel ne peut se démarquer autant en permanence. Cependant, lorsque c'est le cas, cet indice est utilisable avec des implants de faible résolution contenant moins de trois cents électrodes.

Deux limitations méthodologiques ressortent de cette expérimentation : la première est la gestion du temps d'exposition aux environnements. Les objectifs donnés (performance et mémorisation de la configuration spatiale) n'étaient pas ordonnés. Si la plupart des sujets ont favorisé la performance de temps, certains se sont concentrés sur le fait de mémoriser au mieux la configuration de l'environnement. Certains sujets ont par exemple rebroussé chemin pour revenir à une intersection ou fait des détours pour se remémorer un lieu précis de l'environnement. Pour corriger ce biais, une méthode peut être d'imposer aux sujets le même temps d'exposition aux environnements et de hiérarchiser les objectifs. La deuxième limitation est liée à l'utilisation d'environnements virtuels. La familiarité de certains sujets avec ce type d'environnements leur a sans doute permis d'obtenir de meilleures performances que d'autres sujets naïfs. Il semble que les sujets jouant le plus aux jeux vidéo sont ceux qui sont parvenus à récolter les joyaux le plus vite que les autres. La technique d'interaction utilisant les flèches du clavier pour se mouvoir au sein des environnements a pu également favoriser les joueurs et déstabiliser certains non-joueurs. De même, certains sujets non-joueurs ont

commencé par les rendus VPS qu'ils ont eu beaucoup de mal à interpréter et ont été très surpris lorsqu'ils sont passés sur le rendu contrôle. Pour éviter ce phénomène, il est nécessaire de familiariser les sujets avec les environnements utilisés avant de leur faire effectuer la tâche évaluée. Ceci permettra aux sujets d'intégrer les dimensions de l'environnement et d'interpréter plus facilement les rendus VPS. Dans l'expérimentation suivante nous avons essayé de limiter ces deux biais potentiels.

D. Conclusion

Dans cette expérience en vision prothétique simulée, nous avons évalué les performances de sujets sur une tâche d'exploration avec deux rendus de VPS. Deux principaux résultats sont à retenir : premièrement, nous confirmons la possibilité d'explorer activement des environnements inconnus avec un rendu de VPS utilisant moins de trois cents phosphènes et seulement quatre niveaux de gris. Ceci est possible en utilisant un rendu prothétique augmentant la structure : le sol est détecté et affiché avec des phosphènes plus sombres que les murs de l'environnement. Lors de la première expérimentation de navigation en VPS, Cha et collaborateurs ont évalué que six cents phosphènes étaient nécessaires pour naviguer [Cha et al. 1992b]. Cette expérimentation montre qu'en s'appuyant sur la vision par ordinateur, il est possible de naviguer et de se construire une représentation mentale de l'environnement avec moins de trois cents phosphènes. Le deuxième résultat important est qu'avec moins de trois cents électrodes les sujets étaient également capables d'acquérir une représentation mentale de l'environnement lors de la navigation et de la restituer via un dessin.

Chapitre VII.

Différentes méthodes pour augmenter la structure en VPS

Chapitre VII. Augmenter la structure en VPS

L'expérience précédente a montré que l'augmentation du rendu brut en améliorant la perception du contraste entre le sol et les murs permettait d'obtenir de bonnes performances de navigation ainsi qu'une mémorisation correcte de l'environnement. Avec le rendu classique brut les sujets ont eu des performances plus faibles pour la navigation et la mémorisation des environnements. Le rendu brut augmenté simplifiait le rendu brut classique pour les sujets en permettant de distinguer le sol des murs dans les environnements. Les mauvaises performances obtenues avec le rendu brut semblent provenir d'un encombrement dans le champ de la caméra : les éléments présents sont trop nombreux et ne se distinguent pas assez les uns des autres pour être affichés correctement avec une résolution aussi faible. D'autres méthodes de rendu qui permettent de rendre plus visible la structure de l'environnement ont été produites. Il s'agit par exemple des rendus utilisant la distance mesurée entre la caméra et les éléments plutôt que leur luminosité. Des caméras de distance seront bientôt accessibles en mobilité (comme le système Intel® RealSense™⁴ ou le projet Tango de Google⁵) et les rendus utilisant directement ces informations seront donc peu coûteux en calcul.

Différents rendus de VPS se basant sur la distance aux objets ont déjà été proposés [Barnes et al. 2011] et certains d'entre eux ont été évalués dans des tâches de mobilité [McCarthy et al. 2014]. Ces rendus en masquant les éléments lointains de l'environnement peuvent permettre de désencombrer le rendu brut et sans doute faciliter la tâche de navigation. De plus, utiliser la distance aux éléments pour déterminer le niveau de gris des phosphènes est une piste pour mettre en valeur la structure, comme illustré dans la Figure 14.

4 <http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-longrange.html>
5 <https://www.google.com/atap/project-tango/>

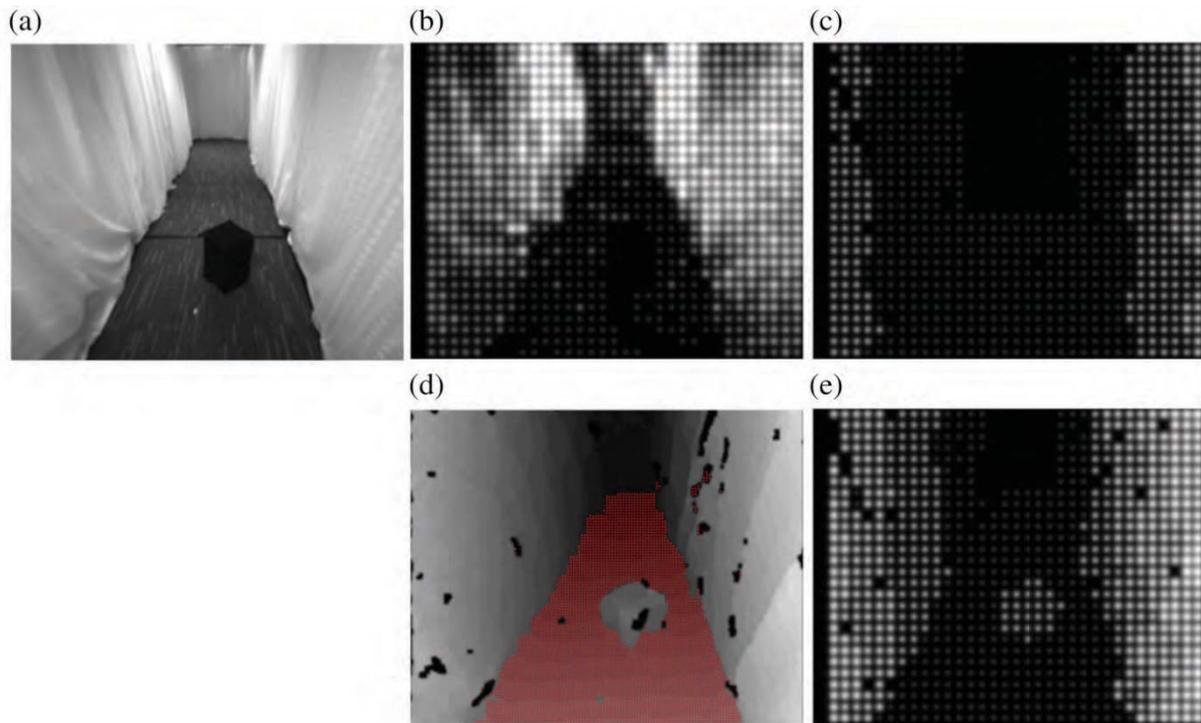


Figure 14 : Exemple de rendu basé sur la distance entre la caméra et les surfaces environnantes. (a) Un couloir contient un obstacle au sol. (b) Rendu brut classique (c) Rendu de distance (d) Le sol est détecté ce qui permet d'isoler l'obstacle qui ne lui appartient pas (e) Le rendu de distance est augmenté par l'isolation du sol et l'ajout de l'obstacle en surbrillance. Source : [McCarthy et al. 2014]

À notre connaissance, aucune étude en VPS n'a encore porté sur l'apport des rendus de distance dans une tâche de navigation et leur utilisabilité n'a pas été évaluée. Plus précisément, nous ne savons pas si ces rendus permettent d'obtenir de bonnes performances pour l'exploration d'un environnement inconnu ni si leur utilisation induit une charge cognitive plus importante ou réduite par rapport à un rendu brut classique. De même, nous ignorons si ces rendus de distance permettent la mémorisation de la disposition des environnements. Dans cette expérimentation nous avons évalué différents rendus de VPS utilisant la distance dans une tâche de navigation. Une pré-expérimentation décrite dans le paragraphe suivant a été menée pour arrêter certains choix dans le design de ces rendus. Un dernier rendu utilisant les possibilités offertes par la segmentation des surfaces [McCarthy et al. 2011] a également été évalué. De plus, pour cette expérimentation nous avons contrôlé si nos sujets étaient joueurs ou non-joueurs de façon à obtenir deux groupes équilibrés pour pouvoir comparer leurs performances. Trois variables dépendantes ont été utilisées : les deux premières analogues à celles de l'expérimentation précédente sont un indice de performance et la qualité des cartes dessinées par les sujets après la tâche de navigation. La dernière est la charge cognitive ressentie par les sujets, calculée à l'aide du Nasa-TLX. Les deux premières variables nous permettaient d'évaluer l'efficacité des rendus tandis que la dernière nous donnait un indice sur leur efficacité.

A. Matériel et méthode

A.1. Pré-expérimentation

Pour désencombrer le rendu brut de faible résolution grâce aux informations de distance, il est possible de masquer les éléments situés au-delà d'une certaine distance de la caméra. Les rendus obtenus de cette façon ont donc une profondeur de champ qui est limitée afin de réduire la quantité d'éléments à afficher. La distance limite utilisée ne peut être fixée arbitrairement : si cette distance est à l'infini le rendu obtenu est le rendu brut avec lequel il est difficile de naviguer. Si cette distance est trop petite, seuls les éléments à proximité immédiate peuvent être affichés ce qui limite également la navigation. Entre ces deux extrêmes, il existe sans doute une valeur (ou une plage de valeurs) dépendante des environnements utilisés qui doit être optimale pour la navigation. Pour nos environnements urbains, cette valeur doit permettre d'observer une rue dans sa largeur. Nous pensions donc que la valeur de cette distance limite devait être légèrement supérieure à la largeur des rues de nos environnements. Afin de sélectionner différentes valeurs de distance à évaluer dans cette expérimentation, nous avons mené une pré-expérimentation en utilisant deux sujets volontaires. Ces sujets ont navigué plusieurs fois avec un rendu prothétique dont la profondeur du champ de vision était limitée à différentes valeurs. De ces passations il ressortait que si la distance était inférieure à trois mètres il n'était plus possible de naviguer pour ces sujets. Les sujets parvenaient à naviguer sur une plage de distance allant de six mètres à neuf mètres. Au-delà de neuf mètres, ils étaient de nouveau en difficulté pour naviguer. Ainsi nous avons décidé d'évaluer ce rendu de VPS avec la valeur de distance fixée à trois mètres, six mètres et neuf mètres. Ces deux sujets n'ont pas participé à la suite de l'expérimentation. Grâce à cette pré-expérimentation nous avons donc décidé de trois valeurs distinctes pour évaluer l'apport de la profondeur de champ limité en VPS.

En vision non-prothétique (rendus Irrlicht) il est également possible de créer des rendus limitant la profondeur du champ. En revanche, il n'y a pas d'encombrement lorsque cette profondeur est à l'infini. Le fait de limiter de plus en plus la profondeur de navigation devrait gêner de plus en plus la navigation. La performance maximale de navigation devrait être obtenue pour la distance maximale (i. e. à l'infini) alors qu'une performance minimale devrait être obtenue pour la plus petite distance. Pour observer ces différentes évolutions entre les rendus avec profondeur de champ limitée en VPS et en Irrlicht nous avons décidé d'évaluer le rendu Irrlicht avec la profondeur de champ limitée aux mêmes valeurs que le rendu VPS qui sont trois mètres, six mètres et neuf mètres.

A.2. Rendus

Trois rendus VPS et trois rendus Irrlicht ont été fixés par la pré-expérimentation. À ces six rendus sont venus s'ajouter les rendus classiques en Irrlicht et en VPS pour lesquels la profondeur n'est pas limitée. Enfin deux autres rendus VPS ont été ajoutés : le premier utilise la distance aux éléments pour les représenter; le dernier affiche la structure de l'environnement

en se basant sur les surfaces détectées. Au final, six rendus de VPS ont été évalués dans cette expérimentation et quatre rendus Irrlicht complétaient ces conditions. Tous les rendus recouvraient un espace visuel de 24x17 degrés, ce qui permet de se rapprocher de la taille visuelle théorique prise par le futur implant Argus III [Stronks & Dagnelie 2014].

Pour faciliter la compréhension, les rendus Irrlicht sont d'abord décrits avant de décrire les rendus de VPS avec le champ de profondeur limité. Le rendu VPS affichant la distance et celui utilisant la segmentation des surfaces sont décrits en derniers.

Pour limiter la profondeur de champ, nous avons défini la distance maximale de vision du rendu Irrlicht avec des outils fournis par le moteur 3D. Ainsi un masque noir est placé à une certaine distance de la caméra virtuelle et empêche de voir les objets situés au-delà du masque. Trois distances différentes ont été évaluées : le masque a été placé à trois mètres, six mètres et neuf mètres. Pour faciliter la lecture, ces trois rendus ont été dénommés **IRR-Lim3**, **IRR-Lim6** et **IRR-Lim9**. Le rendu sans limitation de la profondeur (infini) correspondant au rendu classique Irrlicht a été appelé **IRR-LimInf**. Ces trois distances ont été décidées d'après la pré-expérimentation décrite plus haut.

Pour tous les rendus de vision prothétique, les mêmes caractéristiques sont utilisées pour générer les matrices de phosphènes. Les phosphènes sont des ronds dont le profil de luminance est de type gaussien. Afin de se rapprocher des retours d'expérience des sujets implantés [Chen et al. 2009], seulement quatre niveaux de gris distincts sont utilisés : blanc, gris clair, gris foncé et sombre. De même, la position des phosphènes était altérée de dix pour cent et un dysfonctionnement des électrodes était simulé (phénomène de *dropout*) : durant toute l'expérience, les sujets avaient dix pour cent des phosphènes qui étaient absents. De plus, le phénomène d'adaptation des cellules rétiniennes à la stimulation électrique a été simulé dans cette expérience [Pérez Fornos et al. 2012]. Si un même niveau de gris était attribué à un phosphène pendant une période excédant une demi-seconde, le phosphène en question s'assombrissait complètement pour ne réapparaître qu'après un laps de temps de quelques dixièmes de seconde. Ce phénomène avait comme conséquence visible de produire un certain scintillement lors de l'utilisation des rendus VPS.

Les trois rendus prothétiques avec le champ de profondeur limités ont été générés à partir des trois rendus Irrlicht correspondants. Pour cela l'image a été divisée en deux cent soixante-dix zones et pour chacune d'entre elles la moyenne de la luminance a été utilisée pour attribuer le niveau de gris le plus proche. De cette même manière, le rendu **VPS-LimInf** a été généré à partir du rendu **IRR-LimInf**. Les quatre rendus générés de cette manière sont visibles dans la Figure 16 et ont été dénommés **VPS-Lim3**, **VPS-Lim6** et **VPS-Lim9**.

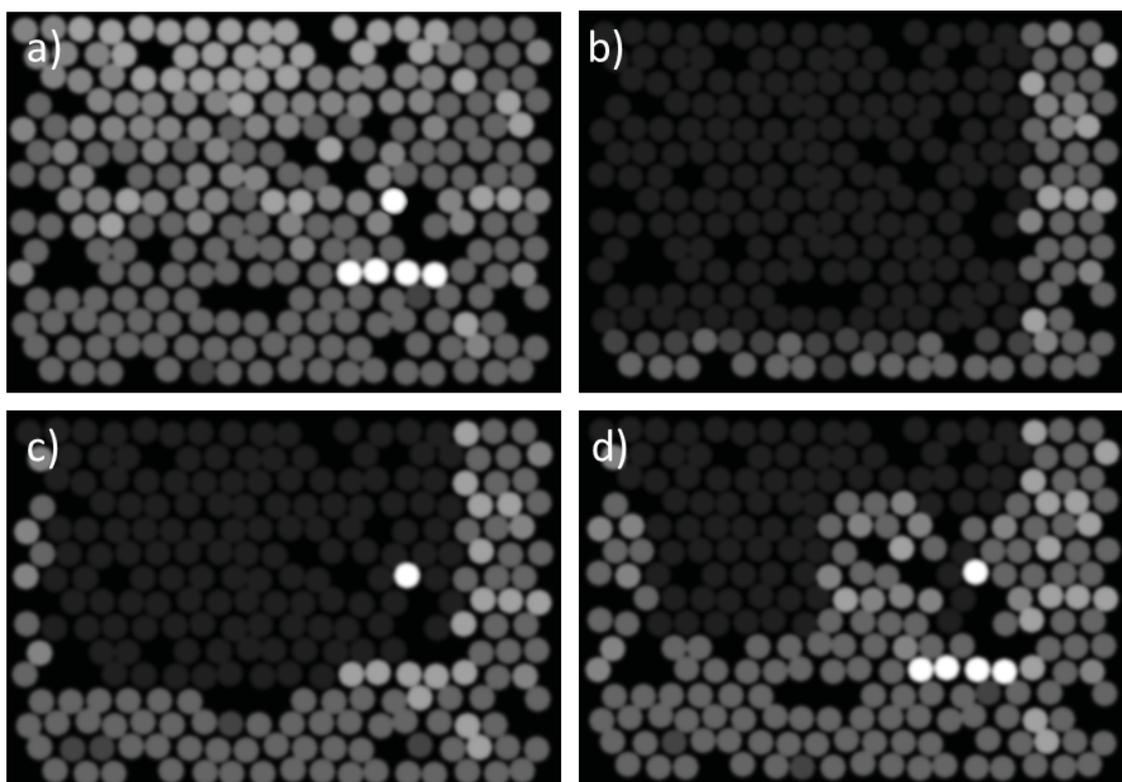


Figure 16 : Rendus prothétiques du même point de vue que Figure 15. a) Rendu VPS-LimInf. b) Rendu VPS-Lim3. c) Rendu VPS-Lim6. d) Rendu VPS-Lim9. Ces rendus sont générés en temps réel à partir des rendus de vision Irrlicht présentés dans la Figure 15.

En plus de ces quatre rendus, deux autres rendus de VPS ont été évalués. Le premier rendu utilise la distance à chaque élément de la scène visuelle. Le niveau de gris de chaque phosphène est déterminé en fonction de la distance moyenne d'une zone donnée dans la scène visuelle. Ce type de rendu a déjà été mis au point et évalué en tâche de mobilité [Tatur 2011; Lieby et al. 2011]. Deux niveaux de gris étaient utilisés pour indiquer respectivement les éléments situés à moins de trois mètres de la caméra, puis les éléments entre trois mètres et neuf mètres de distance. Les phosphènes correspondant à des éléments situés au-delà de neuf mètres étaient affichés avec la luminance la plus sombre. Ce rendu dénommé VPS-Distance est visible dans la Figure 17 c).

Le dernier rendu dénommé VPS-FilDeFer faisait percevoir la structure des environnements sans utiliser la distance. Au lieu d'augmenter le contraste entre les murs et le sol, ce sont les bordures entre les différents plans de la scène visuelle qui étaient augmentées. Ainsi, ce rendu améliorait la perception des intersections entre le sol et les murs, les coins présents dans les intersections ainsi que les bordures hautes des murs. Une vision en « fil-de-fer » était ainsi générée et utilisée en vision prothétique. Seuls trois niveaux de gris étaient nécessaires pour ce rendu : les joyaux et la base de départ étaient indiqués en blanc, les arêtes étaient en gris foncé et toutes les surfaces de l'environnement étaient en gris sombre. Un aperçu de ce rendu est visible dans la Figure 17 d).

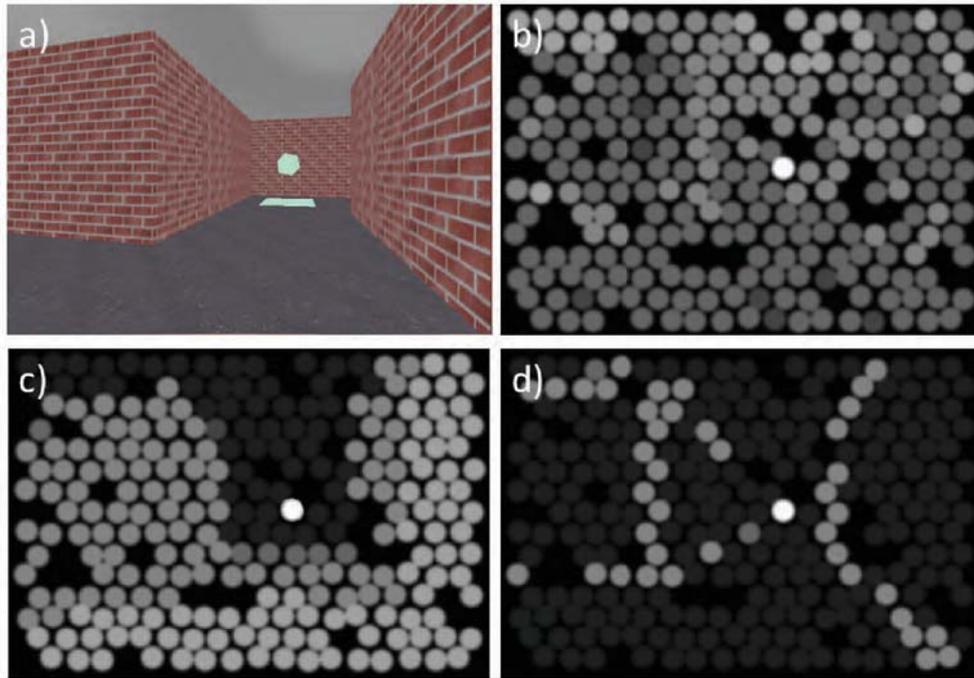


Figure 17 : Rendus prothétiques augmentant la structure de l'environnement. a) Rendu Irrlicht original (IRR-LimInf). b) Rendu brut classique (VPS-LimInf). c) Rendu prothétique basé sur la distance existant entre les éléments et la caméra (VPS-Distance). d) Rendu VPS-FilDeFer.

A.3. Hypothèses

Une première partie des hypothèses porte sur les rendus Irrlicht, la seconde partie des hypothèses porte sur les rendus VPS. Les trois mesures servant de variables dépendantes sont l'indice de performance, les scores des dessins effectués par les sujets ainsi que le Nasa-TLX. D'autres hypothèses, secondaires, portent sur d'autres facteurs pouvant avoir un impact sur les résultats.

A.3.a. Rendus Irrlicht

Pour les quatre rendus Irrlicht, le fait de limiter la profondeur du champ de vision fait disparaître de l'information. Ainsi nous avons supposé que plus la distance de vision est importante, plus les sujets auront de bonnes performances dans la tâche de navigation et plus ils parviendront à produire des dessins de carte précis. Nous avons également fait l'hypothèse que plus la limite de profondeur était éloignée et plus le score du Nasa-TLX devait être faible.

A.3.b. Rendus VPS

Quatre des rendus en vision prothétique limitaient la distance de vision comme les rendus en vision Irrlicht. En raison de l'encombrement du rendu VPS, nous avons fait l'hypothèse que la limite de profondeur allait aider les sujets à percevoir la structure de l'environnement. Cependant, nous avons fait l'hypothèse que si cette distance limite était trop petite, les sujets n'ont plus assez d'information sur l'environnement ce qui limite leur performance et leurs représentations mentales de l'environnement. À l'inverse, si cette limite est trop éloignée le

rendu est trop encombré ce qui engendre des confusions et limite les performances des sujets. Ainsi nous avons fait l'hypothèse que les performances de navigation avec le rendu VPS-Lim6 et VPS-Lim9 allaient être supérieures à celle avec les rendus VPS-Lim3 et VPS-LimInf.

Nous avons fait l'hypothèse que les sujets allaient obtenir de meilleures performances de navigation avec les rendus Distance et Fil-De-Fer qu'avec le rendu VPS-LimInf.

Enfin nous faisons également l'hypothèse d'obtenir des résultats équivalents sur les scores attribués aux dessins ainsi que sur la charge cognitive ressentie par les sujets. S'ils sont plus utilisables, les rendus permettant de naviguer efficacement devaient également permettre de mémoriser correctement la disposition des environnements pour une charge cognitive ressentie plus légère.

A.3.c. Autres facteurs

L'expérience précédente nous a amenés à nous poser des questions sur l'impact de l'utilisation de jeux vidéo par les sujets. Comme dans l'expérience précédente, les sujets devaient interagir avec les flèches directionnelles d'un clavier pour se déplacer dans l'environnement. Les contrôles utilisés par les sujets pour naviguer dans les environnements virtuels étaient donc semblables à ceux utilisés dans le domaine du jeu vidéo. Nous avons fait l'hypothèse que le niveau d'expertise dans l'utilisation des jeux vidéo allait avoir un impact sur les performances des sujets [Larrue 2011]. Notre hypothèse était que les sujets joueurs allaient obtenir de meilleures performances que les autres sujets.

La tâche étudiée était une tâche de navigation avec une importante composante d'orientation. Ainsi, nous faisons l'hypothèse que les résultats des sujets allaient dépendre de leurs compétences propres en orientation mesurées par le questionnaire de Santa-Barbara : *Santa Barbara Sense of Direction Scale* [Hegarty et al. 2002].

A.4. Sujets

Quatorze sujets (cinq femmes et neuf hommes ayant entre 18 et 31 ans; M : 23,7 DS : 3,5) disposant d'une vue normale ou corrigée à la normale ont participé à cette expérimentation. Le comité d'éthique CLERIT de l'université de Toulouse a validé ce protocole et tous les sujets ont signé un formulaire de consentement éclairé. Les sujets ont été recrutés de telle manière à obtenir autant de sujets se qualifiant de « joueurs » que de sujets se qualifiant de « non-joueurs ».

A.5. Matériel

Le matériel utilisé était le même que pour les expériences précédentes. Il comportait un visiocasque NVisor SX-60 (NVIS Inc., Reston, États-Unis) avec une résolution de 1280x1024 s'étalant sur un champ de 44x30 degrés. L'orientation de la caméra virtuelle était asservie à l'orientation du casque grâce à un système de capture de mouvements (OptiTrack, de Natural Point). Les calculs et les rendus des environnements virtuels étaient faits par le moteur 3D Irrlicht. Pour se déplacer dans l'environnement virtuel, les sujets utilisaient les flèches directionnelles d'un clavier classique. La flèche avant permettait d'avancer. Les sujets étaient assis sur

une chaise et pouvaient tourner la tête, ce qui faisait tourner la caméra dans l'environnement virtuel. Ils pouvaient également tourner dans l'environnement en utilisant les flèches directionnelles gauche et droite.

Les agents IVY utilisés sont des évolutions des agents de l'expérimentation précédente et sont donc les suivants :

- ◆ Agent de Motion Capture (C++)
- ◆ Programme de gestion de la tâche utilisant le moteur 3D Irrlicht (C++, librairie Irrlicht)
- ◆ Simulateur de Vision Prothétique (C++, OpenGL)
- ◆ Interface de Contrôle pour l'expérimentateur (JAVA)

Deux ordinateurs étaient utilisés pour faire fonctionner les différents modules. Le premier contenait l'interface expérimentatrice qui a été développée en Java et qui permettait de piloter le moteur 3D Irrlicht ainsi que d'enregistrer toutes les données de l'expérience. Une deuxième machine contenait l'exécutable écrit en C++ utilisant la librairie du moteur 3D Irrlicht avec lequel interagissaient directement les sujets. Ce deuxième ordinateur contenait aussi le simulateur de vision prothétique développé en C++ et OpenCV, qui recevait l'information du moteur Irrlicht pour afficher les phosphènes. Un troisième module s'exécutait sur cette même machine, il s'agissait du logiciel de capture de mouvement qui permettait de localiser et orienter le casque dans l'espace. Cet exécutable utilisait l'API fournie par le système OptiTrack et a été développé en C++. Toutes les communications entre les exécutables se faisaient via le bus open source IVY⁶.

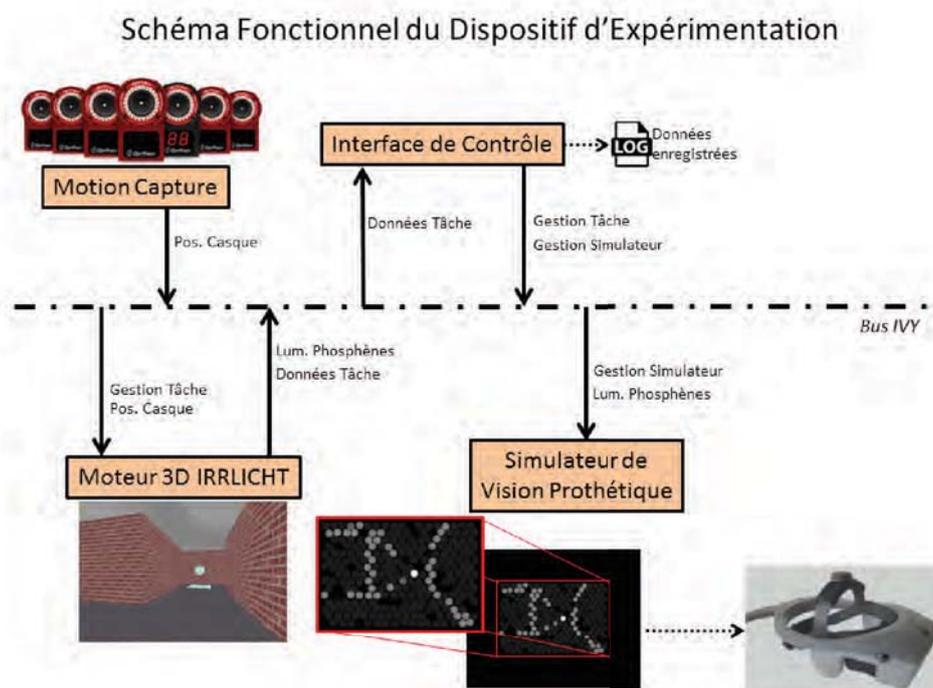


Figure 18 : Schéma fonctionnel des agents utilisés pour cette troisième expérimentation. Même si certains agents ont été recodés en grande partie, les interactions entre agents sont similaires entre cette expérimentation et la précédente.

6 <http://www.eei.cena.fr/products/ivy/> (Consulté le 21/05/2015)

A.6. Environnements

Pour cette expérimentation douze environnements différents ont été créés. C'étaient des environnements urbains simples et constitués de bâtiments séparés par des rues. Toutes les intersections étaient constituées d'angles à 90°. Deux types d'intersections étaient présents dans les environnements : intersection en croix et intersection en « T ». Les rues faisaient cinq mètres de larges, les murs des bâtiments cinq mètres de haut. Chaque environnement était contenu dans une surface au sol de 45x45 mètres carrés. Les environnements contenaient un point de départ (la base) qui était fixe et avait la forme d'un disque jaune très clair dessiné sur le sol. La base de départ était le seul élément pérenne de l'environnement ainsi les seuls points de repère utilisables par les sujets pour se repérer étaient cette base et les intersections. Des bijoux étaient également contenus dans l'environnement, ils étaient matérialisés par un cube bleu très clair flottant dans l'air ainsi que par une base carrée en contact avec le sol. Lors de leur récolte les bijoux disparaissaient ainsi que leur base. Pour limiter l'effet de l'utilisation des différents environnements pour naviguer, le nombre d'intersections de chaque type était très voisin entre les différents environnements (intersection en croix : M : 1; DS : 0; intersection en « T » : M : 5,8; DS : 0,7). Tous les environnements ont été créés en utilisant les logiciels Google Sketchup et CopperCube.

A.7. Tâche et variables observées

Pour chaque condition les sujets effectuaient la tâche de navigation suivante : en premier lieu, les sujets devaient récolter trois bijoux dispersés dans un petit environnement et les ramener à la base de départ. Cet environnement de familiarisation était le même pour toutes les conditions. Lors de cet apprentissage, des explications étaient données aux sujets sur la nature du rendu utilisé. Une aide était également fournie pour expliquer aux sujets les éléments de l'environnement (bijoux, base de départ) ainsi que sa structure (intersections, impasse). Les sujets pouvaient profiter de ce temps-là pour poser des questions et se familiariser avec les interactions. Cette tâche d'apprentissage pouvait être effectuée une deuxième fois à la demande des sujets. Une fois que les sujets se sentaient prêts, la tâche de navigation commençait.

Les sujets disposaient de cinq minutes pour récolter six bijoux dispersés dans un nouvel environnement. Pour récolter un bijou, les sujets devaient le retrouver, le prendre (en passant dessus) puis le déposer à la base de départ. Seuls deux bijoux pouvaient être portés simultanément, ainsi les sujets devaient effectuer au minimum trois allers et retours pour tous les récolter. Lors de leur déplacement, les sujets entendaient des bruits de pas et un son mat était produit lorsqu'ils heurtaient un mur. Des sons spécifiques étaient également joués lors de la récolte d'un bijou ou lors du dépôt de bijoux à la base de départ.

Si les sujets parvenaient à récolter tous les bijoux avant le temps imparti (5 min), la tâche ne s'arrêtait pas pour autant et les sujets utilisaient le temps restant pour mémoriser l'environnement. Pour chaque sujet, chaque rendu était évalué dans un environnement virtuel différent. Pour chaque rendu, les sujets effectuaient deux fois la tâche de navigation dans un même

environnement. Deux objectifs étaient donnés aux sujets, un objectif principal et un objectif secondaire. Le premier objectif était d'effectuer la tâche le plus rapidement possible (récolter les bijoux en un minimum de temps). Le deuxième objectif était de mémoriser la carte de l'environnement. Ainsi il était explicitement demandé aux sujets de se concentrer d'abord sur la récolte des bijoux puis d'utiliser le temps restant pour mémoriser la disposition de l'environnement. Ce fonctionnement permettait à tous les sujets de bénéficier du même temps d'exposition aux environnements.

Une fois que les deux essais étaient passés, les sujets étaient invités à dessiner une carte de l'environnement. La carte de l'environnement devait contenir la base de départ ainsi que la position des différents bijoux.

Pour évaluer la performance des sujets, un indice de performance (PI) a été calculé à la fin de l'expérimentation. Comme pour l'expérience précédente, cet indice était calculé à partir du nombre de bijoux récoltés et du temps mis pour les récolter selon la formule suivante :

$$PI = \frac{J_i - J_{MIN}}{J_{MAX} - J_{MIN}} * 50 + \frac{T_{MAX} - T_i}{T_{MAX} - T_{MIN}} * 50$$

Avec :

J_{MIN} le nombre minimal possible de bijoux collectés (ici $J_{MIN} = 0$),

J_{MAX} le nombre maximal possible de bijoux collectés (ici $J_{MAX} = 6$),

J_i le nombre de bijoux collectés par le sujet,

T_{MIN} le temps minimal enregistré pour accomplir l'expérience, tous sujets et toutes conditions confondus,

T_{MAX} le temps maximal possible (ici $T_{MAX} = 5 \text{ min}$),

T_i le temps mis par les sujets pour accomplir la tâche.

Pour évaluer les dessins des sujets, quatre juges externes à l'expérience et naïfs sur les conditions et hypothèses de l'expérimentation ont noté les dessins produits par les sujets. Chacun des juges a ainsi procédé à l'évaluation de 140 dessins. Les évaluations se sont faites sur une échelle allant de zéro à dix. Les juges avaient comme condition de noter la proximité entre les dessins produits par les sujets et les cartes des environnements. Il a été demandé aux juges de ne pas tenir compte des caractéristiques euclidiennes des dessins mais uniquement de leur topologie.

La charge cognitive ressentie pour naviguer avec chacun des rendus a également été évaluée chez les sujets. Pour ce faire, nous avons utilisé une traduction française du questionnaire Nasa-TLX [Hart & Staveland 1988; Cegarra & Morgado 2009]. Ce questionnaire permettait d'évaluer la charge cognitive d'une tâche selon six dimensions différentes : exigence physique, exigence mentale, exigence temporelle, effort, performance, frustration. Les sujets évaluaient chacune de ces dimensions indépendamment les unes des autres puis devaient les interclasser deux par deux sur l'importance attribuée à chacune. La moyenne pondérée de chacune de ces dimensions donnait le score du Nasa-TLX.

Les sujets devaient aussi remplir deux questionnaires en fin d'expérimentation. Le premier se concentrait sur leur pratique du jeu vidéo et la familiarité vis-à-vis de la navigation dans des environnements virtuels. Le deuxième questionnaire était une version française du questionnaire d'évaluation du sens de l'orientation de Santa Barbara [Hegarty et al. 2002]. Ce questionnaire utilisait des questions sur la vie courante liée à la navigation pour que les sujets puissent évaluer leur sens de l'orientation.

B. Résultats

B.1. Indice de performance

En vision Irrlicht, plus la limite de vision est éloignée, plus l'indice de performance des sujets est important. Des différences significatives sont présentes entre le rendu IRR-Lim3 ($M=59,5$; $DS=13,04$) et les autres rendus : IRR-Lim6 ($M=74,73$; $DS=8,61$; $p < 0,01$), IRR-Lim9 ($M=76,27$; $DS=7,78$; $p < 0,001$) et IRR-LimInf ($M=80,53$; $DS=4,53$; $p < 0,001$). Cela signifie que plus la limite de profondeur est éloignée, plus les sujets sont parvenus à collecter les joyaux rapidement.

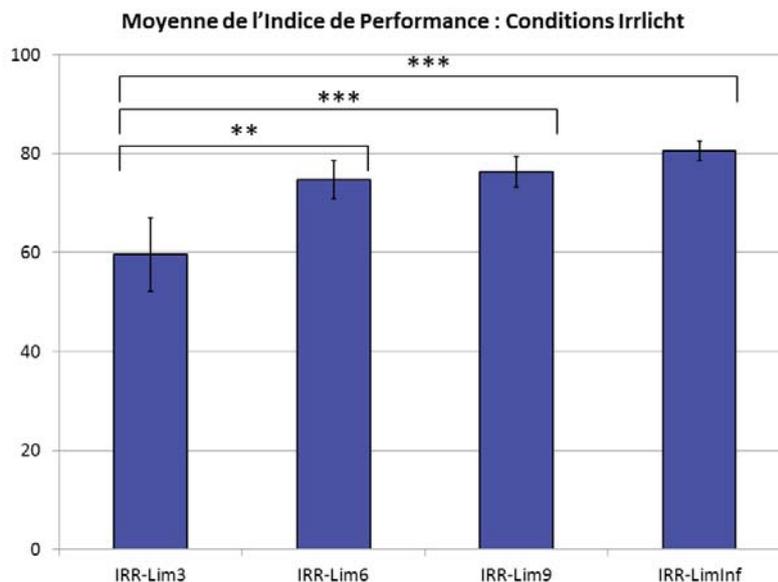


Figure 19 : Moyennes de l'Indice de Performance des conditions de vision Irrlicht. Les différences significatives sont indiquées avec * pour $p < 0,05$; ** pour $p < 0,01$ et *** pour $p < 0,001$. Les barres d'erreurs sont les intervalles de confiance à 95 %

En vision prothétique, la plupart des sujets ne sont pas parvenus à ramener les joyaux dans les temps pour les conditions VPS-LimInf ($M=37,29$; $DS=17,20$) et VPS-Lim3 ($M=38,36$; $DS=18,36$). L'indice de performance est au-delà de 50 pour les rendus VPS-Lim6 ($M=63,61$, $DS=24,01$) et VPS-Lim9 ($M=50,51$; $DS=22,01$). Avec le rendu VPS-Distance ($M=57,20$; $DS=22,29$) et VPS-FilDeFer ($M=65,78$; $DS=19,45$), la moyenne de l'indice de performance est au-dessus de 50. Des différences significatives sont présentes entre la condition VPS-Lim3 et les conditions VPS-Lim6 ($p < 0,001$), VPS-Distance ($p < 0,001$) et VPS-FilDeFer ($p < 0,001$). La différence de moyenne entre la condition VPS-Lim6 et VPS-LimInf est significative également ($p < 0,01$). La moyenne de l'indice de performance pour la condition VPS-Lim9 est statistiquement dif-

férente de celle de la condition VPS-FilDeFer ($p < 0,05$). Enfin la condition VPS-LimInf se démarque significativement des conditions VPS-Distance ($p < 0,01$) et VPS-FilDeFer ($p < 0,001$).

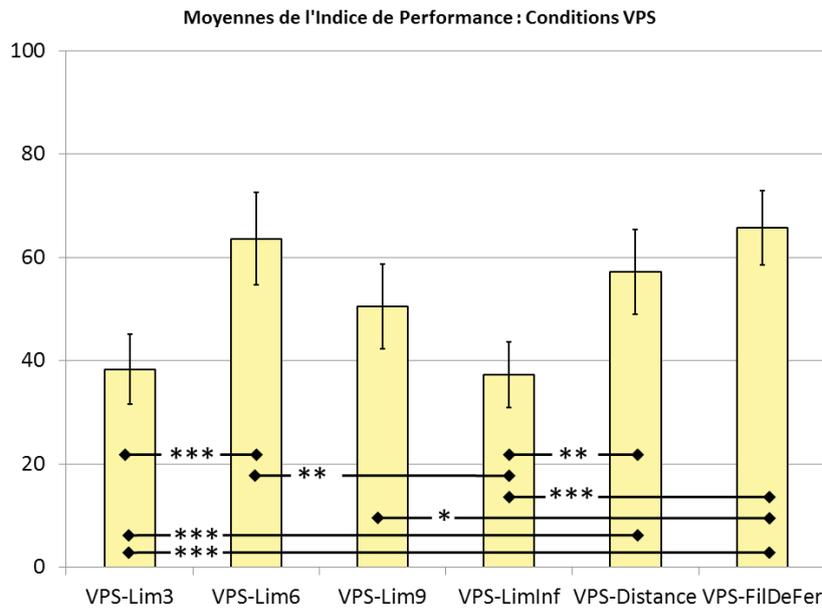


Figure 20 : Moyennes de l'Indice de Performance sur les conditions de vision prothétique. Les principaux résultats statistiques sont indiqués : ** pour $p < 0,01$; *** pour $p < 0,001$. Les barres d'erreurs sont les intervalles de confiance à 95 %.

B.2. Charge cognitive

Parmi les conditions Irrlicht, le rendu ayant la charge cognitive la plus élevée est le IRR-Lim3 ($M=46,06$; $DS=15,01$). Comme pour l'indice de performance, des différences significatives sont présentes entre cette condition et les rendus IRR-Lim6 ($M=27,00$; $DS=10,97$; $p < 0,001$), IRR-Lim9 ($M=25,42$; $DS=12,06$; $p < 0,001$) et IRR-LimInf ($M=23,69$; $DS=11,93$; $p < 0,001$). Ces résultats sont visibles dans la Figure 21.

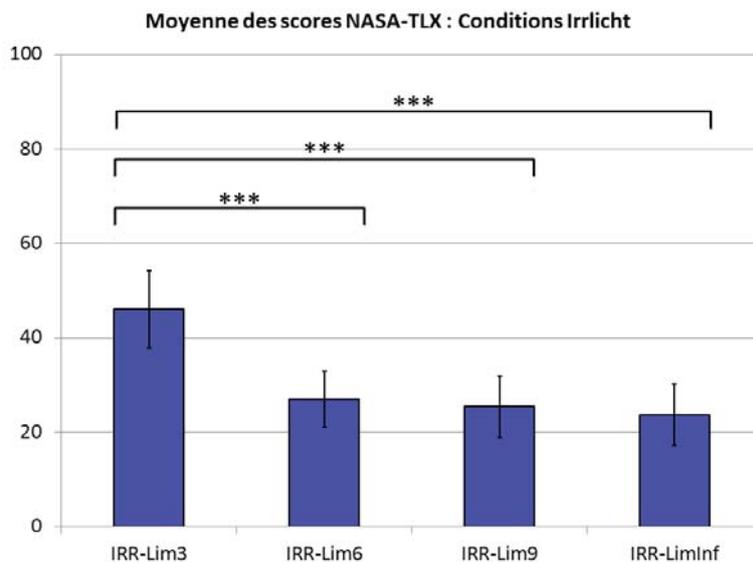


Figure 21 : Moyennes du Nasa-TLX pour les conditions de rendus Irrlicht. Les barres d'erreurs sont les intervalles de confiance à 95 %.

En vision prothétique, la condition pour laquelle les sujets ont ressenti la plus importante charge cognitive est VPS-LimInf (M=74,31; DS=23,69), ce qui correspond au rendu que les personnes implantées utilisent actuellement. La deuxième condition VPS en terme de charge cognitive est VPS-Lim3 (M=61,61; DS=14,51). Le score le plus faible a été obtenu pour la condition VPS-FilDeFer (M=38,59; DS=16,23). Les rendus VPS-Lim6 (M=49,22; DS=13,70), VPS-Lim9 (M=53,43; DS=12,53) ainsi que le rendu VPS-Distance (M=50,01; DS=16,52) présentaient des scores très proches les uns des autres. Des différences significatives apparaissaient entre la condition VPS-Lim3 et les conditions VPS-Distance ($p < 0,05$) et VPS-FilDeFer ($p < 0,001$). De même la condition VPS-LimInf se démarque statistiquement des conditions VPS-Lim6 ($p < 0,01$), VPS-Lim9 ($p < 0,05$), VPS-Distance ($p < 0,001$) et VPS-FilDeFer ($p < 0,001$) Ces résultats sont résumés dans la Figure 22.

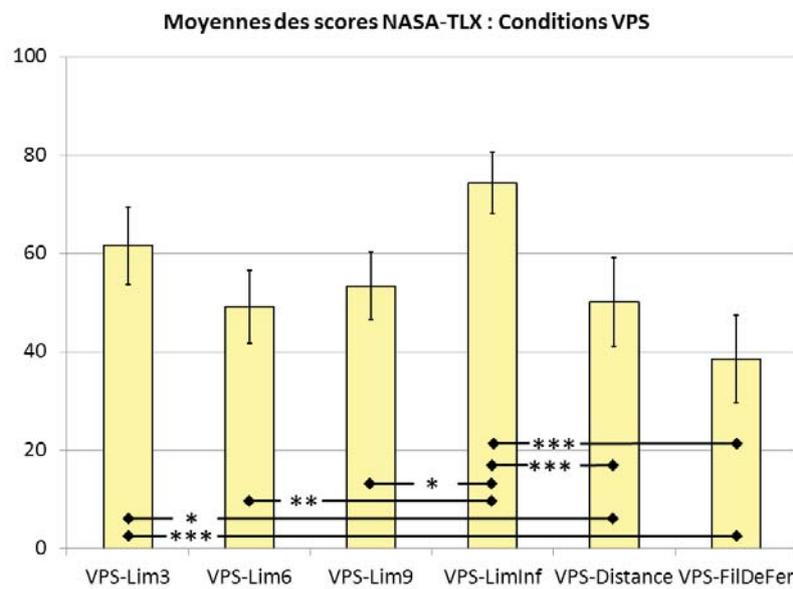


Figure 22 : Moyennes du Nasa-TLX pour les conditions de vision prothétique.
Les barres d'erreur sont les intervalles de confiance à 95 %.

B.3. Évaluation des dessins

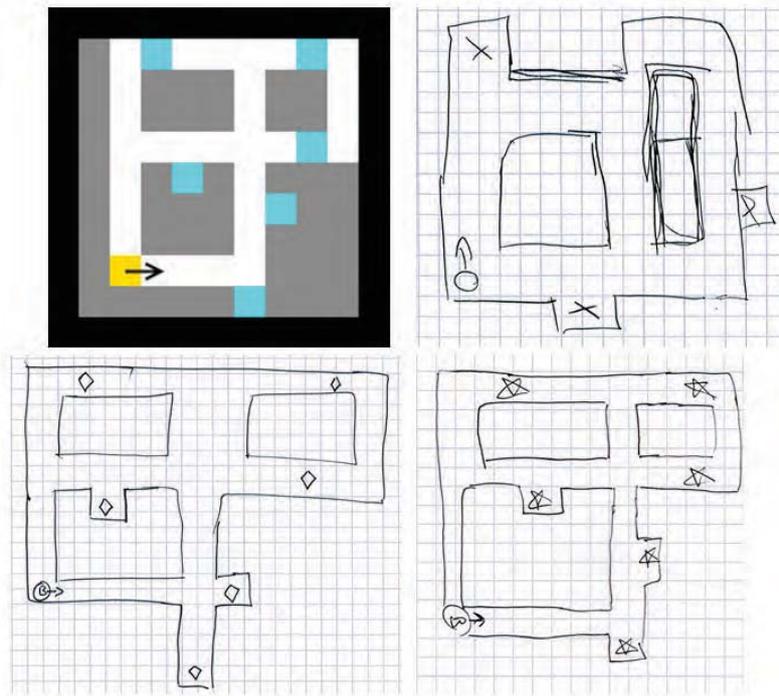


Figure 23 : Exemples de cartes dessinées par les sujets. La carte de l'environnement est représentée en haut à gauche. Les trois autres images sont des représentations produites par des sujets après exploration. Les juges disposaient de la carte de départ pour évaluer tous les dessins.

Pour les rendus Irrlicht plus la limite de profondeur est éloignée de la caméra, plus la note moyenne obtenue par les dessins est grande. IRR-Lim3 ($M=4,27$; $DS=2,39$), IRR-Lim6 ($M=6,80$; $DS=2,56$) IRR-Lim9 ($M=7,57$; $DS=2,62$) ou IRR-LimInf ($M=8,36$; $DS=1,46$). Des différences statistiques sont présentes entre IRR-Lim3 et IRR-Lim9 ($p < 0,001$) ainsi qu'entre IRR-Lim3 et IRR-LimInf ($p < 0,001$).

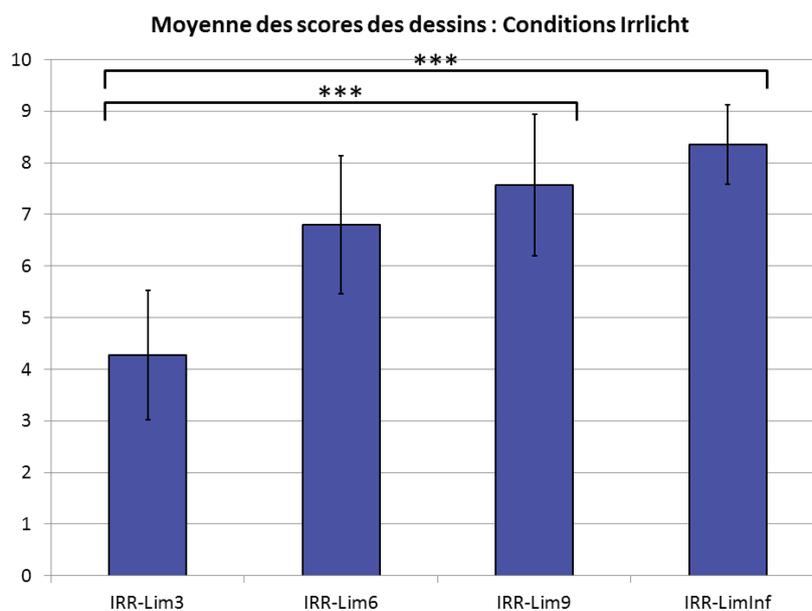


Figure 24 : Moyenne des notes attribuées par les juges sur les dessins pour chacune des conditions Irrlicht. Les barres d'erreurs sont les intervalles de confiance à 95 %.

La condition en vision prothétique qui a permis aux sujets de produire les dessins les mieux évalués est la condition VPS-FilDeFer (M=4,64; DS=3,35) qui se démarque significativement des conditions VPS-Lim3 (M=2,20; DS=2,69; $p < 0,05$) et VPS-LimInf (M=1,27; DS=1,58; $p < 0,01$). La deuxième meilleure est la condition VPS-Distance (M=3,77; DS= 3,58) qui est significativement différente de la condition VPS-LinInf ($p < 0,05$). Les dessins présentant les évaluations les plus basses ont été produits en condition VPS-LimInf (M=1,27; DS=1,58). Les rendus VPS-Lim6 (M=2,52; DS=1,97) et VPS-Lim9 (M=2,79; DS=2,70) ne présentent pas de différence significative avec les autres rendus.

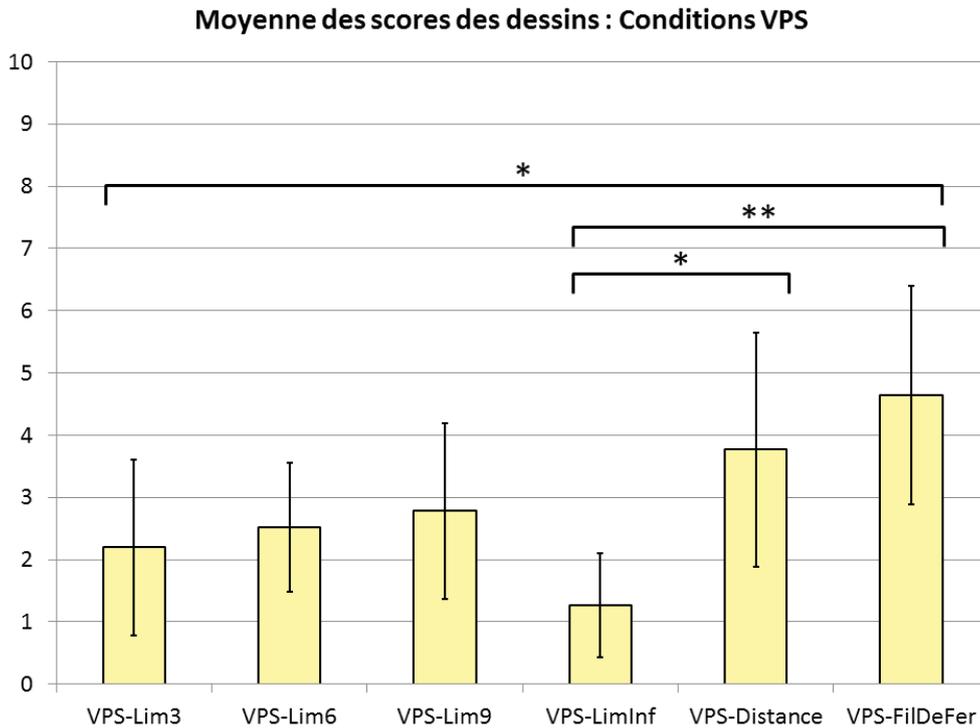


Figure 25 : Moyenne des scores des dessins pour chaque Conditions VPS. Les barres d'erreurs sont les intervalles de confiance à 95 %.

B.4. Sens de l'orientation (score au questionnaire de Santa Barbara)

Avant d'effectuer la tâche, les sujets ont rempli le questionnaire d'évaluation du sens de l'orientation de Santa Barbara [Hegarty et al. 2002]. Tous les sujets ont répondu à ce questionnaire et ont obtenu un score situé sur une échelle entre 1 et 7. Le score moyen obtenu est de 4,7 avec un écart type de 1,6. Pour savoir si ce score correspondant au sens de l'orientation est en corrélation avec les performances des sujets, une régression linéaire portant sur l'ensemble des données a été faite. Il s'avère que le score obtenu au questionnaire de Santa-Barbara des sujets ne corrèle pas avec les scores des dessins ($p > 0,6$; $r = -0,06$; $r^2 = 0,01$). Ce score corrèle également très mal avec l'Indice de Performance ($p > 0,3$; $r = 0,02$; $r^2 = 0,02$) et le Nasa-TLX ($p > 0,8$; $r = 0,04$; $r^2 = 0,11$). La Figure 26 présente les correspondances entre le score de Santa Barbara et ces trois variables dépendantes.

Corrélations du score de Santa Barbara

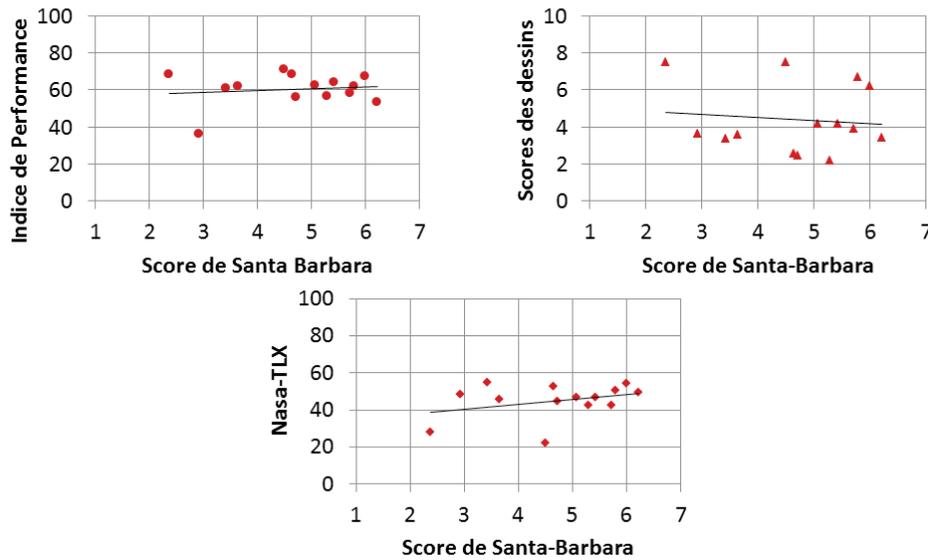


Figure 26 : Le score de Santa Barbara des sujets est affiché en relation avec les trois variables dépendantes de l'expérimentation. Pour chaque graphique, chaque élément correspond à un sujet, en abscisse se trouve l'échelle du questionnaire de Santa-Barbara et en ordonnée se trouve a) l'Indice de Performance b) les scores des dessins c) le Nasa-TLX.

B.5. Influence de la pratique des jeux vidéo et de la connaissance de la réalité virtuelle

Parmi les sujets, deux se considéraient comme des non-joueurs, cinq comme des joueurs occasionnels, quatre comme des joueurs réguliers et trois comme des joueurs passionnés. Ces réponses sont cohérentes avec leurs fréquences d'utilisation des jeux vidéo. Les joueurs passionnés jouent au moins une heure par jour, les joueurs réguliers au moins plusieurs fois par semaine. Les joueurs occasionnels jouent une à plusieurs fois par semaines, les non-joueurs jouaient moins d'une fois par semaine. À partir de ces réponses, les sujets ont été séparés en deux groupes, un groupe contenant les sujets non joueurs et joueurs occasionnels (groupe non-joueur) et un groupe contenant les joueurs réguliers et gros joueurs (groupe joueur). Les résultats de ces deux populations contenant chacune sept sujets ont pu ensuite être comparés en utilisant le test de Mann-Whitney (ou test des rangs de Wilcoxon). L'impact de la pratique des jeux vidéo sur l'Indice de Performance est perceptible pour trois des conditions Irrlicht que sont IRR-Lim3 ($p < 0,05$), IRR-Lim6 ($p < 0,001$) et IRR-Lim9 ($p < 0,001$). La différence des moyennes des joueurs et non-joueurs pour la condition IRR-LimInf est également très proche de la limite de signification ($0,06 < p < 0,05$). Pour les conditions de VPS, des différences significatives sont présentes pour les conditions VPS-Distance ($p < 0,01$) et VPS-FilDeFer ($p < 0,05$). Ces résultats sont visibles dans la Figure 27.

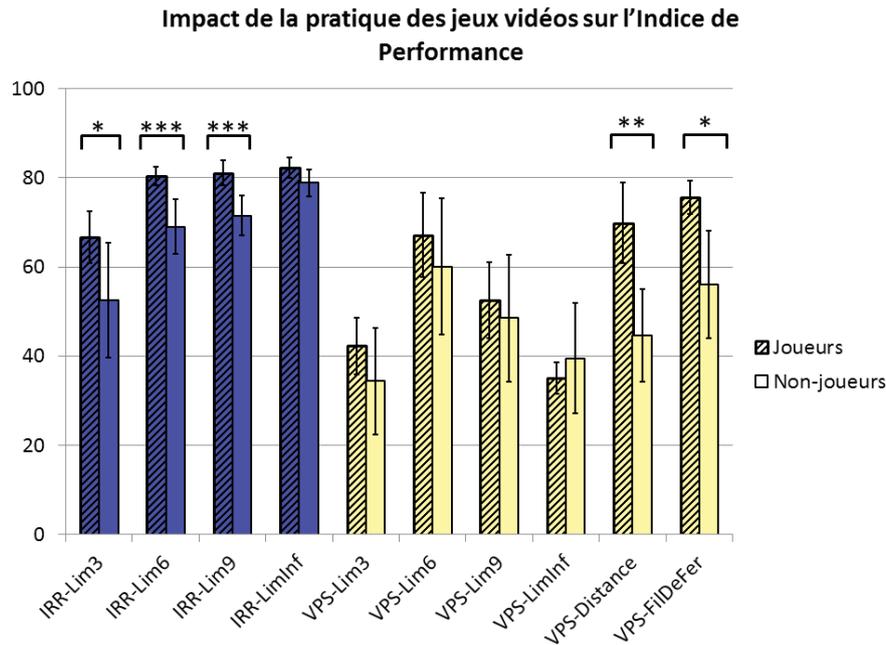


Figure 27 : Comparaison des moyennes des indices de performances entre le groupe des non-joueurs et le groupe des joueurs. En bleu sont représentées les conditions de rendus Irrlicht. En jaune sont les conditions en VPS. Les barres d'erreur sont les intervalles de confiance à 95 %.

Pour le score du Nasa-TLX, aucune des différences entre les moyennes obtenues par les groupes de joueurs et de non-joueurs n'est significative. À titre indicatif, ces résultats sont présentés dans la Figure 28.

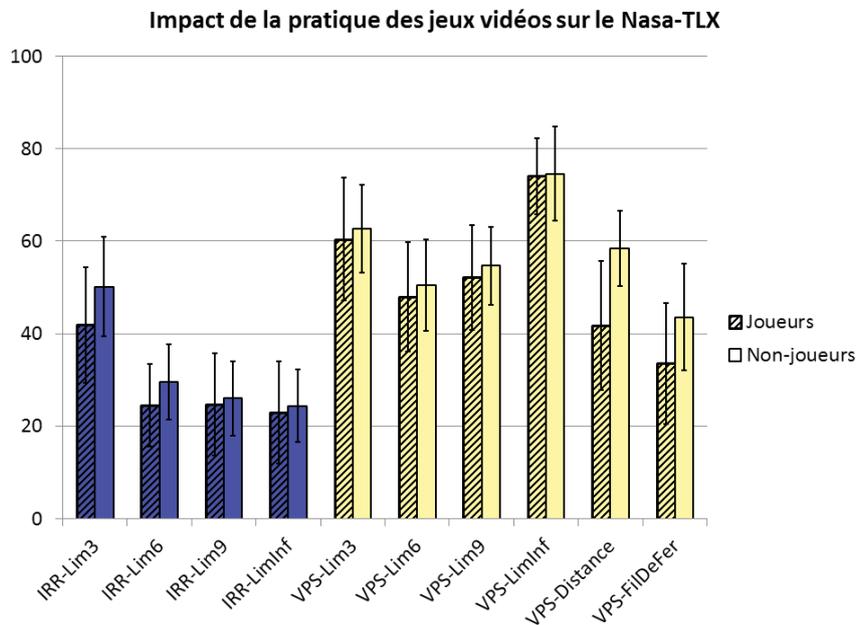


Figure 28 : Moyennes du Nasa-TLX des groupes de joueurs et non-joueurs. Les conditions Irrlicht sont représentées en bleu, les conditions VPS le sont en jaune.

Pour les scores obtenus par les dessins, aucune différence n'est significative en ce qui concerne les conditions Irrlicht. En revanche des différences apparaissent pour les conditions

VPS-Distance ($p < 0,05$), VPS-Lim3 ($p < 0,05$) et VPS-Lim9 ($p < 0,05$). Ces résultats sont visibles dans la Figure 29.

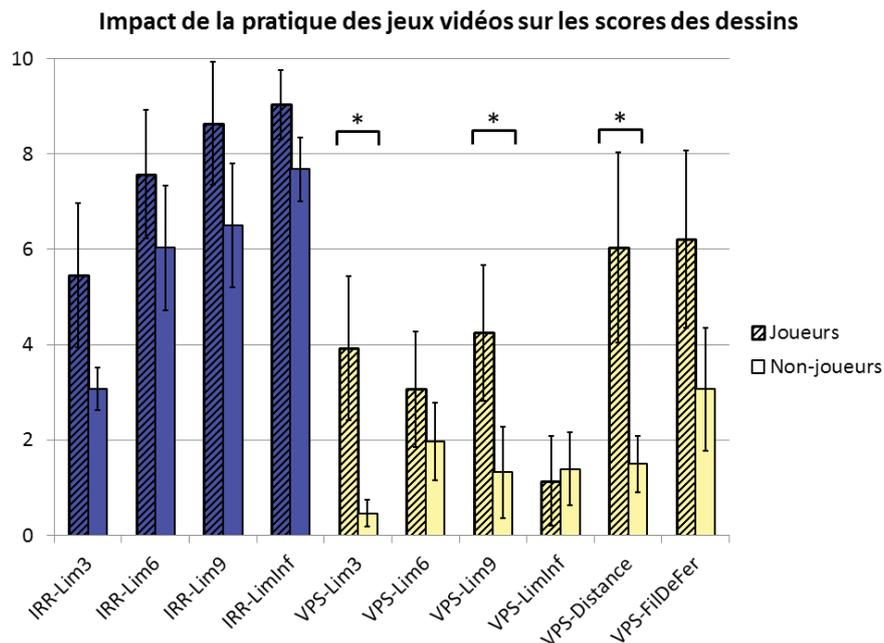


Figure 29 : Moyennes des scores des dessins obtenus par les groupes de joueurs et non-joueurs sur chacune des conditions. Les conditions Irrlicht sont représentées en bleu, celles VPS le sont en jaune.

C. Discussion

L'objectif de cette expérimentation était d'évaluer certaines possibilités offertes par l'utilisation de l'information de distance pour la navigation en VPS. Une première proposition de rendu était de limiter la profondeur du champ visuel à différentes distances. Si la limite de profondeur choisie est à l'infini, le rendu obtenu est le rendu brut. Dans ce cas la scène à représenter était trop riche, la quantité d'information trop importante par rapport à la taille de la matrice de phosphènes utilisée (15x18). Le fait de limiter la profondeur du champ visuel permettait de limiter par la même occasion la quantité d'information à afficher sur cette matrice de faible résolution. Il était alors possible pour les sujets de percevoir de façon plus détaillée les éléments qui étaient proches d'eux. En particulier, lorsque les sujets avançaient la limite entre les éléments visibles et ceux masqués était par définition à une distance constante des sujets et donc se déplaçait avec eux. Le fait qu'elle devançait les sujets leur permettait de détecter des changements dans la structure. Par exemple s'ils suivaient un couloir, la fin du couloir pouvait être détectée lorsqu'elle entrait soudainement dans le champ visible. Il en était de même pour les intersections. Les éléments structurant l'environnement pouvaient donc être détectés ce qui permettait de naviguer et également de mieux comprendre la structure de l'environnement et donc de la mémoriser. Ainsi le fait de limiter cette profondeur de champ permettait d'alléger l'information et compenser la faible résolution de la matrice. Cet allègement facilitait la compréhension des éléments proches et ensuite celle de la disposition des environnements.

Trois distances limites ont été testées : trois mètres (VPS-Lim3), six mètres (VPS-Lim6) et neuf mètres (VPS-Lim9). Elles ont été comparées au rendu où cette distance n'est pas limitée (VPS-LimInf). Pour les rendus VPS, notre hypothèse était que cette limite en désencombrant la perception aider à percevoir les éléments proches et donc la structure de l'environnement d'où une navigation facilitée. En revanche si l'on appliquait cette limite de profondeur à un rendu Irrlicht (donc de haute résolution), il n'y avait pas d'encombrement au départ, la limite devait apporter seulement un manque d'information et donc une gêne dans la navigation. Pour évaluer cette différence, trois conditions analogues aux rendus de VPS limités (IRR-Lim3, IRR-Lim6 et IRR-Lim9) ainsi que le rendu classique Irrlicht (IRR-LimInf) ont également été évalués. Un autre rendu (VPS-Distance) utilisait deux seuils de distance fixés à trois mètres et neuf mètres, pour que chaque phosphène reflète la distance aux éléments. Enfin, un dernier rendu (VPS-FilDeFer) permettait d'augmenter la perception des arêtes entre le sol et les murs, ce qui devait permettre de mieux percevoir la structure de l'environnement.

C.1. Résultats généraux

Pour les rendus Irrlicht, les résultats sur l'Indice de Performance montrent que le fait de limiter la profondeur du champ de vision à trois mètres ($M : 59,5$; $DS : 20,0$) a gêné les sujets pour accomplir la tâche rapidement, ce qui a entraîné une grande différence avec les autres rendus Irrlicht ($p < 0,001$ pour IRR-Lim9 et IRR-LimInf). Ainsi les sujets ne percevant qu'à très courte distance ont eu des difficultés pour anticiper leurs trajets et s'orienter. La charge cognitive ressentie générée par la tâche a également augmenté lorsque la distance limite de profondeur était à trois mètres. Cette condition présentait des différences significatives avec toutes les autres conditions Irrlicht ($p < 0,001$ pour chacune d'elles). Lors de la tâche, les sujets ayant des difficultés pour s'orienter ont été gênés pour mémoriser la structure de l'environnement et donc le restituer. Des différences significatives apparaissent ainsi lorsque l'on compare les dessins produits après avoir navigué en condition IRR-Lim3 et ceux en conditions IRR-Lim9 ($p < 0,001$) et IRR-LimInf ($p < 0,001$). Conformément à notre hypothèse générale et à nos hypothèses opérationnelles, le fait de borner la vision à trois mètres en rendu Irrlicht a entraîné des difficultés pour accomplir la tâche dans les temps et mémoriser la disposition de l'environnement. En revanche nous notons qu'il n'y a pas de différences significatives entre les rendus IRR-Lim6, IRR-Lim9 et IRR-LimInf pour aucune des variables dépendantes. Ceci signifie qu'en vision Irrlicht, le fait de percevoir les éléments jusqu'à seulement six mètres est suffisant pour naviguer et se représenter l'environnement dans lequel s'effectue la tâche. Dans nos environnements, les couloirs mesuraient cinq mètres de large, ainsi lorsque les sujets percevaient jusqu'à six mètres et au-delà, ils percevaient assez facilement les configurations des intersections. Nous supposons qu'ils pouvaient utiliser alors les espaces visibles depuis les intersections comme des *vista spaces*. Lorsque la limite était fixée à trois mètres seulement, il était difficile pour les sujets de repérer les intersections et la configuration qu'elles avaient. Ainsi certains sujets en naviguant ont pu confondre des angles simples avec des intersections et vice-versa. Ces confusions faisaient qu'ils avaient beaucoup de mal à prendre des points de repères en dehors de la base et donc à s'orienter.

Les différences entre les conditions VPS se distinguent des différences entre les conditions Irrlicht. Tout d'abord, contrairement aux rendus Irrlicht, quelle que soit la variable dépendante considérée, il n'y a pas de différence significative entre le rendu classique VPS-LimInf et le rendu VPS-Lim3. Ceci signifie que les sujets ont autant de difficultés pour effectuer la tâche avec la perception de l'environnement limitée à trois mètres de distance qu'avec une distance de perception non limitée. Pour rappel, les implants actuels ont un nombre d'électrodes (6x10) plus faible que le nombre de phosphènes utilisés (15x18) dans cette expérimentation et le rendu le plus courant qui est proposé aux personnes implantées correspond à notre rendu VPS-LimInf. Notre interprétation est que lorsque les sujets perçoivent en VPS sans limite, les éléments présents dans le champ visuel sont difficilement distinguables les uns des autres et encombrant la perception. Il est difficile alors pour eux de repérer les angles et les intersections présents dans les environnements et ils s'égarent assez rapidement. Pour certains sujets, ce rendu était si difficile et perturbant qu'ils ne parvenaient pas à trouver l'axe des couloirs dans lesquels ils étaient et se retrouvaient à heurter à répétition les murs de ces couloirs. Même si les éléments lointains sont présents dans le champ visuel avec ce rendu, il est très difficile pour les sujets de distinguer les intersections et de les utiliser comme points de repère. Les éléments facilement visibles pour les sujets sont les bijoux dont ils essaient de retenir la position par rapport à la base de départ mais avec grande difficulté. Deux dessins d'un même environnement effectués par deux sujets sont visibles dans la Figure 30.

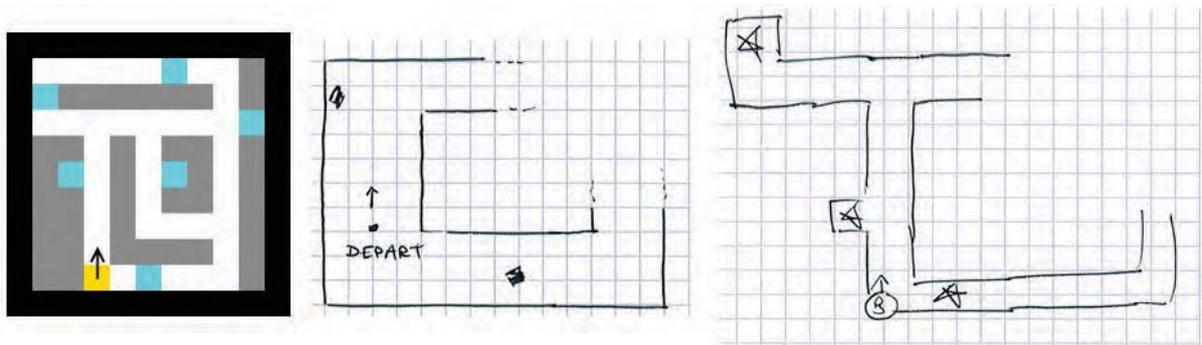


Figure 30 : Exemples de dessins produits par des sujets après exploration avec le rendu VPS-LimInf. Seules la base de départ et la position des bijoux proches sont indiquées.

Avec le rendu VPS-Lim3, les sujets rencontraient également des difficultés pour percevoir l'environnement. Pour rappel, en vision IRR-Lim3 les sujets rencontraient déjà des difficultés. Avec le rendu VPS-Lim3 Les sujets ne parvenaient pas à percevoir les intersections entièrement ou à les détecter. Nous avons observé que certains sujets ont traversé des intersections sans s'en apercevoir à plusieurs reprises. Ce rendu permettait de distinguer uniquement les murs s'ils étaient proches de la caméra. Les sujets avaient ainsi tendance à suivre les murs sans penser à leur orientation et se perdaient donc très rapidement.

Le rendu VPS-Lim6 et dans une moindre mesure le rendu VPSLim9 se démarquent des deux rendus VPS précédemment cités. Avec le rendu VPS-Lim6 les sujets parvenaient à effectuer plus rapidement la tâche qu'avec les rendus VPS-Lim3 ($p < 0,01$) et VPS-LimInf ($p < 0,001$). La charge cognitive ressentie était également plus faible en VPS-Lim6 et VPS-Lim9 qu'avec

le rendu VPS-LimInf ($p < 0,01$ et $p < 0,05$). En revanche, les dessins produits par les sujets en VPS-Lim6 ou VPS-Lim9 ne se distinguent pas des rendus VPS-Lim3 et VPS-LimInf. Avec le rendu VPS-Lim6 les éléments situés au-delà du masque situé à six mètres n'étaient pas visibles. Ce masque placé à six mètres permettait de visualiser les intersections dans leur globalité, ce qui a aidé les sujets pour accomplir la tâche de navigation rapidement. Avec le rendu VPS-Lim9, les sujets ne parvenaient pas au même niveau de performance qu'avec le rendu VPS-Lim6. Notre hypothèse est que tout comme le VPS-Lim6, ce rendu permettait de reconnaître la structure des intersections mais était déjà trop encombré pour permettre aux sujets de se déplacer facilement vers une intersection qu'ils venaient de détecter. Le rendu VPS-Lim6 permet de naviguer plus efficacement dans les environnements virtuels mais ne facilite pas nécessairement la mémorisation de la disposition des environnements. Sans doute que pour naviguer efficacement et éviter de se perdre, les sujets ont besoin de percevoir au-delà de six mètres dans les environnements proposés en VPS. Dans nos environnements contenant des couloirs larges de cinq mètres, il est possible avec le rendu VPS-Lim6 de repérer si les intersections sur lesquelles on arrive sont en croix ou en « T », en revanche il est impossible d'estimer à quelle distance se situe l'intersection suivante ou de quel type elle est. Or, le rendu VPS-Lim9 ne permet pas d'obtenir de meilleures performances pour la mémorisation que le rendu VPS-Lim6. Ainsi, autour du rendu VPS-Lim6 se situe peut-être l'équilibre pour nos environnements entre l'intérêt apporté par un allègement du percept et l'inconvénient de ne pas percevoir l'information distante. Si la profondeur est limitée au-delà de six mètres, il est plus difficile de se déplacer efficacement dans l'environnement, comme avec le rendu VPS-Lim9. Si la distance est plus courte, un manque d'information apparaît pour comprendre la structure de l'environnement et donc la mémoriser, comme avec le rendu VPS-Lim3. Le rendu VPS-Lim6 semble donc être proche de la valeur optimale pour effectuer cette tâche dans nos environnements.

Les rendus prothétiques les plus efficaces pour naviguer sont les rendus VPS-Distance et VPS-FilDeFer. Les sujets ont été plus efficaces pour effectuer la tâche de récolte des bijoux avec ces deux rendus qu'avec les rendus VPS-Lim3 et VPS-LimInf ($p < 0,01$ au maximum). De même la charge cognitive a été évaluée plus faible pour les rendus VPS-FilDeFer et VPS-Distance que pour les rendus VPS-Lim3 et VPS-Lim9. Enfin la différence entre la moyenne des scores des dessins en VPS-Distance et la moyenne des scores des dessins en VPS-LimInf est significative ($p < 0,05$). Il en est de même pour la différence entre les dessins produits en VPS-FilDeFer et ceux produits en VPS-LimInf ($p < 0,01$) ou VPS-Lim3 ($p < 0,05$). Ainsi l'hypothèse supposant l'avantage des rendus VPS-Distance et VPS-FilDeFer sur les rendus VPS-Lim3 et VPS-LimInf est validée, cependant il n'y a au final pas de différence significative entre VPS-FilDeFer ou VPS-Distance et VPS-Lim6 pour l'Indice de Performance, le score Nasa-TLX ou les scores des dessins. Les sujets ont eu des résultats similaires avec les deux rendus VPS-Distance et VPS-FilDeFer pour les trois variables dépendantes.

Le rendu VPS-Distance utilise trois niveaux de gris pour distinguer les objets proches, ceux entre trois et neuf mètres et finalement ceux situés au-delà de neuf mètres. Ce rendu

permet de détecter des intersections assez rapidement et d'identifier leurs configurations comme les rendus VPS-Lim6 et VPS-Lim9. De plus, ce rendu permettait également de détecter les murs si les sujets s'en rapprochaient à moins de trois mètres. Cette détection a grandement dû aider pour se déplacer au sein des environnements ou pour emprunter correctement une intersection. Ce qui apporte une aide supplémentaire par rapport au rendu VPS-Lim9. Le fait de détecter correctement les intersections et de se déplacer facilement dans l'environnement devait permettre aux sujets de mémoriser plus facilement les différentes intersections pour se construire une représentation mentale de l'environnement. Les espaces autour des intersections pouvaient alors servir de *vista space* pour structurer leur représentation mentale comme indiqué par le modèle de Meilinger [Meilinger 2008]. Le rendu VPS-FilDeFer était le seul rendu VPS à ne pas se baser sur l'information de distance mais sur la structure même de l'environnement en améliorant la détection des différentes surfaces présentes. En effet, ce rendu affichait les arêtes horizontales entre le sol et les murs ainsi que les arêtes verticales entre les différents murs. Ce rendu permettait ainsi facilement de percevoir la forme des couloirs avec les arêtes visibles correspondant au haut des murs. Avec ce rendu, la profondeur du champ de vision n'était pas limitée et les sujets pouvaient donc détecter les intersections depuis une plus grande distance qu'avec le rendu VPS-Distance ou VPS-Lim9. Une fois que les sujets étaient à proximité des intersections, il était également facile pour eux d'identifier la configuration de l'intersection. De plus il était également facile de se déplacer avec ce rendu, les intersections entre les murs et le sol étant également affichées. Ainsi le rendu VPS-FilDeFer présentait trois avantages principaux que les autres rendus en VPS n'avaient pas simultanément. Premièrement les intersections pouvaient être détectées de plus loin qu'avec tous les autres rendus prothétiques car la profondeur du champ visuel n'était pas limitée et de plus ce rendu en affichant uniquement la structure limitait considérablement l'encombrement très présent pour le rendu VPS-Liminf. Deuxièmement, le fait d'afficher uniquement les arêtes permettait d'identifier très rapidement les configurations des intersections car les angles mêmes de ces intersections étaient rendus. Enfin, à courte distance, les murs pouvaient également être détectés grâce aux arêtes horizontales visibles et délimitant le sol des murs. Ainsi pour ne pas heurter un mur il suffisait de se fier aux arêtes visibles au niveau du sol.

C.2. Avantage des joueurs

Les comparaisons de l'Indice de Performance des joueurs et celui des non-joueurs montrent que les joueurs sont parvenus à effectuer la tâche plus rapidement que les non-joueurs pour les conditions Irrlicht IRR-Lim3, IRR-Lim6 et IRR-Lim9. Cette différence peut s'expliquer par deux éléments principaux. Contrairement au groupe des non-joueurs, les sujets joueurs étaient habitués à naviguer dans des environnements virtuels. Ainsi ils étaient déjà familiers avec ces types de rendu Irrlicht. De plus les rendus limitant la profondeur sont assez courants dans le jeu vidéo car le brouillard artificiel permet à la fois de limiter les ressources machines nécessaires et d'instaurer une ambiance anxieuse. Ce procédé a par exemple été utilisé pour le jeu *Silent Hill* (Konami, 1999) afin de générer une ambiance menaçante dans un milieu ouvert. La profondeur

est également limitée lorsque le personnage incarné par un joueur porte une torche qui lui permet de voir à proximité mais qui ne lui permet pas de voir au loin. Cet effet très classique est par exemple utilisé dans le jeu *Amnesia : The Dark Descent* (Frictional Games, 2010) pour angoisser le joueur qui se doit d'explorer l'environnement. Le jeu *Tomb Raider* (Square Enix, 2013) utilise des jeux de lumières autour des torches pour instaurer une ambiance et guider le joueur. Enfin le mod *Dynamic Lights* développé par AtomicStryker pour le jeu *Minecraft* (Mojang, 2011) permet de se déplacer dans le noir à l'aide de la lumière d'une torche pour ajouter du réalisme au jeu. Ces exemples sont illustrés dans la Figure 31. Ainsi pour les rendus Irrlicht limités, les joueurs avaient sans doute déjà été en contact de rendus voisins ou reposant sur les mêmes mécaniques.

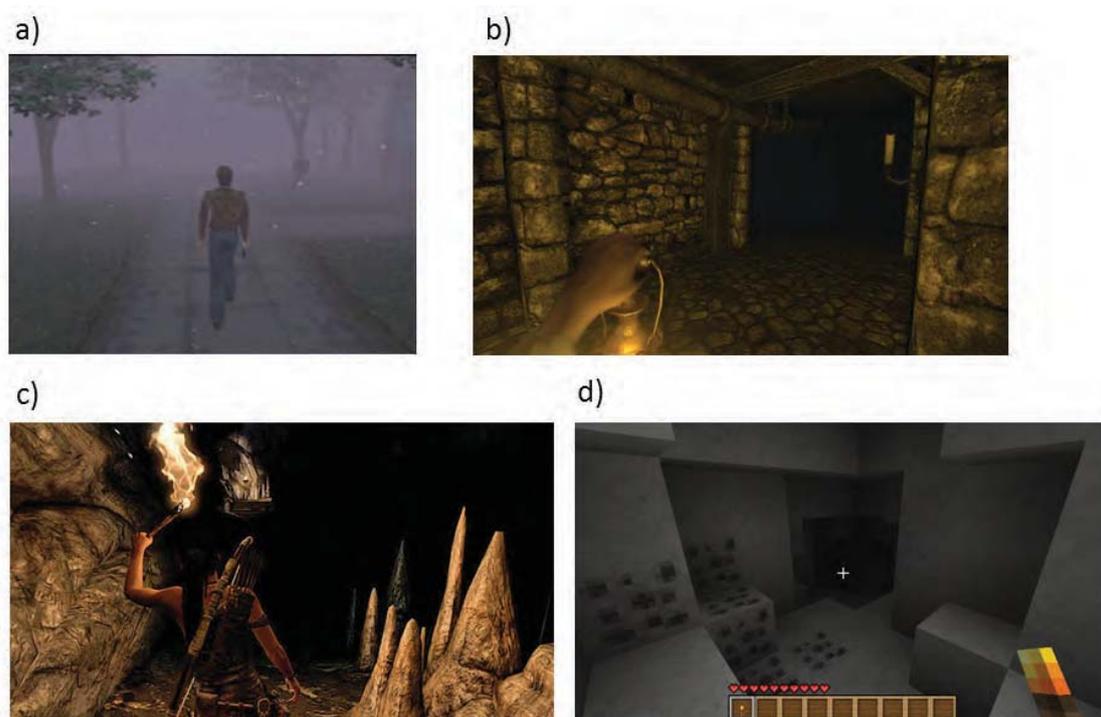


Figure 31 : Différents jeux vidéo classiques limitant la distance de vision. a) Les jeux *Silent Hill* et *Amnesia : The Dark Descent* utilisent la limitation de la profondeur pour amener de l'angoisse aux joueurs. c) Le jeu *Tomb Raider* l'utilise pour donner une ambiance propice à l'exploration et pour diriger le joueur vers des points lumineux. d) Le mod *Dynamic Light* développé pour *Minecraft* permet de rajouter une touche de réalisme au jeu de départ.

Cette habitude de l'utilisation de tels rendus leur a probablement permis de se distinguer des non-joueurs et de se déplacer plus efficacement dans les environnements virtuels.

Le deuxième élément qui a également sans doute aidé les joueurs repose sur les contrôles utilisés pour l'expérimentation. Pour se déplacer dans les environnements, les joueurs devaient utiliser les flèches du clavier. Cette interaction est extrêmement classique dans le domaine du jeu vidéo mais a pu surprendre, voire dérouter, les non-joueurs qui ne retrouvaient pas le rythme de la marche dans un appui continu sur une touche même si des bruits de pas étaient présents.

Nous supposons que la combinaison de ces deux facteurs a pu amener les joueurs à se distinguer avec les rendus Irrlicht IRR-Lim3, IRR-Lim6 et IRR-Lim9. De même en VPS, les joueurs ont sans doute compris très rapidement comment le rendu VPS-Distance fonctionnait et ont pu l'exploiter rapidement et facilement. Même si une courte familiarisation a été faite,

l'absence d'utilisation de tels rendus dans un autre contexte par les sujets non-joueurs les a sans doute gênés pour effectuer la tâche de navigation proposée. Des différences significatives sur des rendus utilisant la notion de distance sont également présentes pour les scores des dessins produits par les sujets. Les conditions VPS-Lim3, VPS-Lim9 et VPS-Distance présentent de telles différences. Les ressources cognitives des sujets joueurs parvenant à se déplacer facilement dans les environnements étaient ainsi sans doute focalisées sur la mémorisation de la disposition de l'environnement. À l'inverse, celles des sujets non-joueurs étaient probablement dévouées à la concentration sur les contrôles et l'interprétation des rendus.

Une autre différence significative entre joueurs et non-joueurs concerne l'Indice de Performance pour la condition VPS-FilDeFer. Ce rendu ne se base pas sur la distance mais sur la structure même des environnements. Ce rendu en fil de fer a été utilisé au début du développement des jeux vidéo, l'affichage de textures n'étant pas encore possible dans des environnements 3D avec le matériel de l'époque. Ainsi certains jeux comme Akalabeth : World of Doom (California Pacific Computer Company, 1979), Battlezone (Atari, 1980) ou StarWars : The Arcade Game (Atari, 1983) utilisaient uniquement ce rendu pour afficher des environnements 3D. Ce rendu a également été utilisé pour des jeux plus récents tels que Rez HD (United Game Artists, 2001). Des images de ces différents jeux sont visibles dans la Figure 32.

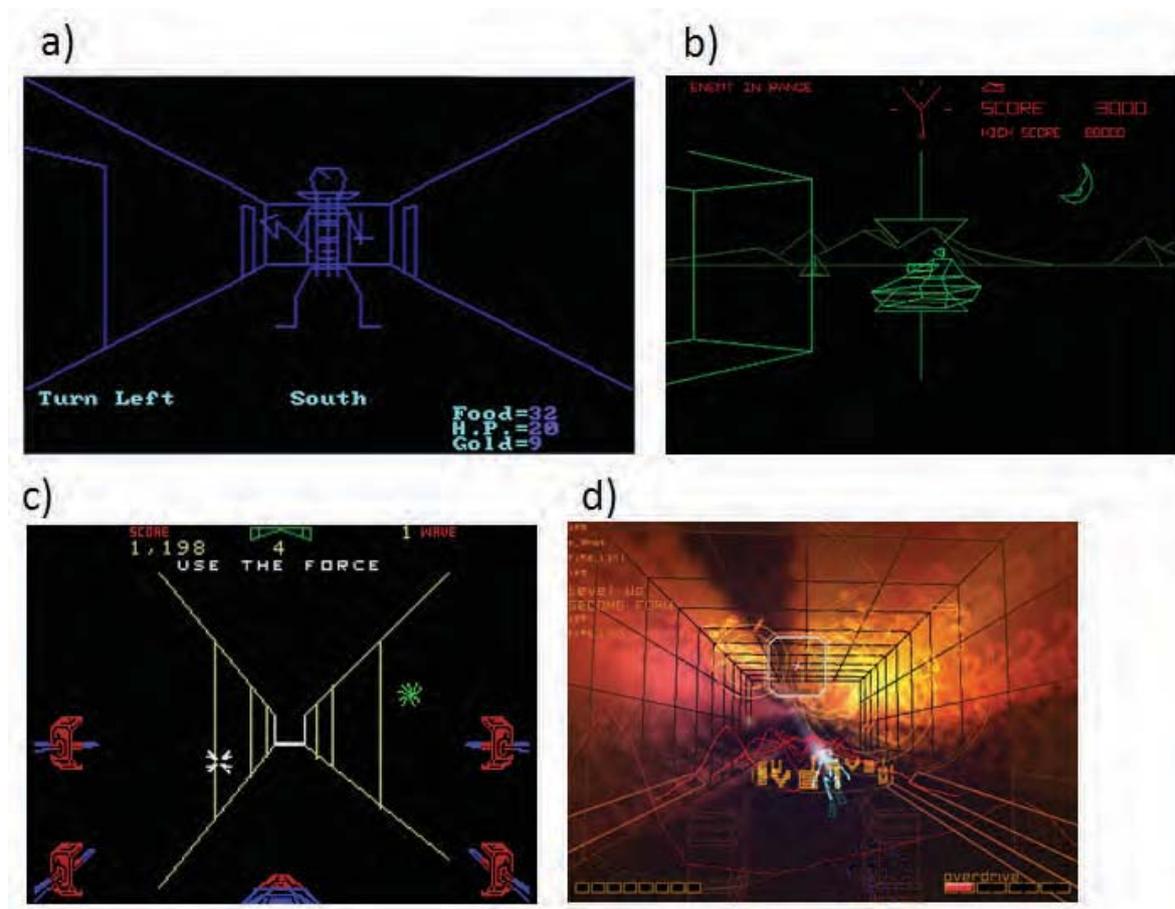


Figure 32 : Jeux vidéo utilisant un rendu en fil de fer. a) Akalabeth : World of Doom est un des premiers jeux de rôle dans lequel des donjons sont à explorer. b) Battlezone a connu un grand succès et proposait des combats entre tanks. c) Star Wars : The Arcade Game est un simulateur spatial également très populaire. d) REZ est un railshooter en 3D utilisant des rendus en fil de fer pour la plupart de ses éléments.

Même si le rendu en fil de fer n'est pas très utilisé dans les jeux vidéo récents qui ne souffrent plus des contraintes techniques des jeux des années quatre-vingt, il reste un rendu familier des joueurs. Lors de l'expérimentation les joueurs ont donc pu comprendre très rapidement le fonctionnement de ce rendu et être très à l'aise pour l'interpréter et se déplacer dans le labyrinthe. Les non-joueurs qui découvraient ce rendu lors de la familiarisation ont sans doute eu plus de difficultés pour l'interpréter d'où la différence sur l'Indice de Performance. Les joueurs et non-joueurs tendent aussi à se différencier sur les scores des dessins, même si cet écart est non significatif dans notre expérimentation ($p < 0,61$).

Ainsi pour certains rendus, les joueurs ont obtenu de meilleurs résultats que les non-joueurs. Si nous effectuons des tests statistiques en incluant toutes les conditions, nous remarquons que les joueurs et non-joueurs se distinguent significativement pour l'Indice de Performance ($p < 0,05$) et les scores des dessins ($p < 0,05$). En revanche ces deux groupes ne se distinguent ni pour la charge cognitive mesurée par le Nasa-TLX ($p > 0,8$) ni pour le sens de l'orientation ($p > 0,7$). Ces résultats sont visibles dans la Figure 33. Ceci signifie principalement deux choses. Que les sujets soient familiers avec des rendus ou qu'ils ne les connaissent que sommairement, la difficulté ressentie est à peu près la même. Par exemple, les non-joueurs n'ont pas ressenti beaucoup plus de frustration que les joueurs alors qu'ils étaient moins performants pour effectuer la tâche. Sur le plan fonctionnel, chez des patients implantés, il serait ainsi envisageable d'utiliser de tels rendus en environnement virtuel pour procéder à un apprentissage avant d'utiliser ces rendus en extérieur. Le deuxième élément est que les joueurs et non-joueurs ne se démarquent pas particulièrement au niveau de leur sens de l'orientation. L'effet de la pratique des jeux vidéos a peut-être limité l'influence du sens de l'orientation pour cette expérimentation qui ne corrèle pas avec les variables dépendantes. Il est également possible que nos environnements soient trop simples pour que le sens de l'orientation ait un impact sur les performances des sujets.

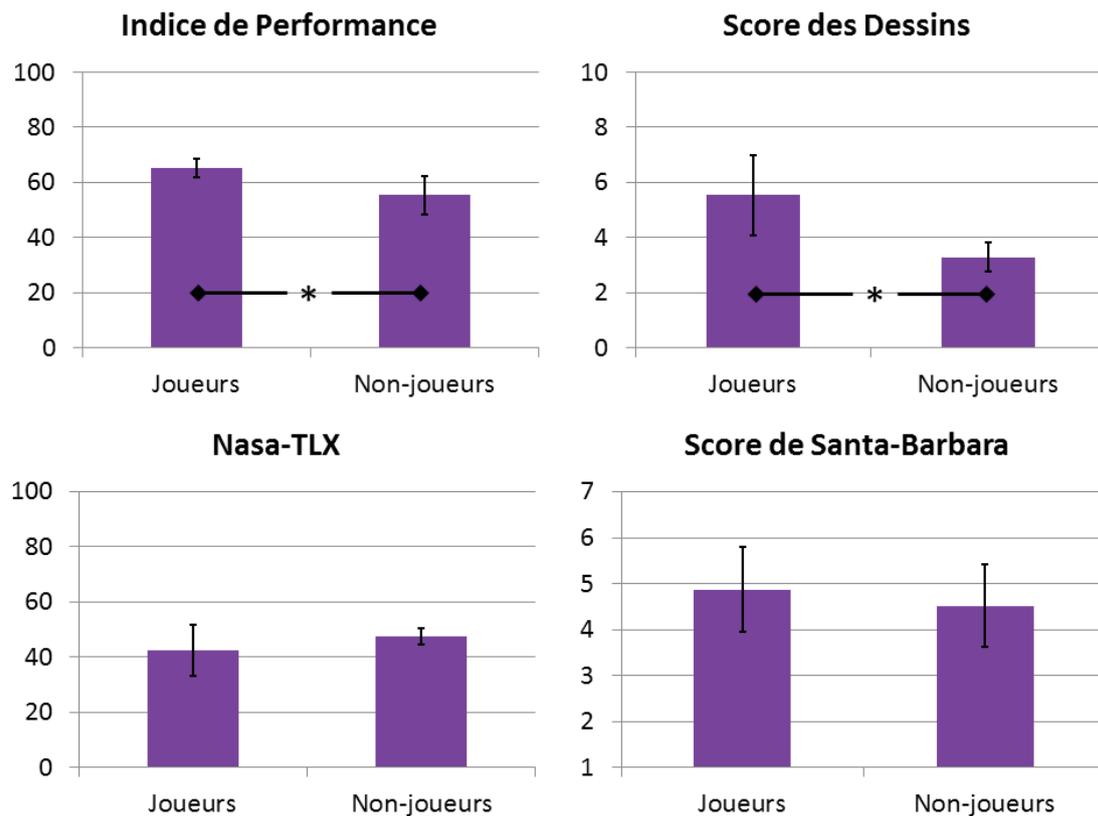


Figure 33 : Différences entre les groupes de Joueurs et Non-joueurs sur les variables dépendantes et le score de Santa-Barbara. Les moyennes ont été obtenues en moyennant sur toutes les conditions. Les barres d'erreurs sont les intervalles de confiance à 95 %.

D. Conclusion

Cette expérience a permis d'étudier en profondeur l'impact de différents types de rendus visuels dans une tâche de navigation dans un environnement inconnu. La comparaison des quatre conditions en vision Irrlicht a montré que la limitation de la profondeur disponible dans le champ visuel gênait la navigation des sujets. En revanche ces mêmes limitations de la distance de vision, appliquées à la vision prothétique, ont produit des résultats très différents. En effet, si cette limitation était fixée à une courte distance (trois mètres), les sujets étaient mis en difficulté et obtenaient de mauvaises performances. En revanche si cette distance était située à une distance « adéquate » (six mètres pour nos environnements) les sujets parvenaient à naviguer et à obtenir de meilleures performances qu'avec le rendu VPS classique sans limite. Ainsi le fait de limiter la profondeur à six mètres a permis aux sujets de percevoir plus facilement la structure des environnements et de s'y déplacer plus aisément. En revanche, la faible portée de ce rendu limitait l'orientation des sujets ainsi que la mémorisation des environnements dans lesquels ils naviguaient. Le fait d'améliorer la perception de la distance aux éléments, même avec un nombre limité de niveaux de gris (quatre pour cette expérience) a permis aux sujets de s'orienter et de mieux mémoriser les environnements. Un autre rendu basé non sur la distance mais sur la structure même de l'environnement et qui facilitait la perception des arêtes entre les surfaces de l'environnement (comme entre le sol et les murs)

a été aussi expérimenté. En VPS, les sujets ont obtenu les meilleurs résultats avec ce dernier rendu qui leur a permis d'effectuer la tâche rapidement en plus de mémoriser correctement la structure des environnements. Ainsi l'information de la distance ou l'affichage des arêtes de la structure de l'environnement semble être de bons candidats pour tenter de restituer des capacités de navigation à des personnes disposant d'implants de faible résolution.

Chapitre VIII.

Navigation et points de repères

La première expérimentation effectuée dans ces travaux utilisait un environnement virtuel contenant différents points de repères qui étaient représentés en VPS par un code visuel afin de les distinguer les uns des autres. Les environnements dans lesquels se sont déroulées les expérimentations suivantes étaient très simples et ne contenaient pas de repères en dehors du point de départ et de la structure même de l'environnement (intersections, coudes). Les rendus que nous avons évalués étaient adaptés pour ces environnements simples et ont permis d'obtenir de bons résultats en faisant percevoir plus facilement la structure de l'environnement. Pour des environnements plus complexes en revanche il n'est pas garanti qu'ils soient suffisants. Par exemple si une personne implantée souhaite se déplacer en ville de nombreuses intersections peuvent être empruntées et il est peu probable que toutes ces intersections se distinguent les unes des autres uniquement par leur structure. Il faudrait alors compléter ces rendus en affichant des repères supplémentaires reconnus par exemple avec des algorithmes de traitement d'images. De plus, il ne faudra pas afficher trop de points de repères simultanément pour éviter de surcharger la matrice de faible résolution et retomber dans les limites du rendu brut [Rheede et al. 2010]. Dans un premier temps, les éléments affichés en plus doivent permettre de compléter les *vista spaces* des patients implantés afin que ces derniers parviennent plus facilement à les reconnaître. Il faut donc identifier des points de repères permettant potentiellement de caractériser ces lieux. Les *vista spaces* qui semblent les plus probablement utilisés par des personnes voyantes naviguant en ville sont les grands espaces comme les places. Une méthode pour sélectionner des points de repères est de demander à plusieurs personnes d'indiquer ceux qu'elles utilisent pour identifier un lieu ou se repérer sur un même trajet. Nous pouvons alors faire l'hypothèse que les repères les plus rapportés sont de bons candidats pour être indiqués aux personnes implantées.

A. Protocole

Pour avoir quelques pistes sur les points de repères à utiliser pour des rendus de faible résolution, nous avons mis au point un questionnaire exploratoire en ligne portant sur deux places du centre de Toulouse : la place du Capitole et la place Esquirol. Nous avons demandé aux participants d'indiquer les éléments qui caractérisent pour eux ces lieux ainsi que les repères importants attachés à ces lieux. Nous leur avons également demandé de nous décrire le trajet allant de la place Esquirol à la place du Capitole. Il est extrêmement difficile d'inférer la construction mentale d'une personne à partir d'une description qu'elle donne d'un trajet [Levelt 1982; Denis et al. 1999] ou même de la description d'un point de repère [Lee & Tversky 2005]. Ainsi nous ne souhaitons pas analyser les représentations mentales des participants.

Notre objectif était simplement d'obtenir une liste d'éléments qui pourraient servir de repères et pourraient peut-être s'avérer pertinents à augmenter pour un rendu VPS. Les résultats obtenus ici n'ont pas vocation à être généralisés et nous éviterons donc de les sur-interpréter. Ce travail est très exploratoire et sa destination est principalement de préparer un autre questionnaire plus large et plus précis sur différents lieux de Toulouse ainsi que de nourrir la discussion sur les points de repère à représenter en VPS.

B. Sujets

Pour ce premier questionnaire, sept personnes voyantes ont répondu au questionnaire (V1-V7). Nous avons également contacté l'Institut des Jeunes Aveugles de Toulouse ce qui nous a permis de faire également participer à ce questionnaire exploratoire deux personnes non-voyantes (NV1-NV2) et une personne malvoyante (MV1).

C. Résultats

Pour la place du Capitole, les sujets ont cités de nombreux éléments. Sur dix sujets, neuf ont cité l'hôtel de ville. Le restaurant rapide McDonald's présent sur la place a été cité par sept sujets dont les deux non-voyants. Deux autres éléments se détachaient en étant cités par six sujets : les arcades et l'entrée de la rue Saint-Rome. Ces éléments sont visibles dans la Figure 34. Ainsi pour aider les sujets implantés à reconnaître cette place, il semble que des éléments intéressants à augmenter soient la façade de l'hôtel de ville ainsi que l'enseigne de restauration rapide.



Figure 34 : Place du Capitole. Les éléments cités par les sujets sont représentés. Les points rouges, bleus et violet correspondent aux positions prises par les sujets pour décrire cet environnement. Certains sujets ont décrit l'environnement depuis une position allocentrée et ne sont donc pas représentés.

Pour la place Esquirol, l'élément qui est cité le plus par les sujets est l'enseigne Midica. La Figure 35 indique les différents éléments cités par les sujets.

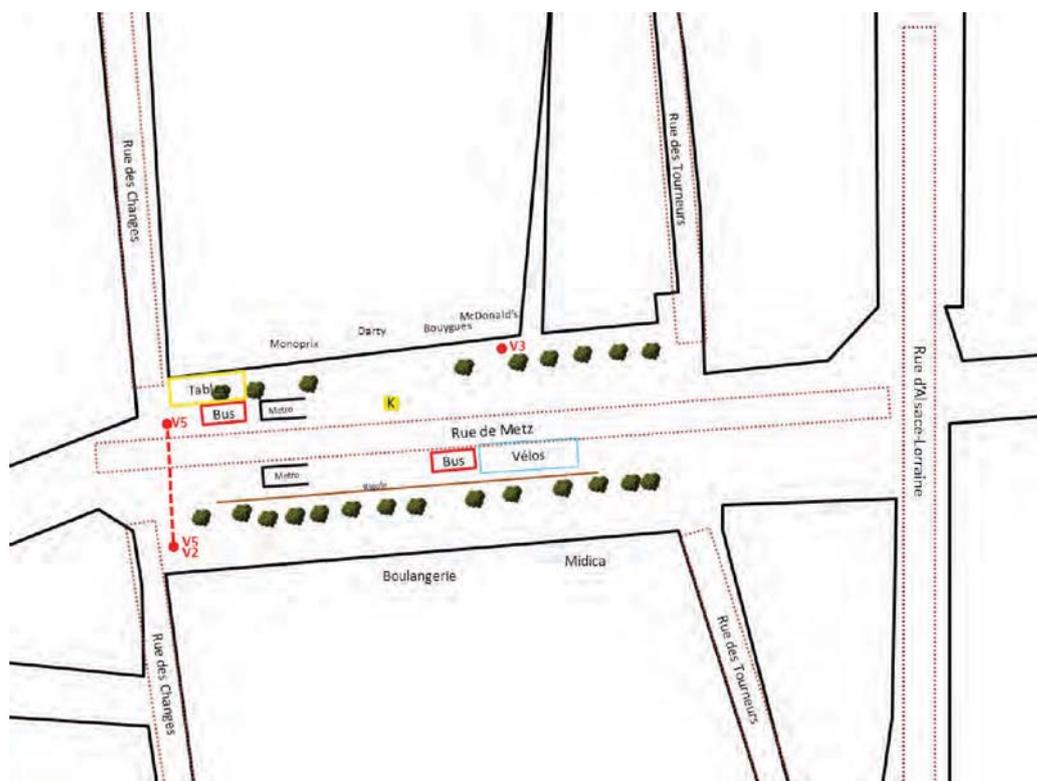


Figure 35 : Place Esquirol. Les éléments cités par les participants sont représentés sur ce plan.

La place Esquirol est coupée en deux par une voie de circulation, ce qui gêne sans doute le fait de se la représenter comme un seul *vista space*. Ainsi les réponses des sujets pour ce lieu sont très variables, et certains décrivent cette place en s'imaginant naviguer à l'intérieur. Ainsi, la structure de cette place est sans doute trop complexe pour la considérer comme un seul *vista space* potentiel. Nous avons demandé également aux sujets de nous décrire le trajet entre la place Esquirol et la place du Capitole. Ce trajet consiste essentiellement à suivre une voie piétonne en ligne droite et très peu d'éléments descriptifs nous ont été donnés par les sujets sur ce trajet.

Il semblerait que les éléments les plus pertinents à représenter pour des personnes implantées souhaitant naviguer dans le centre de Toulouse soient la façade de l'hôtel de ville pour la place du Capitole et l'enseigne Midica pour la place Esquirol. Les résultats détaillés de ce questionnaire sont disponibles dans les annexes de ce manuscrit.

D. Discussion et Conclusion

Ce petit questionnaire exploratoire nous a permis de poser les bases pour développer un nouveau questionnaire qui sera plus complet sur différents lieux de Toulouse. De plus, de nouveaux éléments de réflexions sur la représentation de points de repères en plus de la structure pour des rendus de VPS peuvent être amenés par ces premiers résultats.

Le trajet que les sujets devaient décrire a été décidé pour que des non-voyants puissent également remplir le questionnaire. Le problème est qu'il était alors un peu trop simple pour des voyants pour qu'ils nous proposent des points de repère utilisables facilement. Le trajet passant essentiellement par une rue commerçante, la plupart des sujets ont utilisés des termes génériques pour identifier les boutiques présentes dans cette rue. Ainsi si d'autres trajets sont à étudier, il faudrait qu'ils soient un peu plus complexe et impliquent des prises de décision en des lieux précis. Il y a des chances que ces lieux précis soient utilisés comme des *vista spaces* pour les personnes voyantes connaissant ce trajet. Il pourrait être demandé à des non-voyants quels sont les points de repère éventuels qu'ils utilisent en ces points de décision sans pour autant leur demander de décrire l'ensemble du trajet. Pour les rendus VPS supportant la navigation, il semble que pour les places du Capitole et Esquirol, les points de repère les plus utilisés correspondent à certaines façades. Il serait possible de rendre les environnements en VPS en utilisant la base du rendu FilDeFer en fond et de rendre avec des phosphènes plus lumineux les façades utilisées comme point de repères principaux sur ces places. Ce questionnaire exploratoire nous a permis de mieux concevoir un futur questionnaire portant sur les points de repère pertinent à représenter en VPS. Il semble que les trajets dont la description est demandée aux sujets doivent être suffisamment complexes pour contenir plusieurs points de décisions. Il pourrait être demandé des informations complémentaires sur les points de décision pour préciser les points de repère nécessaires à l'orientation en ville.

Les différents résultats obtenus dans ces expérimentations sont discutés dans la partie suivante ce qui nous permet d'ouvrir des perspectives sur l'utilisation fonctionnelle des implants rétiniens de prochaine génération dans le cadre de la navigation.

PARTIE 3. DISCUSSION
ET PERSPECTIVES

Actuellement, les neuroprothèses visuelles ne semblent pas permettre d'effectuer des tâches de navigation. Les personnes implantées sont évaluées uniquement sur des tâches de mobilité qui restent difficiles à effectuer pour les patients [Humayun et al. 2012]. Cette grande difficulté est due à la faible résolution des implants, les implants actuels disposent de moins d'une centaine d'électrodes. Les implants de génération suivante disposeront de plus de cent électrodes [Stronks & Dagnelie 2014] mais ce nombre reste en dessous de la résolution évaluée comme minimale pour effectuer une tâche de navigation [Cha et al. 1992]. En 2008, des prévisions étaient faites sur le développement des implants suffisamment rapide pour arriver à des résolutions dépassant les deux cents électrodes et concernant dix mille personnes dès 2010 [Chader et al. 2009]. Il s'avère que le développement des implants a été plus long que prévu (voir Figure 1). En septembre 2015, cent cinquante personnes disposent de l'implant Argus® II qui contient soixante électrodes. L'implant IRIS® développé par l'entreprise Pixium Vision dispose de cent cinquante électrodes mais ne sera pas commercialisé avant 2016. La navigation ne semblait pas accessible pour les personnes implantées, les implants actuels ne satisfaisant pas les prérequis nécessaires pour la navigation.

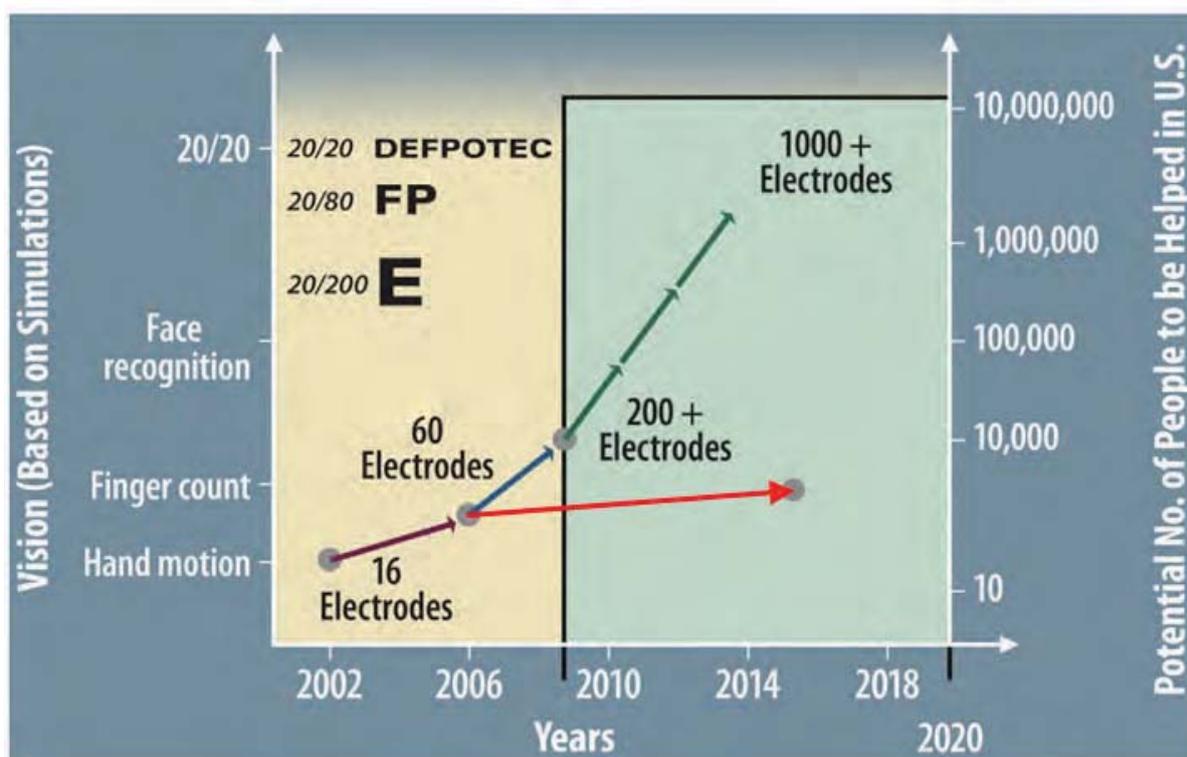


Figure 36 : Prévision faite en 2009 sur le développement des implants. L'échelle de gauche illustre les capacités obtenues avec ces résolutions et l'échelle de droite, le nombre potentiel de personnes implantées. La flèche rouge qui a été ajoutée sur cette image indique la situation actuelle. Source : [Chader et al. 2009]

La plupart des neuroprothèses visuelles dont les implants épi-rétiniens qui sont les plus avancés aujourd'hui utilisent une caméra portée par les sujets. Cette caméra permet de capter l'information visuelle qui est transmise à l'implant. Il est donc possible d'utiliser des traitements de vision par ordinateur sur ces images afin d'optimiser le rendu prothétique final.

Il est extrêmement difficile de prévoir quand arriveront des implants de haute résolution, permettant la navigation sans traitement de l'image. En revanche il est possible de prévoir des traitements de l'image permettant de fournir des rendus prothétiques optimisés pour la navigation, même avec une résolution faible des implants. Ces rendus peuvent être évalués grâce à l'outil de la Vision Prothétique Simulée (VPS). Les travaux effectués en VPS se rapprochant de la navigation, se concentrent principalement sur la tâche de mobilité [McCarthy et al. 2014] et proposent différents rendus pour afficher des obstacles [Dowling et al. 2005]. À de rares exceptions près [Wang et al. 2008; Zapf et al. 2014; Parikh et al. 2013] les matrices de phosphènes utilisées correspondent à des implants de haute résolution. La navigation à l'aide de matrices de faible résolution n'a été que très peu étudiée. De même, les représentations mentales construites par les sujets lors de la navigation en VPS n'avaient pas encore été étudiées.

Chapitre IX.

Principaux résultats

Les travaux de cette thèse ont porté sur la tâche de navigation avec des rendus prothétiques de faibles résolutions. Notre axe de recherche était de proposer des approches permettant aux patients implantés de naviguer ainsi que de construire une représentation mentale de l'environnement. Trois approches différentes ont été évaluées pendant ces travaux. La première approche évaluée était de faire percevoir uniquement les points de repères (ou *landmarks*) reconnus par un algorithme de traitement d'image. La deuxième approche était d'utiliser la distance aux éléments pour limiter la profondeur du champ de vision ou faire percevoir la distance aux éléments rendus. Enfin, la dernière approche était d'utiliser la reconnaissance de la structure des lieux par un traitement de vision par ordinateur puis de rendre cette structure en vision prothétique.

Le premier rendu de VPS non classique que nous avons évalué est une application de l'approche par localisation à la navigation. Seuls les éléments présents dans l'environnement pouvant servir de points de repère étaient affichés. Ce rendu représentait les éléments de l'environnement grâce à un code visuel pour chacun des types de points de repère et n'affichait pas la structure de l'environnement. L'hypothèse principale qui fondait cette première expérience utilisait le modèle de Siegel et White [Siegel & White 1975] et était que seuls les éléments présents dans l'environnement sont utilisés en tant que points de repère et que les points de repère sont peut-être suffisants pour permettre la navigation. Lors de cette expérience un rendu par localisation proposant seulement les points de repère à afficher a été testé. Les résultats de cette première étude exploratoire ont montré que la représentation de la structure de l'environnement est nécessaire pour permettre l'orientation et la mobilité. De plus si trop de points de repère sont représentés simultanément, les sujets n'arrivent plus à interpréter les codes visuels qui se confondent. Ainsi l'approche par localisation semble très limitée en VPS dans une tâche de navigation.

La deuxième expérimentation a évalué le rendu d'augmentation de structure. Le sol et les murs étaient représentés avec des phosphènes de luminosités différentes afin de créer un contraste. Des rendus de ce type ont déjà été proposés en VPS [Wang et al. 2008; Srivastava et al. 2009] mais n'ont jamais été évalués dans une tâche d'exploration libre. Ce rendu a permis aux sujets de mieux comprendre l'environnement et d'identifier plus facilement les intersections. En rendu brut il était possible pour un sujet de confondre un croisement en « T » et un angle, ce qui l'amenait à manquer un embranchement. Ce type d'erreurs n'a quasiment pas été observé avec le rendu augmentant la structure. Ce rendu a permis aux sujets d'explorer plus rapidement l'environnement et de construire une représentation mentale de l'environnement.

Pour la troisième expérimentation, la théorie de Meilinger du *Network of Reference Frames* [Meilinger 2008] ainsi que la notion de *vista space* introduite par Montello [Montello 1993] furent utilisés pour aider au design de certains rendus. La faible résolution des implants et le peu de niveaux de gris disponibles font que le rendu classique brut est rapidement encombré par un nombre important d'éléments à afficher. Ces éléments se distinguent mal les uns des autres et il est difficile de comprendre la scène ou la portion de l'environnement captée par la caméra. Notre hypothèse de travail était que le fait de masquer les éléments situés au-delà d'une certaine distance permettait d'alléger le rendu tout en permettant la navigation. Le principal intérêt de ce rendu est de faciliter la mobilité, les sujets parvenant à se déplacer dans l'environnement plus facilement qu'avec le rendu brut. L'inconvénient principal est que la limite de perception de la profondeur impose aux sujets un effort d'intégration temporel dans la mémorisation de l'environnement par rapport à des rendus permettant de percevoir plus loin comme le rendu fil-de-fer.

L'utilisation du rendu en fil-de-fer qui fait apparaître les arêtes entre le sol et les murs a permis aux sujets d'obtenir de meilleures performances de navigation, de mieux retenir la disposition de l'environnement et ce, pour une charge cognitive moindre. Une explication possible de l'efficacité de ce rendu repose sur la profondeur de champ disponible et sa simplicité. En effet les seuls éléments de l'environnement qui sont représentés avec ce rendu sont les arêtes verticales et horizontales des différentes surfaces de l'environnement (murs et sol). Ce rendu permet de simplifier considérablement la perception de l'environnement par rapport au rendu brut. De plus, contrairement aux autres rendus de l'expérience utilisant la notion de distance, la profondeur du champ de vision n'est pas limitée pour ce rendu. Cette absence de limite a peut-être permis aux sujets de construire des *vista space* de taille plus large qu'avec les autres rendus [Meilinger 2008]. Si un environnement est représenté à l'aide de *vista spaces* plus larges, dans ce cas le nombre de *vista spaces* est également plus faible. D'après le modèle de Meilinger ceci implique un nombre d'éléments à mémoriser et à manipuler plus faible pour naviguer et donc une charge cognitive moins importante.

Les rendus de distance consistent à utiliser des phosphènes clairs pour présenter des éléments qui sont proches de l'utilisateur et des phosphènes plus sombres pour des éléments plus éloignés. L'information de distance permet d'afficher un dégradé, comptant autant de nuances que de niveaux de gris, en fonction de la distance de l'élément affiché. Les résultats de cette expérimentation indiquent que l'utilisation de rendu exploitant la distance peut permettre de percevoir l'environnement plus clairement qu'avec le rendu brut. Dans une navigation en ville, de courtes distances seuils (inférieures à vingt mètres) peuvent être nécessaires pour percevoir les intersections dans une rue étroite. En débouchant sur une grande place, de plus grandes distances devront sans doute être utilisées pour pouvoir détecter les limites de cette place. Ainsi pour que ce rendu soit utile les distances utilisées comme seuils ne peuvent être décidées au hasard, il est nécessaire de les fixer en fonction des dimensions des environnements dans lesquels la navigation s'effectue. Une possibilité est de laisser l'utilisateur choisir lui-même ces seuils au fur et à mesure de la navigation. Une autre possibilité est de détecter la distance moyenne des éléments environnants et de proposer

automatiquement des valeurs qui en dépendent. Un rendu adaptatif de ce type permettrait alors de naviguer en ville et de représenter la structure des quartiers dans leurs diversités.

Lors de la deuxième et troisième expérimentation, nous avons évalué l'apport de différents rendus pour la navigation en vision prothétique simulée. En particulier, nous avons montré que rendre la structure de l'environnement même simplement avec peu de niveaux de gris est suffisant pour naviguer et se construire une représentation mentale de l'environnement. Ces rendus n'affichaient donc aucun point de repère à proprement parler, les seuls repères utilisables étant la structure des intersections elle-même ou le point de départ dans la tâche de navigation. La première expérimentation a montré que rendre uniquement des points de repère ne favorisaient pas la navigation. Le modèle de Meilinger indique que les points de repère sont utilisés et présent dans les *vista spaces*, donc nous supposons que les rendus affichant la structure peuvent être augmentés avec des points de repère pour faciliter la navigation et la mémorisation de l'environnement. Il est possible d'imaginer un rendu prothétique utilisant un des rendus évalués en arrière-plan pour afficher la structure et de l'augmenter en affichant des points de repère sélectionnés au préalable.

La quatrième expérimentation reposait sur cette hypothèse de facilitation de la navigation par l'augmentation de rendus prothétiques avec des points de repère pertinents. Notre souhaitons interroger des personnes naviguant régulièrement dans Toulouse pour avoir un premier retour sur les points de repère qui seraient pertinents à rendre en vision prothétique. Nous avons fait passer un questionnaire exploratoire portant sur la navigation dans le centre-ville de Toulouse. L'objectif de ce questionnaire en ligne était de préciser quels étaient les types de points de repère que l'on pourrait afficher pour des personnes souhaitant se déplacer en ville à l'aide d'un implant. Des questions étaient posées sur deux places du centre-ville de Toulouse ainsi que sur le trajet pour aller de l'une à l'autre. Pour les deux places sur lesquelles les sujets ont répondu, les points de repère principaux semblent être certaines façades permettant d'identifier précisément ces places. Il semble que le trajet proposé était trop simple pour obtenir des points de repère pertinents. Ces résultats donnent des indications pour élaborer un questionnaire plus précis sur différents lieux de Toulouse où les sujets pourront décrire des trajets plus complexes. Nous espérons que ce nouveau questionnaire permettra d'identifier les points de repère les plus utilisés afin d'en extraire des caractéristiques communes.

Dans leur ensemble, ces résultats montrent que la navigation avec une matrice de résolution très limitée (15x18) est possible, à condition d'afficher clairement et simplement la structure de l'environnement. De plus, les sujets ont montré qu'ils étaient également capables de construire des représentations mentales des environnements qu'ils avaient traversés. Les rendus utilisés pour parvenir à de telles performances partageaient différentes caractéristiques. Tout d'abord que ce soit sur les contrastes des surfaces ou la mise en valeur des arêtes, ces rendus représentaient distinctement la structure à tout instant. La faible résolution des rendus obtenus en VPS impose de sélectionner les informations à afficher si l'on souhaite éviter d'encombrer l'affichage. Les rendus utilisés affichaient une quantité d'information assez faible tout en permettant d'observer les éléments distants.

Chapitre X.

Limitations des différents résultats observés

Chapitre X. Limitations des différents résultats...

A. L'évaluation de la navigation en VPS

Il s'avère que pour la mobilité en VPS, différentes méthodes d'évaluation ont déjà été utilisées. Les premières que l'on retrouve le plus fréquemment proviennent de l'étude de la mobilité des malvoyants et non-voyants, et reposent sur la mesure de la vitesse de marche adoptée par les sujets appelée « *Preferred Walking Speed* » [Ralston 1958] ainsi que sur le pourcentage de cette vitesse dénommé « *Percentage of Preferred Walking Speed* » [Clark-Carter et al. 1986]. Les sujets ont comme consigne de traverser un couloir ou un environnement simple contenant ou non des obstacles et leurs déplacements sont chronométrés sur une dizaine de mètres. De cette manière, la rapidité avec laquelle se déplacent les sujets est mesurée ce qui permet de comparer l'efficacité de différents rendus prothétiques. Cette évaluation porte exclusivement sur la mobilité (évitement d'obstacles) et ne concerne pas l'orientation. Elle ne semble pas adaptée pour nos travaux et nous ne l'avons donc pas utilisée. La deuxième méthode un peu moins utilisée dans la littérature consiste à donner aux sujets des instructions à suivre pour naviguer. Les sujets doivent alors parcourir un trajet précis dans un environnement plus ou moins complexe [Rheede et al. 2010]. Leur temps de parcours est utilisé comme mesure de performance. Ayant reçu des instructions, ils peuvent également commettre des erreurs de navigation qui peuvent être comptées. Plus proche de la navigation, nous avons en partie utilisé cette méthode d'évaluation dans la première expérimentation. Avec cette méthode, les sujets obtiennent des informations sur l'environnement dans lequel ils vont naviguer par les instructions qu'ils reçoivent. Ainsi la navigation n'est pas la seule source d'informations sur l'environnement. De plus, cette tâche peut être vue par les sujets comme une séquence d'actions à menées et n'incitent pas forcément à retenir la structure de l'environnement. Enfin, une dernière méthode consiste à faire parcourir aux sujets une séquence de salles connectées l'une après l'autre et de chronométrer le temps de parcours. Dans chaque salle, les sujets doivent trouver la suivante jusqu'à arriver à la dernière [Wang et al. 2008; Dagnelie et al. 2007]. Lors de cette dernière tâche, les sujets doivent parvenir à trouver leur chemin dans une succession de salles sans obstacles ni difficulté de mobilité particulière. Sur les études en VPS, ces expérimentations sont celles qui se rapprochent le plus de la tâche de navigation. Cependant, aucun choix dans la navigation ne s'impose aux sujets qui doivent uniquement trouver la salle suivante et s'y diriger. Il n'y a donc pas d'exploration de l'environnement et le trajet effectué est

également imposé. Il faut remarquer que la qualité des représentations mentales construites par les sujets lors de ces navigations n'a jamais été évaluée dans ces différentes études. Ces représentations nous intéressaient pour évaluer dans quelle mesure les sujets pouvaient mémoriser la structure des environnements, s'orienter au sein de ces environnements et y naviguer. De même, la difficulté ressentie par les sujets pour naviguer n'a jamais été évaluée dans la littérature. Or il nous paraissait important de comparer la difficulté ressentie par les sujets pour nos rendus proposés. Si un rendu permet de bonnes performances de navigation mais implique une grande charge cognitive, nous supposons qu'il y a peu de chance qu'il soit utilisé quotidiennement par une personne implantée. Enfin, dans les différentes études en VPS, les sujets n'exploraient jamais un environnement inconnu avec des choix de navigation possibles et n'étaient pas amenés à mémoriser la disposition des lieux. Dans nos expérimentations, nous avons essayé de mettre en place une évaluation portant sur la navigation et l'exploration. En particulier dans les deuxième et troisième expérimentations, nous souhaitons que les sujets naviguent activement et n'aient pas un trajet imposé.

Notre première expérimentation s'appuie sur une méthode déjà utilisée en VPS [Dowling et al. 2005]: les sujets sont invités à suivre des instructions précises et leur temps d'exécution est chronométré. Ensuite, ils dessinent sur un fond de carte le trajet qu'ils viennent d'effectuer. La tâche évaluée n'est pas une tâche de navigation libre mais plutôt une tâche de suivi d'itinéraire. Le trajet étant imposé, les sujets ne sont pas amenés à prendre des décisions pour optimiser leur trajet mais à suivre une route imposée le plus rapidement possible, sans forcément mémoriser la disposition des lieux ou le trajet lui-même. Lors de la restitution, certains sujets ont utilisé les instructions qu'ils avaient retenues pour dessiner le trajet sur le fond de carte. Ainsi cette méthode ne permet pas de distinguer la restitution des consignes de la restitution de ce qui a réellement été mémorisé lors de la navigation elle-même.

Pour évaluer la navigation et la construction de cartes cognitives [Sanchez et al. 2009] a mis au point un protocole d'apprentissage et d'évaluation qui a influencé la méthode utilisée dans les deux expériences suivantes. Nous avons tenté d'améliorer la méthode précédemment utilisée pour évaluer une capacité d'exploration et de navigation dans un environnement inconnu. La tâche choisie impose aux sujets d'explorer l'ensemble de l'environnement afin d'y retrouver des objets disséminés. Au cours de cette tâche, ils sont amenés à retourner au moins trois fois à leur point de départ. Cette obligation les incite à faire l'effort de se repérer dans l'environnement et donc d'en retenir la configuration. Pour les sujets les plus compétents, il était possible de repérer la localisation de certains objets qu'ils ne peuvent pas emporter (limite atteinte). Cela permettait d'anticiper leurs déplacements futurs et de mémoriser le plus court chemin pour y revenir. Comme aucun obstacle n'est présent dans l'environnement, les sujets se concentraient sur la navigation à grande échelle et non sur la mobilité. L'indice de performance que nous avons calculé mesure principalement des capacités de navigation à grande échelle. La production des dessins par les sujets mesure la capacité des sujets à mémoriser la carte de l'environnement et à la restituer. Cette méthode de restitution provient de la littérature portant sur la navigation [Billinghurst & Weghorst 1995; Coluccia et al. 2007] et évalue

la restitution des représentations construites lors de cette navigation. En partant d'une feuille blanche, les sujets sont amenés à se rappeler la structure des environnements pour la restituer sur papier. L'évaluation de ces dessins est effectuée par des juges qui ne connaissent ni l'expérience ni les hypothèses opérationnelles. Dans la troisième expérimentation, il est demandé aux juges d'évaluer les cartes produites par les sujets sans tenir compte des distances, ce qui a permis d'observer des différences entre plusieurs rendus de VPS. Pour améliorer encore un peu plus cette évaluation, il est possible de demander aux juges de fournir deux évaluations distinctes des dessins, l'une tenant compte des distances et l'autre non. Ceci permettrait d'affiner les résultats lorsque les dessins des sujets sont fortement déformés. Certains sujets produisant en effet des dessins topologiquement justes mais avec une précision faible dans les rapports de distance.

B. Validation de la Simulation de Vision Prothétique

La première problématique soulevée porte sur la validité des expérimentations effectuées en VPS. En effet, même si la simulation de la vision prothétique est construite en incluant des paramètres pour la rendre plus réaliste (dysfonctionnement des électrodes, bruit sur la position des électrodes, etc.), elle ne pourra jamais correspondre parfaitement à l'expérience perceptive des personnes implantées [Chen et al. 2009]. Ceci est dû au simple fait que les différents paramètres de simulation proviennent des descriptions faites par les sujets et que les perceptions elles-mêmes ne sont pas directement observables. De plus certaines déformations dues à la structure de la rétine elle-même restent à évaluer pour simuler correctement la vision prothétique [Fine & Boynton 2015]. Les rendus de VPS sont donc toujours à améliorer en tenant compte de toute la phénoménologie des phosphènes déjà rapportée et des recommandations des articles de cette thématique pour se rapprocher le plus possible de l'expérience perceptive des personnes implantées.

C. Techniques d'interaction utilisées (clavier et casque) pour naviguer dans un environnement virtuel

Pour la navigation elle-même, toutes les expériences menées se sont déroulées dans des environnements virtuels. Les sujets utilisaient les touches de clavier pour se déplacer et s'orienter dans l'environnement. Les mouvements de la tête se répercutaient sur le mouvement de la caméra et les sujets pouvaient également tourner grâce aux flèches du clavier. Ainsi seules les informations vestibulaires d'orientation de la tête pouvaient être utilisées. Les déplacements étant virtuels, les sujets ne pouvaient pas non plus utiliser les signaux kinesthésiques et proprioceptifs liés à la marche pour intégrer leurs déplacements. L'utilisation du corps lors de la marche favorise la mémorisation d'un trajet effectué par rapport à un joystick [Ruddle et al. 2011]. En particulier, Klatzky et collaborateurs ont montré l'importance du système vestibulaire pour une tâche de complétion de triangle. En partant d'un point de départ, les sujets devaient parcourir deux segments de droite physiquement, en se l'imagi-

nant à l'aide d'une description orale ou en regardant quelqu'un le faire. Leur tâche était alors d'évaluer l'orientation à prendre depuis la fin du trajet pour se diriger vers le point de départ. Ils ont trouvé que le fait de marcher physiquement pour cette tâche limitait les erreurs des sujets [Klatzky et al. 1998]. Cette étude montre que la marche dans son ensemble (information vestibulaire, kinesthésique, proprioceptive) a un impact sur l'orientation et la mémorisation d'un trajet. En revanche l'implication particulière des systèmes vestibulaire, kinesthésiques et proprioceptifs n'est pas encore établie [Montello & Waller 2004; Larrue et al. 2012]. Par exemple les informations vestibulaires liées à l'accélération ont un impact limité dans la mémorisation de la structure d'un environnement que l'on explore [Waller et al. 2003]. En vision prothétique simulée, une étude portait sur différentes techniques d'interaction dans une tâche de mobilité et aucune différence n'a été trouvée entre une simulation de marche, le couple clavier-souris ou l'utilisation d'une manette [Zapf et al. 2014].

Quoi qu'il en soit, il serait intéressant de répliquer nos expériences dans une situation immersive plus réaliste. Il est par exemple possible d'utiliser des tapis roulants uni-directionnels ou même omni-directionnels [Souman et al. 2011] qui permettent de faire marcher le sujet dans toutes les directions. Des alternatives intéressantes et moins coûteuses utilisent des surfaces glissantes [Avila & Bailey 2014]. Pour une revue de ces dispositifs voir [Terziman 2012]. De plus ces dispositifs pourraient permettre de placer des patients implantés dans un milieu contrôlé permettant de concevoir des situations d'apprentissage avant de naviguer effectivement en ville [Larrue 2011].

D. Joueurs et non-joueurs, interaction et rendus utilisés

Les joueurs se sont retrouvés favorisés par notre technique d'interaction impliquant l'utilisation des touches de clavier pour se déplacer dans les environnements virtuels. La plupart de nos joueurs avaient déjà navigué dans des environnements virtuels en 3D et étaient donc déjà familiarisés avec la technique d'interaction choisie. En revanche cette familiarité avec la technique d'interaction ne suffisait pas à expliquer toutes les différences entre joueurs et non-joueurs. En effet, lors de notre troisième expérimentation, l'avantage des joueurs n'était pas visible uniformément sur toutes les conditions. Les environnements virtuels peuvent expliquer un autre biais lié à la technique d'interaction. En plus de l'utilisation du clavier, les rendus en vision Irrlicht ou VPS ont sans doute favorisé les joueurs qui étaient déjà familiers de ce type de rendu. En particulier, certains rendus VPS évalués comme le rendu de distance ou en fil-de-fer sont des rendus qui se retrouvent régulièrement dans les jeux vidéo. Ainsi les joueurs se sont sans doute familiarisés plus vite que les non-joueurs sur ces rendus VPS. Pour limiter ces écarts, une solution serait d'effectuer un réel apprentissage contrôlé au début de l'expérience, les sujets devant parvenir à une performance minimale avant d'effectuer la tâche évaluée. Ce contrôle des performances permettra de combler en partie les difficultés rencontrées par les non-joueurs et devrait limiter les différences de performances entre ces deux catégories de sujets.

E. Variabilité inter-individuelle

Un autre problème soulevé par la majorité des expériences présentées ici est la variabilité présente parmi les sujets. En effet, sur la plupart des rendus prothétiques utilisés, les performances des sujets couvraient un spectre large, certains trouvant la tâche trop facile alors que d'autres étaient mis en grande difficulté. Cette variabilité interindividuelle peut influencer la signification des tests statistiques et masquer les effets que l'on cherche à évaluer dans les données recueillies. Pour limiter cette influence, différentes méthodes peuvent être utilisées. La première est d'effectuer un apprentissage centré sur la navigation pour aider les sujets en difficulté à se rapprocher des sujets à l'aise avec cette tâche. Le principal inconvénient de cette solution est qu'un temps d'apprentissage est nécessaire avant de pouvoir effectivement faire passer les sujets sur les différentes conditions de rendu. La deuxième méthode consiste à sélectionner les sujets qui passent l'expérimentation de manière à ce qu'ils aient des performances similaires. Dans ce cas il est possible de sélectionner uniquement des sujets très à l'aise avec la tâche, ou au contraire des sujets éprouvant plus de difficultés. L'avantage de cette méthode est de limiter l'apprentissage avant de passer à l'expérimentation. Cependant la sélection doit être prise en compte dans le temps de passation de l'expérimentation. Le principal inconvénient de cette méthode est que les résultats ne seront alors plus généralisables à l'ensemble de la population mais seront limités à une sous-catégorie de la population. Une troisième méthode intermédiaire consiste à diviser en plusieurs groupes de performances les sujets. Il sera ainsi possible de comparer les performances liées aux différents rendus pour chacun de ces groupes et de procéder également à des comparaisons entre groupes. Même si l'utilisation de cette méthode implique de faire passer un plus grand nombre de sujets pour ne pas perdre en puissance statistique, il reste tout de même un bon compromis permettant de conserver la diversité de la population tout en limitant les effets de cette variabilité.

F. Les résultats sont-ils applicables à des patients implantés ?

Les patients implantés sont aveugles ou malvoyants alors que les sujets utilisés dans le cadre de ces travaux sont voyants. Il est donc possible que les patients implantés n'aient pas les mêmes habitudes de navigation que les sujets voyants. Ces différences peuvent avoir comme conséquences que les résultats obtenus en VPS ne se retrouvent pas chez les sujets aveugles implantés. Pour répondre à cette remarque, nous pouvons signaler qu'actuellement les personnes implantées ont eu des pathologies qui leur ont fait perdre la vue mais ont eu une expérience visuelle normale jusqu'à l'âge adulte. Le fait de cibler des personnes ayant perdu la vue pour des raisons connues, permet de poser les implants sur des structures du système visuel qui ne sont pas atteintes. La question reste entière en ce qui concerne l'utilisation de tels implants chez des sujets aveugles congénitaux et donc sans expérience visuelle préalable à l'implantation. Pour le moment, les interventions sont lourdes et les gains fonctionnels assez faibles, implanter des personnes n'ayant jamais vu implique des risques encore trop importants par rapport aux gains fonctionnels. Le développement des neuroprothèses visuelles en est encore à un stade préliminaire. On peut rapprocher les évolutions des neuroprothèses

visuelles aux évolutions déjà connues des implants cochléaires. Ces implants étaient au départ destinés aux personnes ayant perdu l'ouïe et leur apport fonctionnel se limitait à une aide pour lire sur les lèvres [Brackmann 1976]. Petit à petit, les progrès faits sur ses implants ont été si importants que maintenant des enfants n'ayant jamais entendu peuvent être implantés avec un gain fonctionnel important [Cheng et al. 2000]. Dans le cadre des implants cochléaires, quatre à huit canaux de stimulations sont suffisants pour amener un progrès fonctionnel [Margalit et al. 2002]. Pour les neuroprothèses visuelles, le nombre d'électrodes nécessaire est beaucoup plus important pour arriver à des progrès visibles [Humayun et al. 2012]. De nombreux progrès sont encore à faire pour augmenter le nombre d'électrodes utilisables simultanément mais nous pouvons imaginer que si ces recherches se poursuivent et que le nombre de canaux de stimulations augmente significativement, nous pourrions obtenir des progrès fonctionnel chez les patients implantés. Cette évolution pourrait aller jusqu'à ce que des enfants nés aveugles puissent profiter d'une vision prothétique limitant leur handicap.

Chapitre XI.

Discussion générale et perspectives

Ces travaux ont permis de montrer qu'il était possible de naviguer en VPS de basse résolution à condition d'utiliser la vision par ordinateur. Les résultats obtenus dans ces expérimentations devraient être validés par des sujets non-voyants implantés en s'appuyant également sur la réalité virtuelle. Au-delà de la validation des rendus proposés, l'utilisation d'environnements virtuels pourrait permettre aux sujets implantés de se familiariser avec les rendus eux-mêmes avant de les utiliser en réel. Enfin, ces résultats nous permettent d'imaginer ce que pourrait donner l'utilisation quotidienne de ces implants dans le cadre de la navigation.

A. Utilisabilité des implants à basse résolution pour naviguer

Les personnes non-voyantes connaissent de grandes difficultés pour naviguer en autonomie, ce qui peut entraîner une grande dépendance vis-à-vis de leurs proches ou un isolement. Les non-voyants qui parviennent à se déplacer seuls ne le font que sur des trajets qu'ils connaissent très bien et ne vont pratiquement jamais explorer seuls des lieux qui leur sont inconnus. Les neuroprothèses visuelles actuelles et celles arrivant sur le marché à moyen terme ne permettront pas de restaurer une autonomie complète et immédiate chez les personnes non-voyantes. Leur résolution et le nombre de niveaux de gris discernables sont encore trop réduits pour permettre de telles avancées fonctionnelles [Stronks & Dagnelie 2014]. En terme de mobilité, les implants actuels permettent seulement de suivre difficilement une ligne blanche au sol et de se diriger vers une porte noire sur un mur blanc [Zhou et al. 2013]. Même si les gains fonctionnels dans le cadre de la navigation autonome sont encore faibles, ces expérimentations montrent qu'ils sont présents et qu'il est possible de les améliorer. Les travaux de cette thèse se sont focalisés sur des implants de basse résolution (implants disposant de moins de trois cents électrodes), implants en développement ou qui vont être commercialisés à moyen terme. Les personnes disposant de ces implants pourraient donc bénéficier de l'aide apportée par les différents rendus proposés. Nous pouvons tirer de ces résultats quelques grandes lignes pour optimiser le design de rendus de VPS destinés à la navigation. Ces remarques sont déduites des résultats des expériences menées durant ces travaux de thèse et ne concernent que des rendus de vision prothétique en faible résolution. Nous supposons que pour des matrices d'électrodes de taille plus importantes, les contraintes seront différentes et ces recommandations ne seront sans doute plus aussi adéquates. Il est à noter que ces propositions de design concernent en priorité la navigation à grande échelle et ne sont pas destinées à palier des difficultés liées à la mobilité. Ces rendus n'ont pas été évalués avec l'ambition de remplacer

le chien-guide ou la canne blanche mais bien dans l'objectif de permettre aux personnes implantées de se repérer plus facilement et de naviguer en autonomie. Les propositions déduites des expérimentations sont les suivantes :

- ◆ **Représenter distinctement la structure présente dans le champ caméra.** Toutes les expériences en VPS effectuées dans le cadre de cette thèse indiquent que même si la structure n'est représentée que sommairement, elle est capitale. Son absence amène à une perte rapide de repères et d'orientation, ainsi il vaut mieux choisir de représenter la structure qui offre une certaine continuité que des points de repères ponctuels.
- ◆ **Équilibrer le rapport entre la richesse de l'environnement et la taille de l'espace visible.** Il semble nécessaire de rendre l'espace présent dans le champ caméra avec une profondeur de champ étendue. Cependant, pour que les informations lointaines soient pertinentes et compréhensibles, il est souhaitable de limiter leur quantité. Ceci permet d'éviter une surcharge qui peut avoir pour conséquence une incompréhension de la scène de la part de l'utilisateur en vision prothétique.
- ◆ **Garantir une continuité.** La première expérimentation a montré que l'utilisation de points de repère identifiés par un code visuel n'était pas suffisante pour permettre aux sujets de se repérer pour suivre des instructions de navigation. Si seulement des éléments discrets sont représentés, la continuité du rendu n'est pas assurée. Sans cette continuité, il est plus difficile de comprendre ses rotations et déplacements dans l'environnement et il est difficile de conserver son orientation. Même si le rendu brut n'est pas compréhensible seul, il permet d'assurer une continuité dans le rendu ce qui permet de mieux comprendre ses rotations et déplacements dans l'environnement.

B. Sujets implantés et sujets voyants

Les sujets qui ont participé aux expériences dans ces travaux sont jeunes par rapport aux patients qui seront effectivement implantés dans les prochaines années. Nos sujets avaient dans leur grande majorité une vingtaine d'années alors que les personnes qui profiteront des implants auront certainement soixante ans ou plus. Les populations ciblées pour les implants sont essentiellement des personnes atteintes de rétinite pigmentaire et perdent la vision à partir de 40 ans. Or une différence d'âge implique une différence dans l'habileté spatiale [Driscoll et al. 2005]. Ainsi nous pouvons supposer que si nous avions fait passer des sujets ayant l'âge des personnes implantées actuellement, les performances des sujets seraient moindres. Cependant il est à noter que la mesure que nous avons faite de l'habileté spatiale de nos sujets ne semble pas avoir de relations avec les résultats obtenus. Effectivement, cette mesure donnée par le questionnaire de Santa Barbara [Hegarty et al. 2002] est une mesure d'habileté spatiale rapportée par les sujets sur leur sens de l'orientation dans leur vie quotidienne et non dans un environnement virtuel en utilisant un clavier. Plus précisément, nous pouvons supposer que des personnes plus âgées effectuant la même tâche que dans nos expérimentations obtiendraient des performances moindres. Ceci pourrait être dû à l'habileté spatiale mais sur-

tout à la technique d'interaction utilisée (clavier et casque) ainsi qu'aux rendus à apprendre. Nos sujets ne profitaient pas d'un apprentissage complet avant d'effectuer la tâche, il est certain que les patients implantés suivront un long apprentissage avant de pouvoir utiliser leur implant pour naviguer [Zhou et al. 2013]. Nous supposons qu'avec un apprentissage suffisamment long et faisant appel à la réalité virtuelle, certains sujets pourraient progresser dans leur navigation [Moffat 2009]. Plus particulièrement dans ces travaux, la résolution des matrices utilisées correspond aux implants de prochaine génération comme l'Argus III. Une fois que ces implants seront utilisés par des non-voyants, il peut être intéressant de leur faire réaliser la même tâche de navigation en environnement virtuel. Ceci permettrait d'évaluer les possibilités réelles offertes par les différents rendus qui ne sont testés qu'avec la VPS pour le moment. Si les performances des sujets implantés avec un apprentissage suffisant sont proches de celles des sujets voyants effectuant la même expérience en VPS, les rendus testés seraient validés et pourraient être évalués dans des conditions de plus en plus réalistes. Pour gagner en réalisme et se rapprocher de leur quotidien, le nombre de repères dans les environnements peut être augmenté et des obstacles peuvent être ajoutés. Ceci amènera à combiner des difficultés d'orientation et de mobilité pour se rapprocher de leurs conditions de navigation. À l'inverse, si les patients implantés ne parviennent pas à se rapprocher des performances des sujets voyants en VPS même avec un long apprentissage, ceci peut signifier que le rendu en question ne peut être utilisé directement en condition réelle. Dans tous les cas il semble fondamental que le retour des personnes implantées soit pris en compte lors de l'apprentissage pour permettre une optimisation du rendu.

C. Importance des algorithmes de vision embarquée et des nouvelles technologies

Comme vu en introduction de cette partie, le développement des implants est long et il est très difficile de prévoir quand arriveront les implants de très haute résolution [Chader et al. 2009]. Ces travaux ont montré la possibilité de naviguer avec des implants de faible résolution. Il est donc possible d'imaginer de quelle manière les personnes implantées navigueront avec ces implants de faible résolution.

Les développements techniques actuels permettent l'émergence de dispositifs capables d'analyser l'espace de plus en plus précisément et en temps réel. Le domaine des algorithmes dédiés à la reconnaissance de scènes est extrêmement dynamique. Leur utilisation la plus immédiate se fera dans le domaine automobile [Scharwächter & Franke 2015] mais il est probable qu'à moyen terme de la reconnaissance de scènes soient disponibles sur des outils mobiles [Cavigelli et al. 2015]. En utilisant des dispositifs comprenant de mieux en mieux la scène, il sera alors possible d'ajouter aux rendus spécifiques à la navigation des éléments visuels pour soutenir la mobilité. Le fait que ces outils puissent fonctionner en temps réel est primordial pour éviter un décalage temporel entre les différents mouvements des sujets et la perception de l'environnement qu'ils auront [Lambrey et al. 2002]. Ces dispositifs techniques permettent d'imaginer de très nombreuses façons de naviguer et de représenter le monde. L'intégration

dans les smartphones de certaines technologies permettant de capter des informations 3D (comme le capteur RealSense 3D de la société Intel) commence à apparaître.

Le projet Tango de Google¹ a comme objectif le développement d'un smartphone disposant d'un grand nombre de capteurs et de caméras. Ce boîtier équipé de détecteurs de distance devrait être capable de capter et mémoriser la disposition des lieux au fur et à mesure de la navigation. Une personne implantée pourrait donc utiliser un tel dispositif pour naviguer à l'intérieur des bâtiments. En utilisant en plus un système d'information géographique, le dispositif serait en mesure de connaître sa position dans le bâtiment et d'indiquer le trajet à la personne qui l'utilise. C'est d'ailleurs un des arguments mis en avant dans la présentation du projet². Avec un tel dispositif, il devient imaginable pour des personnes disposant d'un implant de basse résolution d'avoir un guide visuel leur permettant de suivre un trajet précis à l'intérieur d'un bâtiment. La direction à prendre pourrait être indiquée en permanence par un phosphène très clair guidant l'utilisateur.

En extérieur, l'utilisation de la localisation GPS en plus de dispositif gérant l'orientation de la caméra permettra aussi sans doute de grandes avancées pour les personnes implantées. De la même façon qu'en intérieur avec le projet Tango, il est possible de représenter un guide lumineux servant à parcourir un trajet prédéfini. Pour le moment, les personnes implantées parviennent difficilement à suivre des lignes blanches au sol [Humayun et al. 2012]. Cette tâche sera sans doute plus facile avec les implants de prochaine génération et pourrait être utilisée dans le cadre de la navigation. Plutôt que d'afficher un phosphène indiquant la route à suivre, il sera possible de créer artificiellement le rendu d'une ligne au sol qu'il suffira ensuite de suivre pour arriver à destination.

Les deux rendus précédents pourraient être utiles pour parcourir des trajets, un guide visuel à hauteur des yeux ou au niveau du sol peut également gêner les utilisateurs si le rendu sur lequel il vient se greffer contient déjà de nombreux éléments. Ainsi, une autre méthode pour indiquer des directions, serait de placer l'aide à la navigation en hauteur. Les sujets pourraient profiter d'un rendu facilitant la compréhension de la structure lors de la navigation et lever les yeux pour que l'indice de navigation apparaisse et indique alors la direction à suivre.

Avec les implants disposant de moins de cent électrodes, il paraît difficile de réserver certaines électrodes pour une fonction spécifique [Parikh et al. 2013]. En revanche la perte sera moins importante pour les implants de prochaine génération. Nous pouvons imaginer alors l'affichage d'une boussole virtuelle en utilisant les électrodes situées à aux bords de la matrice. Une méthode serait de stimuler trois électrodes simultanément sur ce cadre pour indiquer par exemple le nord. Ainsi si l'utilisateur est tourné vers l'ouest, trois électrodes situées à droite de la matrice pourraient lui indiquer qu'il doit tourner à droite pour faire de nouveau face au Nord.

Seule la tâche de navigation a été étudiée durant ces travaux et les rendus étudiés avaient comme objectif unique de permettre l'exploration et la navigation dans un environnement

1 <https://www.google.com/atap/project-tango/about-project-tango/> (Dernier accès le 16/09/2015)

2 <https://www.youtube.com/watch?v=Qe10ExwzCqk> (Dernier accès le 17/09/2015)

inconnu. Ces rendus ne sont sans doute pas exploitables directement pour d'autres tâches. Or l'utilisation de la vision par ordinateur a un potentiel dans de nombreuses tâches en VPS comme la reconnaissance de visage [Wang et al. 2014], de texte [Denis 2014], la saisie d'objets [Macé et al. 2015] etc. Lors de l'utilisation quotidienne de l'implant, les personnes équipées souhaiteront sans doute changer et modifier les rendus à volonté en fonction des tâches à effectuer et des objectifs qu'elles se donnent. De nombreux dispositifs d'interaction pourraient être proposés afin de changer les rendus en temps réel. Par exemple le prototype de Grégoire Denis [Denis 2014] propose un système utilisant un écran tactile mobile (smartphone, *smartwatch*, etc.) et permettant de changer de rendu visuel à la demande. La détection des différentes parties du corps depuis une *smartwatch* permet d'interagir avec les rendus sans forcément utiliser de support [Lv et al. 2014]. Les utilisateurs pourraient sélectionner le rendu qu'ils souhaitent à un temps donné en effectuant des gestes simples devant eux et en gardant les deux mains libres. La résolution des implants actuels limite peut-être encore un peu trop les rendus possibles pour voir apparaître un gain fonctionnel. En revanche, lorsque les implants de prochaine génération seront utilisés, de nombreux outils et rendus seront déjà disponibles pour aider les personnes implantées à effectuer différentes tâches quotidiennes dont la navigation. Cette génération pourrait réaliser un important gain fonctionnel dans la vie quotidienne des personnes implantées.

Bibliographie

Ahuja A.K., Dorn J.D., Caspi A., McMahon M.J., Dagnelie G., da Cruz L., Stanga P.E., Humayun M.S. & Greenberg R.J. 2011. "Blind Subjects Implanted with the Argus II Retinal Prosthesis Are Able to Improve Performance in a Spatial-Motor Task." *The British Journal of Ophthalmology* 95 (4) (April 1): 539–43. doi:10.1136/bjo.2010.179622.

Allen G.L., ed. 2004. *Human Spatial Memory: Remembering Where*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Arno P., Capelle C., Wanet-Defalque M.C., Catalan-Ahumada M. & Veraart C. 1999. "Auditory Coding of Visual Patterns for the Blind." *Perception* 28 (8) (January): 1013–29.

Arno P., De Volder A.G., Vanlierde A., Wanet-Defalque M.-C., Streeel E., Robert A., Sanabria-Bohórquez S. & Veraart C. 2001. "Occipital Activation by Pattern Recognition in the Early Blind Using Auditory Substitution for Vision." *NeuroImage* 13 (4) (April): 632–45. doi:10.1006/nimg.2000.0731.

ATIH. 2015. "Classification Statistique Internationale Des Maladies et Des Problèmes de Santé Connexes." Paris, France.

Avila L. & Bailey M. 2014. "Virtual Reality for the Masses." *IEEE Computer Graphics and Applications*: 103–104.

Ayton L.N., Blamey P.J., Guymer R.H., Luu C.D., Nayagam D. a. X., Sinclair N.C., Shivdasani M.N., et al. 2014. "First-in-Human Trial of a Novel Suprachoroidal Retinal Prosthesis." *PLoS ONE* 9: e115239. doi:10.1371/journal.pone.0115239.

Bach-y-Rita P., Collins C.C., Saunders F.A., White B. & Scadden L. 1969. "Vision Substitution by Tactile Image Projection." *Nature* 50 (5184) (March 8): 83–91. doi:10.1038/221963a0.

Bach-y-Rita P., Kaczmarek K.A., Tyler M.E. & Garcia-Lara J. 1998. "Form Perception with a 49-Point Electrotactile Stimulus Array on the Tongue: A Technical Note." *Journal of Rehabilitation Research and Development* 35 (4) (October): 427–30.

Bak M., Girvin J.P., Hambrecht F.T., Kufra C. V., Loeb G.E. & Schmidt E.M. 1990. "Visual Sensations Produced by Intracortical Microstimulation of the Human Occipital Cortex." *Medical and Biological Engineering and Computing* 28 (May): 257–259. doi:10.1007/BF02442682.

Barnes N., Lieby P., Dennet H., Walker J., McCarthy C., Liu N. & Li Y. 2011. "Investigating the Role of Single-Viewpoint Depth Data in Visually-Guided Mobility." *Journal of Vision* 11 (11) (September 23): 926–926. doi:10.1167/11.11.926.

Barry M.P. & Dagnelie G. 2012. "Use of the Argus II Retinal Prosthesis to Improve Visual Guidance of Fine Hand Movements." *Investigative Ophthalmology & Visual Science* 53 (9) (August 1): 5095–101. doi:10.1167/iovs.12-9536.

Berthoz A. & Viaud-Delmon I. 1999. "Multisensory Integration in Spatial Orientation." *Current Opinion in Neurobiology* 9 (6): 708–712. doi:S0959-4388(99)00041-0 [pii].

- Billinghamurst M. & Weghorst S. 1995. "The Use of Sketch Maps to Measure Cognitive Maps of Virtual Environments." *Proceedings of the Virtual Reality Annual International Symposium '95*: 40–47. doi:10.1109/VRAIS.1995.512478.
- Borich G.D. & Bauman P.M. 1972. "Convergent and Discriminant Validation of the French and Guilford-Zimmerman Spatial Orientation and Spatial Visualization Factors." In *American Psychological Association, 80th Annual Meeting*. Honolulu, Hawaii.
- Boyle J.R., Maeder A. & Boles W. 2002. "Image Enhancement for Electronic Visual Prostheses." *Australasian Physics & Engineering Sciences in Medicine* 25 (2) (June): 81–86. doi:10.1007/BF03178470.
- Brackmann D.E. 1976. "The Cochlear Implant: Basic Principles." *The Laryngoscope* 86 (3): 373–388.
- Bradley D.C., Troyk P.R., Berg J.A., Bak M.J., Cogan S.F., Erickson R., Kufta C. V, et al. 2005. "Visuotopic Mapping through a Multichannel Stimulating Implant in Primate V1." *Journal of Neurophysiology* 93 (3) (March): 1659–70. doi:10.1152/jn.01213.2003.
- Brelén M.E., Vince V., Gérard B., Veraart C. & Delbeke J. 2010. "Measurement of Evoked Potentials after Electrical Stimulation of the Human Optic Nerve." *Investigative Ophthalmology & Visual Science* 51 (10): 5351–5355. doi:10.1167/iovs.09-4346.
- Brelén M.Å.E., De Potter P., Gersdorff M., Cosnard G., Veraart C. & Delbeke J. 2006. "Intraorbital Implantation of a Stimulating Electrode for an Optic Nerve Visual Prosthesis." *Journal of Neurosurgery* 104: 593–597.
- Brelén M.Å.E., Duret F., Gérard B., Delbeke J. & Veraart C. 2005. "Creating a Meaningful Visual Perception in Blind Volunteers by Optic Nerve Stimulation." *Journal of Neural Engineering* 2 (1) (March): S22–8. doi:10.1088/1741-2560/2/1/004.
- Brindley G.S. & Lewin W.S. 1968. "The Sensations Produced by Electrical Stimulation of the Visual Cortex." *The Journal of Physiology* 196 (2) (May): 479–93.
- Buchs G., Maidenbaum S. & Amedi A. 2015. "Augmented Non-Visual Distance Sensing With the EyeCane." In *AH '15 Proceedings of the 6th Augmented Human International Conference*, 209–210.
- Cai S., Fu L., Zhang H., Hu G. & Liang Z. 2005. "Prosthetic Visual Acuity in Irregular Phosphene Arrays under Two down-Sampling Schemes: A Simulation Study." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 5 (January): 5223–6. doi:10.1109/IEMBS.2005.1615656.
- Capelle C., Trullemans C., Arno P. & Veraart C. 1998. "A Real-Time Experimental Prototype for Enhancement of Vision Rehabilitation Using Auditory Substitution." *IEEE Transactions on Biomedical Engineering* 45 (10): 1279–1293. doi:10.1109/10.720206.
- Caspi A., Dorn J.D., McClure K.H., Humayun M.S., Greenberg R.J. & McMahon M.J. 2009. "Feasibility Study of a Retinal Prosthesis." *Archives of Ophthalmology* 127 (4): 398–401.
- Caspi A. & Zivotofsky A.Z. 2015. "Assessing the Utility of Visual Acuity Measures in Visual Prostheses." *Vision Research* 108: 77–84. doi:10.1016/j.visres.2015.01.006.

- Cavigelli L., Magno M. & Benini L. 2015. "Accelerating Real-Time Embedded Scene Labeling with Convolutional Networks." In *Proceedings of the 52nd Annual Design Automation Conference on - DAC '15*, 1–6. New York, New York, USA: ACM Press. doi:10.1145/2744769.2744788.
- Cegarra J. & Morgado N. 2009. "Étude Des Propriétés de La Version Francophone Du NASA-TLX." In *EPIQUE 2009 : 5ème Colloque de Psychologie Ergonomique*, 233–239.
- Cha K., Horch K.W. & Normann R.A. 1992a. "Simulation of a Phosphene-Based Visual Field: Visual Acuity in a Pixelized Vision System." *Annals of Biomedical Engineering* 20 (4) (January): 439–49.
- Cha K., Horch K.W. & Normann R.A. 1992b. "Mobility Performance with a Pixelized Vision System." *Vision Research* 32 (7) (July): 1367–72.
- Cha K., Horch K.W., Normann R.A. & Boman D.K. 1992c. "Reading Speed with a Pixelized Vision System." *Journal of the Optical Society of America. A, Optics and Image Science* 9 (5) (May): 673–7.
- Chader G.J., Weiland J.D. & Humayun M.S. 2009. "Artificial Vision: Needs, Functioning, and Testing of a Retinal Electronic Prosthesis." *Progress in Brain Research* 175 (January): 317–32. doi:10.1016/S0079-6123(09)17522-2.
- Chai X., Li L., Wu K., Zhou C., Cao P. & Ren Q. 2008. "C-Sight Visual Prostheses for the Blind." *IEEE Engineering in Medicine and Biology Magazine : The Quarterly Magazine of the Engineering in Medicine & Biology Society* 27 (5): 20–8. doi:10.1109/MEMB.2008.923959.
- Chai X., Yu W., Wang J., Zhao Y., Cai C. & Ren Q. 2007. "Recognition of Pixelized Chinese Characters Using Simulated Prosthetic Vision." *Artificial Organs* 31 (3) (March): 175–82. doi:10.1111/j.1525-1594.2007.00362.x.
- Chang M.H., Kim H.S., Shin J.H. & Park K.S. 2012. "Facial Identification in Very Low-Resolution Images Simulating Prosthetic Vision." *Journal of Neural Engineering* 9 (4) (August): 046012. doi:10.1088/1741-2560/9/4/046012.
- Chebat D.-R., Schneider F.C., Kupers R. & Ptito M. 2011. "Navigation with a Sensory Substitution Device in Congenitally Blind Individuals." *Neuroreport* 22 (7) (May 11): 342–7. doi:10.1097/WNR.0b013e3283462def.
- Chen S.C., Hallum L.E., Lovell N.H. & Suaning G.J. 2005. "Visual Acuity Measurement of Prosthetic Vision: A Virtual-Reality Simulation Study." *Journal of Neural Engineering* 2 (1) (March): S135–45. doi:10.1088/1741-2560/2/1/015.
- Chen S.C., Hallum L.E., Suaning G.J. & Lovell N.H. 2006. "Psychophysics of Prosthetic Vision: I. Visual Scanning and Visual Acuity." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 1 (January): 4400–3. doi:10.1109/IEMBS.2006.260816.
- Chen S.C., Hallum L.E., Suaning G.J. & Lovell N.H. 2007. "A Quantitative Analysis of Head Movement Behaviour during Visual Acuity Assessment under Prosthetic Vision Simulation." *Journal of Neural Engineering* 4 (1) (March): S108–23. doi:10.1088/1741-2560/4/1/S13.

- Chen S.C., Lovell N.H. & Suaning G.J. 2004. "Effect on Prosthetic Vision Visual Acuity by Filtering Schemes, Filter Cut-off Frequency and Phosphene Matrix : A Virtual Reality Simulation." In *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, 4201–4204.
- Chen S.C., Suaning G.J., Morley J.W. & Lovell N.H. 2009. "Simulating Prosthetic Vision: I. Visual Models of Phosphenes." *Vision Research* 49 (12) (June): 1493–1506. doi:10.1016/j.visres.2009.02.003.
- Cheng a K., Rubin H.R., Powe N.R., Mellon N.K., Francis H.W. & Niparko J.K. 2000. "Cost-Utility Analysis of the Cochlear Implant in Children." *JAMA : The Journal of the American Medical Association* 284 (7): 850–856. doi:10.1001/jama.284.7.850.
- Choset H. & Nagatani K. 2001. "Topological Simultaneous Localization and Mapping (SLAM): Toward Exact Localization without Explicit Localization." *IEEE Transactions on Robotics and Automation* 17 (2) (April): 125–137. doi:10.1109/70.928558.
- Chow A. 2013. "Retinal Prostheses Development in Retinitis Pigmentosa Patients—Progress and Comparison." *The Asia-Pacific Journal of Ophthalmology* 2 (4): 253–268. doi:10.1097/APO.0b013e3182a0b4fe.
- Chow A.Y. & Chow V.Y. 1997. "Subretinal Electrical Stimulation of the Rabbit Retina." *Neuroscience Letters* 225 (1): 13–16. doi:10.1016/S0304-3940(97)00185-7.
- Chow A.Y., Chow V.Y., Packo K.H., Pollack J.S., Peyman G. a & Schuchard R. 2004. "The Artificial Silicon Retina Microchip for the Treatment of Vision Loss from Retinitis Pigmentosa." *Archives of Ophthalmology* 122 (4): 460–469. doi:10.1001/archopht.122.4.460.
- Clark-Carter D.D., Heyes a D. & Howarth C.I. 1986. "The Efficiency and Walking Speed of Visually Impaired People." *Ergonomics* 29 (6): 779–789. doi:10.1080/00140138608968314.
- Collignon O., Lassonde M., Lepore F., Bastien D. & Veraart C. 2006. "Functional Cerebral Reorganization for Auditory Spatial Processing and Auditory Substitution of Vision in Early Blind Subjects." *Cerebral Cortex* 17 (2): 457–465. doi:10.1093/cercor/bhj162.
- Coluccia E., Iosue G. & Brandimonte M.A. 2007. "The Relationship between Map Drawing and Spatial Orientation Abilities: A Study of Gender Differences." *Journal of Environmental Psychology* 27 (2) (June): 135–144. doi:10.1016/j.jenvp.2006.12.005.
- Cornoldi C., Tinti C., Mammarella I.C., Re A.M. & Varotto D. 2009. "Memory for an Imagined Pathway and Strategy Effects in Sighted and in Totally Congenitally Blind Individuals." *Acta Psychologica* 130 (1): 11–16. doi:10.1016/j.actpsy.2008.09.012.
- Coulombe J., Sawan M. & Gervais J.-F. 2007. "A Highly Flexible System for Microstimulation of the Visual Cortex: Design and Implementation." *IEEE Transactions on Biomedical Circuits and Systems* 1 (4) (December): 258–69. doi:10.1109/TBCAS.2007.916026.
- da Cruz L., Coley B.F., Dorn J.D., Merlini F., Filley E., Christopher P., Chen F.K., et al. 2013. "The Argus II Epiretinal Prosthesis System Allows Letter and Word Reading and Long-Term Function in Patients with Profound Vision Loss." *The British Journal of Ophthalmology* 97 (5) (May): 632–6. doi:10.1136/bjophthalmol-2012-301525.

Dagnelie G., Barnett D.G., Humayun M.S. & Thompson R.W. 2006a. "Paragraph Text Reading Using a Pixelized Prosthetic Vision Simulator: Parameter Dependence and Task Learning in Free-Viewing Conditions." *Investigative Ophthalmology & Visual Science* 47 (3) (March): 1241–50. doi:10.1167/iovs.05-0157.

Dagnelie G., Keane P., Narla V., Yang L., Weiland J.D. & Humayun M.S. 2007. "Real and Virtual Mobility Performance in Simulated Prosthetic Vision." *Journal of Neural Engineering* 4 (1) (March): S92–101. doi:10.1088/1741-2560/4/1/S11.

Dagnelie G., Thompson R., Barnett D.G. & Zhang W. 2000. "Visual Perception and Performance under Conditions Simulating Prosthetic Vision." *Perception* 29.

Dagnelie G., Thompson R.W., Barnett D.G. & Zhang W. 2001. "Simulated Prosthetic Vision: Perceptual and Performance Measures." *Vision Science and Its ...*

Dagnelie G., Walter M. & Yang L. 2006b. "Playing Checkers: Detection and Eye–hand Coordination in Simulated Prosthetic Vision." *Journal of Modern Optics* 53 (9) (June 15): 1325–1342. doi:10.1080/09500340600619197.

Dai C., Lu M., Zhao Y., Lu Y., Zhou C., Chen Y., Ren Q. & Chai X. 2010. "Correction for Chinese Character Patterns Formed by Simulated Irregular Phosphene Map." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2010 (January): 5887–90. doi:10.1109/IEMBS.2010.5627528.

De Balthasar C., Patel S., Roy A., Freda R., Greenwald S.H., Horsager A., Mahadevappa M., et al. 2008. "Factors Affecting Perceptual Thresholds in Epiretinal Prostheses." *Investigative Ophthalmology & Visual Science* 49 (6): 2303–2314.

Dejos M. 2012. "Approche Ecologique de L'évaluation de La Memoire Episodique et de La Navigation Spatiale Dans La Maladie D'alzheimer." Bordeaux 2 Segalen.

Delbeke J., Oozeer M. & Veraart C. 2003. "Position, Size and Luminosity of Phosphenes Generated by Direct Optic Nerve Stimulation." *Vision Research* 43 (9) (April): 1091–1102. doi:10.1016/S0042-6989(03)00013-0.

Den Brinker B. 2003. "EBU Policy Statement on Low Vision." Amsterdam.

Denis G. 2014. "Apport de La Vision Par Ordinateur Dans L'utilisabilité Des Neuroprothèses Visuelles."

Denis G., Jouffrais C., Vergniew V. & Macé M.J.-M. 2013. "Human Faces Detection and Localization with Simulated Prosthetic Vision." In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 61–66.

Denis G., Macé M.J.-M. & Jouffrais C. 2012. "Simulated Prosthetic Vision: Object Recognition and Localization Approach." In *Proceedings of the 4th International Conference on Neuroprosthetic Devices (ICNPD 2012)*, 40–41.

Denis M., Pazzaglia F., Cornoldi C. & Bertolo L. 1999. "Spatial Discourse and Navigation: An Analysis of Route Directions in the City of Venice." *Applied Cognitive Psychology* 13 (2): 145–174. doi:10.1002/(SICI)1099-0720(199904)13:2<145::AID-ACP550>3.0.CO;2-4.

Dobelle W.H. 2000. "Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex." *ASAIO Journal (American Society for Artificial Internal Organs)* 46 (1): 3–9.

Dobelle W.H. & Mladejovsky M.G. 1974. "Phosphenes Produced by Electrical Stimulation of Human Occipital Cortex, and Their Application to the Development of a Prosthesis for the Blind." *The Journal of Physiology* 243 (2) (December): 553–76.

Dobelle W.H., Mladejovsky M.G. & Girvin J.P. 1974. "Artificial Vision for the Blind: Electrical Stimulation of Visual Cortex Offers Hope for a Functional Prosthesis." *Science* 183 (4123) (February 1): 440–444. doi:10.1126/science.183.4123.440.

Dorn J.D., Ahuja A.K., Caspi A., da Cruz L., Dagnelie G., Sahel J.-A., Greenberg R.J. & McMahon M.J. 2012. "The Detection of Motion by Blind Subjects With the Epiretinal 60-Electrode (Argus II) Retinal Prosthesis." *Archives of Ophthalmology* (October 8): 1–7. doi:10.1001/2013.jamaophthalmol.221.

Dowling J. 2007. "Mobility Enhancement Using Simulated Artificial Human Vision." Queensland University of Technology.

Dowling J.A., Boles W. & Maeder A. 2005. "Mobility Assessment Using Simulated Artificial Human Vision." In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, 3:32–32. IEEE. doi:10.1109/CVPR.2005.494.

Dowling J.A., Boles W.W. & Maeder A.J. 2006. "Simulated Artificial Human Vision : The Effects of Spatial Resolution and Frame Rate on Mobility." In *Proceedings Active Media Technology*, edited by Yuefeng Li, Mark Looi, and Ning Zhong, 138–143. Brisbane.

Dowling J.A., Maeder A. & Boles W. 2004. "Mobility Enhancement and Assessment for a Visual Prosthesis." In *Medical Imaging*, edited by Amir A. Amini and Armando Manduca, 780–791. doi:10.1117/12.534866.

Dramas F., Thorpe S.J. & Jouffrais C. 2010. "Artificial Vision For The Blind: A Bio-Inspired Algorithm For Objects And Obstacles Detection." *International Journal of Image and Graphics* 10 (4): 531–544.

Driscoll I., Hamilton D.A., Yeo R.A., Brooks W.M. & Sutherland R.J. 2005. "Virtual Navigation in Humans: The Impact of Age, Sex, and Hormones on Place Learning." *Hormones and Behavior* 47 (3): 326–335. doi:10.1016/j.yhbeh.2004.11.013.

Ekman P. 1993. "Facial Expression and Emotion." *American Psychologist* 48 (4): 384–392. doi:10.1037/0003-066x.48.4.384.

Fang X., Sakaguchi H., Fujikado T., Osanai M., Kanda H., Ikuno Y., Kamei M., et al. 2004. "Direct Stimulation of Optic Nerve by Electrodes Implanted in Optic Disc of Rabbit Eyes." *Graefe's Archive for Clinical and Experimental Ophthalmology* 243: 49–56. doi:10.1007/s00417-004-0957-0.

Fernandez E., Greger B., House P. a., Aranda I., Botella C., Albusua J., Soto-SÁnchez C., Alfaro A. & Normann R. a. 2014. "Acute Human Brain Responses to Intracortical Microelectrode Arrays: Challenges and Future Prospects." *Frontiers in Neuroengineering* 7 (July): 24. doi:10.3389/fneng.2014.00024.

- Fine I. & Boynton G.M. 2015. "Pulse Trains to Percepts: The Challenge of Creating a Perceptually Intelligible World with Sight Recovery Technologies." *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1677) (September 19): 20140208. doi:10.1098/rstb.2014.0208.
- Fletcher J.F. 1980. "Spatial Representation in Blind Children. 2: Development Compared to Sighted Children." *Journal of Visual Impairment and Blindness* 74 (10): 381–85.
- Foerster O. 1929. "Beitrage Zur Pathophysiologie Der Sehbahn Und Der Sehspahre." *J. Psychol. Neurol.* 39: 463–485.
- Fortin M., Voss P., Lord C., Lasseonde M., Pruessner J., Saint-Amour D., Rainville C. & Lepore F. 2008. "Wayfinding in the Blind: Larger Hippocampal Volume and Supranormal Spatial Navigation." *Brain* 131 (11) (June 21): 2995–3005. doi:10.1093/brain/awn250.
- Fu L., Cai S., Zhang H., Hu G. & Zhang X. 2006. "Psychophysics of Reading with a Limited Number of Pixels: Towards the Rehabilitation of Reading Ability with Visual Prosthesis." *Vision Research* 46 (8-9) (April): 1292–301. doi:10.1016/j.visres.2005.11.011.
- Giudice N. a. & Legge G.E. 2008. "Blind Navigation and the Role of Technology." In *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence*, 479–500. doi:10.1002/9780470379424.ch25.
- Golledge R.G. 1992. "Place Recognition and Wayfinding: Making Sense of Space." *Geoforum* 23 (2): 199–214.
- Grumet A.E., Wyatt J.L. & Rizzo J.F. 2000. "Multi-Electrode Stimulation and Recording in the Isolated Retina." *Journal of Neuroscience Methods* 101: 31–42. doi:10.1016/S0165-0270(00)00246-6.
- Guilford J.P. & W.S. Z. 1956. *Guilford-Zimmerman Aptitude Survey*. Beverly Hills: Sheridan Supply Company.
- Guo H., Qin R., Qiu Y., Zhu Y. & Tong S. 2010. "Configuration-Based Processing of Phosphene Pattern Recognition for Simulated Prosthetic Vision." *Artificial Organs* 34 (4) (April): 324–30. doi:10.1111/j.1525-1594.2009.00863.x.
- Hart S.G. & Staveland L.E. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Advances in Psychology* 52. *Advances in Psychology*: 139–183. doi:10.1016/S0166-4115(08)62386-9.
- Hayes J.S., Yin V.T., Piyathaisere D. V, Weiland J.D., Humayun M.S. & Dagnelie G. 2003. "Visually Guided Performance of Simple Tasks Using Simulated Prosthetic Vision." *Artificial Organs* 27 (11) (November): 1016–28.
- Hegarty M., Montello D.R., Richardson A.E., Ishikawa T. & Lovelace K. 2006. "Spatial Abilities at Different Scales: Individual Differences in Aptitude-Test Performance and Spatial-Layout Learning." *Intelligence* 34 (2) (March): 151–176. doi:10.1016/j.intell.2005.09.005.
- Hegarty M., Richardson A.E., Montello D.R., Lovelace K. & Subbiah I. 2002. "Development of a Self-Report Measure of Environmental Spatial Ability." *Intelligence* 30: 425–447. doi:10.1016/S0160-2896(02)00116-2.
- Hornig R., Zehnder T. & Velikay-Parel M. 2007. "The IMI Retinal Implant System." In *Artificial Sight*, 111–128. doi:10.1007/978-0-387-49331-2_6.

- Hu J., Xia P., Gu C., Qi J., Li S. & Peng Y. 2013. "Recognition of Similar Objects Using Simulated Prosthetic Vision." *Artificial Organs* (September 22): n/a–n/a. doi:10.1111/aor.12147.
- Humayun M.S., Dorn J.D., Ahuja A.K., Caspi A., Filley E., Dagnelie G., Salzmann J., et al. 2009. "Preliminary 6 Month Results from the Argus II Epiretinal Prosthesis Feasibility Study." In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009:4566–8. doi:10.1109/IEMBS.2009.5332695.
- Humayun M.S., Dorn J.D., da Cruz L., Dagnelie G., Sahel J.-A., Stanga P.E., Cideciyan A. V, et al. 2012. "Interim Results from the International Trial of Second Sight's Visual Prosthesis." *Ophthalmology* 119 (4) (May): 779–88. doi:10.1016/j.ophtha.2011.09.028.
- Humayun M.S., Weiland J.D., Fujii G.Y., Greenberg R.J., Williamson R., Little J., Mech B. V, et al. 2003. "Visual Perception in a Blind Subject with a Chronic Microelectronic Retinal Prosthesis." *Vision Research* 43 (24): 2573–2581.
- Iachini T. & Ruggiero G. 2010. "The Role of Visual Experience in Mental Scanning of Actual Pathways: Evidence from Blind and Sighted People." *Perception* 39 (7): 953–969. doi:10.1068/p6457.
- Ishikawa T. & Montello D.R. 2006. "Spatial Knowledge Acquisition from Direct Experience in the Environment: Individual Differences in the Development of Metric Knowledge and the Integration of Separately Learned Places." *Cognitive Psychology* 52 (2): 93–129. doi:10.1016/j.cogpsych.2005.08.003.
- Jouffrais C. 2011. "Les Nouvelles Technologies Au Service de La Cognition Spatiale Des Déficients Visuels." Université of Toulouse.
- Ju J.S.S., Ko E. & Kim E.Y.Y. 2009. "EYECane: Navigating with Camera Embedded White Cane for Visually Impaired Person." In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, 237–238. ACM.
- Kaczmarek K.A. 2011. "The Tongue Display Unit (TDU) for Electrotactile Spatiotemporal Pattern Presentation." *Scientia Iranica* 18 (6): 1476–1485.
- Kane S.R., Cogan S.F., Ehrlich J., Plante T.D., McCreery D.B. & Troyk P.R. 2013. "Electrical Performance of Penetrating Microelectrodes Chronically Implanted in Cat Cortex." *IEEE Transactions on Biomedical Engineering* 60 (8): 2153–2160. doi:10.1109/TBME.2013.2248152.
- Katz B.F.G., Dramas F., Parseihian G., Gutierrez O., Kammoun S., Brilhault A., Brunet L., et al. 2012. "NAVIG: Guidance System for the Visually Impaired Using Virtual Augmented Reality." *Technology and Disability* 24 (2): 163–178. doi:10.3233/TAD-2012-0344.
- Kelly S.K., Shire D.B., Chen J., Doyle P., Gingerich M.D., Drohan W.A., Theogarajan L.S., Cogan S.F., Wyatt J.L. & Rizzo J.F. 2009. "Realization of a 15-Channel, Hermetically-Encased Wireless Subretinal Prosthesis for the Blind." *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2009 (January): 200–3. doi:10.1109/IEMBS.2009.5333619.
- Kiral-Kornek F.I., O Sullivan-Greene E., Savage C.O., McCarthy C., Grayden D.B. & Burkitt A.N. 2014. "Improved Visual Performance in Letter Perception through Edge Orienta-

tion Encoding in a Retinal Prosthesis Simulation.” *Journal of Neural Engineering* 11: 066002. doi:10.1088/1741-2560/11/6/066002.

Kitiratschky V.B.D., Stingl K., Wilhelm B., Peters T., Besch D., Sachs H., Gekeler F., Bartz-Schmidt K.U. & Zrenner E. 2014. “Safety Evaluation of ‘retina Implant Alpha IMS’—a Prospective Clinical Trial.” *Graefe’s Archive for Clinical and Experimental Ophthalmology* 253 (3): 381–387. doi:10.1007/s00417-014-2797-x.

Klatzky R.L., Loomis J.M., Beal A.C., Chance S.S. & Golledge R.G. 1998. “Spatial Updating of Self-Position and Orientation during Real, Imagined, and Virtual Locomotion.” *Psychological Science* 9: 293–298.

Klauke S., Goertz M., Rein S., Hoehl D., Thomas U., Eckhorn R., Bremmer F. & Wachtler T. 2011. “Stimulation with a Wireless Intraocular Epiretinal Implant Elicits Visual Percepts in Blind Humans.” *Investigative Ophthalmology & Visual Science* 52 (1): 449–455.

Knapp M.L., Hall J.A. & Terrence G.H. 2013. *Nonverbal Communication in Human Interaction, Heighth Edition*. Cengage Le. Boston, MA: Monica Eckman.

Kotecha A., Zhong J., Stewart D. & da Cruz L. 2014. “The Argus II Prosthesis Facilitates Reaching and Grasping Tasks: A Case Series.” *BMC Ophthalmology* 14 (1): 71. doi:10.1186/1471-2415-14-71.

Krause F. & Schum H. 1931. “Die Epileptischen Erkrankungen.” *Neue Deutsche Chirurgie* 49a (April).

Kuipers B.J. 1977. “Representing Knowledge of Large-Scale Space.” Massachusetts Institute of Technology.

Lambrey S., Viaud-Delmon I. & Berthoz A. 2002. “Influence of a Sensorimotor Conflict on the Memorization of a Path Traveled in Virtual Reality.” *Cognitive Brain Research* 14 (1): 177–186. doi:10.1016/S0926-6410(02)00072-1.

Lane F.J., Huyck M., Troyk P. & Schug K. 2012. “Responses of Potential Users to the Intracortical Visual Prosthesis: Final Themes from the Analysis of Focus Group Data.” *Disability and Rehabilitation: Assistive Technology*. doi:10.3109/17483107.2011.629326.

Larrue F. 2011. “Influence Des Interfaces Dans Le Transfert Du Virtuel Au Réel.” Université de Bordeaux 2.

Larrue F., Sauzéon H., Aguilova L., Lotte F., Hachet M. & NKaoua B. 2012. “Brain Computer Interface vs Walking Interface in VR.” In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology - VRST ’12*, 113. New York, New York, USA: ACM Press. doi:10.1145/2407336.2407359.

Lee P.U. & Tversky B. 2005. “Interplay Between Visual and Spatial: The Effect of Landmark Descriptions on Comprehension of Route/Survey Spatial Descriptions.” *Spatial Cognition & Computation* 5 (2): 163–185. doi:10.1207/s15427633scc052&3_4.

Leroy C. 1755. “Mémoire Où L’on Rend Compte de Quelques Tentatives Que L’on a Faites Pour Guérir Plusieurs Maladies Par l’Électricité.” In *Histoire de l’Académie Royale Des Sciences*, edited by Académie des Sciences, 60–98. Paris.

Levelt W.J.M. 1982. "Linearization in Describing Spatial Networks." In *Processes, Beliefs, and Questions*, 199–220. Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-7668-0_7.

Lewis P.M., Ackland H.M., Arthur J. & Rosenfeld J. V. 2015. "Restoration of Vision in Blind Individuals Using Bionic Devices: A Review with a Focus on Cortical Visual Prostheses." *Brain Research* 1595: 51–73. doi:10.1016/j.brainres.2014.11.020.

Li S., Hu J., Chai X. & Peng Y. 2011. "Image Recognition with a Limited Number of Pixels for Visual Prostheses Design." *Artificial Organs* 36 (3) (September 29): 266–274. doi:10.1111/j.1525-1594.2011.01347.x.

Li W.-H., Tang T.J.-J. & Lui W.L.D. 2013. "Going beyond Vision to Improve Bionic Vision." In *20th IEEE International Conference on Image Processing (ICIP)*, 1555–1558. Melbourne, Australia.

Lieby P., Barnes N., McCarthy C., Liu N., Dennett H., Walker J.G., Botea V. & Scott A.F. 2011. "Substituting Depth for Intensity and Real-Time Phosphene Rendering: Visual Navigation under Low Vision Conditions." In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011:8017–20. doi:10.1109/IEMBS.2011.6091977.

Loomis J.M., Klatzky R.L., Golledge R.G., Cicinelli J.G., Pellegrino J.W. & Fry P. a. 1993. "Nonvisual Navigation by Blind and Sighted: Assessment of Path Integration Ability." *Journal of Experimental Psychology. General* 122 (1) (March): 73–91.

Lowery A.J. 2013. "Introducing the Monash Vision Group's Cortical Prosthesis." In *20th IEEE International Conference on Image Processing (ICIP)*, 1536–1539. Melbourne, Australia: IEEE.

Lu Y., Wang J., Wu H., Li L., Cao X. & Chai X. 2013. "Recognition of Objects in Simulated Irregular Phosphene Maps for an Epiretinal Prosthesis." *Artificial Organs*. doi:10.1111/aor.12174.

Lui W.L.D., Browne D., Kleeman L., Drummond T. & Li W.H. 2012. "Transformative Reality: Improving Bionic Vision with Robotic Sensing." *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2012* (January): 304–7. doi:10.1109/EMBC.2012.6345929.

Luo Y.H.-L., Zhong J.J. & da Cruz L. 2014. "The Use of Argus® II Retinal Prosthesis by Blind Subjects to Achieve Localisation and Prehension of Objects in 3-Dimensional Space." *Graefe's Archive for Clinical and Experimental Ophthalmology*. doi:10.1007/s00417-014-2912-z.

Lv Z., Feng L., Li H. & Feng S. 2014. "Hand-Free Motion Interaction on Google Glass." In *SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications on - SA '14*, 1–1. New York, New York, USA: ACM Press. doi:10.1145/2669062.2669066.

Lynch K. 1960. *The Image of the City*. Cambridge, UK: MIT Press.

Macé M.J.-M., Guivarch V., Denis G. & Jouffrais C. 2015. "Simulated Prosthetic Vision: The Benefits of Computer-Based Object Recognition and Localization." *Artificial Organs* (6): 12. doi:10.1111/aor.12476.

Mahadevappa M., Weiland J.D., Yanai D., Fine I., Greenberg R.J. & Humayun M.S. 2005. "Perceptual Thresholds and Electrode Impedance in Three Retinal Prosthesis Subjects." *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society* 13 (2) (June): 201–6. doi:10.1109/TNSRE.2005.848687.

Mandel Y., Goetz G., Lavinsky D., Huie P., Mathieson K., Wang L., Kamins T., et al. 2013. "Cortical Responses Elicited by Photovoltaic Subretinal Prostheses Exhibit Similarities to Visually Evoked Potentials." *Nature Communications* 4: 1980. doi:10.1038/ncomms2980.

Mathieson K., Loudin J., Goetz G., Huie P., Wang L., Kamins T.I., Galambos L., et al. 2012. "Photovoltaic Retinal Prosthesis with High Pixel Density." *Nature Photonics*. doi:10.1038/nphoton.2012.104.

McCarthy C. & Barnes N. 2012. "Time-to-Contact Maps for Navigation with a Low Resolution Visual Prosthesis." *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2012* (January): 2780–3. doi:10.1109/EMBC.2012.6346541.

McCarthy C., Barnes N. & Lieby P. 2011. "Ground Surface Segmentation for Navigation with a Low Resolution Visual Prosthesis." In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011:4457–60. doi:10.1109/IEMBS.2011.6091105.

McCarthy C., Walker J.G., Lieby P., Scott A. & Barnes N. 2015. "Mobility and Low Contrast Trip Hazard Avoidance Using Augmented Depth." *Journal of Neural Engineering* 12 (1) (November 26): 016003. doi:10.1088/1741-2560/12/1/016003.

Meijer P.B. 1992. "An Experimental System for Auditory Image Representations." *IEEE Transactions on Bio-Medical Engineering* 39 (2) (February): 112–21. doi:10.1109/10.121642.

Meilinger T. 2008a. "The Network of Reference Frames Theory : A Synthesis of Graphs and Cognitive Maps." *Spatial Cognition VI: Learning, Reasoning, and Talking about Space*: 344–360.

Meilinger T. 2008b. "Strategies of Orientation in Environmental Spaces." Max Planck Institute.

Menzel-Severing J., Laube T., Brockmann C., Bornfeld N., Mokwa W., Mazinani B., Walter P. & Roessler G. 2012. "Implantation and Explantation of an Active Epiretinal Visual Prosthesis: 2-Year Follow-up Data from the EPIRET3 Prospective Clinical Trial." *Eye (London, England)* 26 (4) (April): 501–9. doi:10.1038/eye.2012.35.

Moffat S.D. 2009. "Aging and Spatial Navigation: What Do We Know and Where Do We Go?" *Neuropsychology Review* 19 (4): 478–489. doi:10.1007/s11065-009-9120-3.

Mohammadi H.M., Ghafar-Zadeh E. & Sawan M. 2012. "An Image Processing Approach for Blind Mobility Facilitated through Visual Intracortical Stimulation." *Artificial Organs* 36 (7) (July): 616–28. doi:10.1111/j.1525-1594.2011.01421.x.

Mokwa W., Goertz M., Koch C., Krisch I., Trieu H.K. & Walter P. 2008. "Intraocular Epiretinal Prosthesis to Restore Vision in Blind Humans." *Conference Proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2008*: 5790–5793. doi:10.1109/IEMBS.2008.4650530.

Monash Vision Group. 2013. "Monash Vision Group, Annual Report." Clayton, Australia.

Money J., Walker H.T. & Alexander D. 1965. *A Standardized Road-Map Test of Direction Sense: Manual*. Johns Hopkins Press.

Montello D. & Waller D. 2004. "Spatial Memory of Real Environments, Virtual Environments, and Maps." In *Human Spatial Memory*, edited by Gary L Allen, 251–285. Lawrence Erlbaum Associates.

Montello D.R. 1993. "Scale and Multiple Psychologies of Space." In *Spatial Information Theory: A Theoretical Basis for GIS*, edited by Andrew U Frank and I. Campari, 312–321. Berlin: Springer.

Montello D.R. 1998. "A New Framework for Understanding the Acquisition of Spatial Knowledge in Large-Scale Environments." In *Spatial and Temporal Reasoning in Geographic Information Systems*, In M.J Ege, 143–154. New York: Oxford University Press.

Nishida K., Sakaguchi H., Kamei M., Cecilia-Gonzalez C., Terasawa Y., Velez-Montoya R., Fujikado T., et al. 2015. "Visual Sensation by Electrical Stimulation Using a New Direct Optic Nerve Electrode Device." *Brain Stimulation*. doi:10.1016/j.brs.2015.03.001.

Normann R.A. 2007. "Technology Insight: Future Neuroprosthetic Therapies for Disorders of the Nervous System." *Nature Clinical Practice. Neurology* 3 (8) (August): 444–52. doi:10.1038/ncpneuro0556.

Normann R.A., Greger B.A., House P.A., Romero S.F., Pelayo F. & Fernández E. 2009. "Toward the Development of a Cortically Based Visual Neuroprosthesis." *Journal of Neural Engineering* 6 (3) (June): 035001. doi:10.1088/1741-2560/6/3/035001.

Normann R.A., Maynard E.M., Rousche P.J. & Warren D.J. 1999. "A Neural Interface for a Cortical Vision Prosthesis." *Vision Research* 39 (15) (July): 2577–87.

Parikh N., Itti L., Humayun M.S. & Weiland J. 2013. "Performance of Visually Guided Tasks Using Simulated Prosthetic Vision and Saliency-Based Cues." *Journal of Neural Engineering* 10 (2) (April 1): 026017. doi:10.1088/1741-2560/10/2/026017.

Pérez Fornos A., Sommerhalder J., da Cruz L., Sahel J.A., Mohand-Said S., Hafezi F. & Pelizzone M. 2012. "Temporal Properties of Visual Perception on Electrical Stimulation of the Retina." *Investigative Ophthalmology & Visual Science* 53 (6) (May 4): 2720. doi:10.1167/iovs.11-9344.

Pérez Fornos A., Sommerhalder J. & Pelizzone M. 2011. "Reading with a Simulated 60-Channel Implant." *Frontiers in Neuroscience* 5 (January): 57. doi:10.3389/fnins.2011.00057.

Pérez Fornos A., Sommerhalder J., Pittard A., Safran A.B. & Pelizzone M. 2008. "Simulation of Artificial Vision: IV. Visual Information Required to Achieve Simple Pointing and Manipulation Tasks." *Vision Research* 48 (16) (July): 1705–18. doi:10.1016/j.visres.2008.04.027.

Pérez Fornos A., Sommerhalder J., Rappaz B., Safran A.B. & Pelizzone M. 2005. "Simulation of Artificial Vision, III: Do the Spatial or Temporal Characteristics of Stimulus Pixelization Really Matter?" *Investigative Ophthalmology & Visual Science* 46 (10) (October): 3906–12. doi:10.1167/iovs.04-1173.

Peters M., Laeng B., Latham K., Jackson M., Zaiyouna R. & Richardson C. 1995. "A Redrawn Vandenberg and Kuse Mental Rotations Test: Different Versions and Factors That Affect Performance." *Brain and Cognition* 28: 39–58.

Piaget J. 1937. *La Construction Du Réel Chez L'enfant / The Construction of the Real in the Child*.

- Piaget J. & Inhelder B. 1948. *La Représentation de L'espace Chez L'enfant*.
- Presson C.C. & Montello D.R. 1988. "Points of Reference in Spatial Cognition Stalking the Elusive Landmark." *British Journal of Developmental Psychology* 6 (1960): 378–381. doi:10.1111/j.2044-835X.1988.tb01113.x.
- Rainville C., Marchand N. & Passini R. 2002. "Performances of Patients with a Dementia of the Alzheimer Type in the Standardized Road-Map Test of Direction Sense." *Neuropsychologia* 40 (5): 567–573. doi:10.1016/S0028-3932(01)00133-6.
- Ralston H.J. 1958. "Energy-Speed Relation and Optimal Speed during Level Walking." *Internationale Zeitschrift Für Angewandte Physiologie Einschliesslich Arbeitsphysiologie* 17 (4): 277–283. doi:10.1007/BF00698754.
- Ran L., Helal S. & Moore S. 2004. "Drishti: An Integrated Indoor/outdoor Blind Navigation System and Service." *Proceedings - Second IEEE Annual Conference on Pervasive Computing and Communications, PerCom*: 23–30. doi:10.1109/PERCOM.2004.1276842.
- Ren Q. 2014. "Visual Prosthesis, Optic Nerve Approaches." *Encyclopedia of Computational Neuroscience*. doi:10.1007/978-1-4614-7320-6.
- Renier L., Collignon O., Poirier C., Tranduy D., Vanlierde A., Bol A., Veraart C. & De Volder A.G. 2005. "Cross-Modal Activation of Visual Cortex during Depth Perception Using Auditory Substitution of Vision." *NeuroImage* 26 (2): 573–580. doi:10.1016/j.neuroimage.2005.01.047.
- Rheede J.J. Van, Kennard C. & Hicks S.L. 2010. "Simulating Prosthetic Vision: Optimizing the Information Content of a Limited Visual Display." *Journal of Vision* 10 (14): 1–14. doi:10.1167/10.14.32.Introduction.
- Riecke B.E., Bodenheimer B., McNamara T.P., Williams B., Peng P. & Feuereissen D. 2010. "Do We Need to Walk for Effective Virtual Reality Navigation? Physical Rotations Alone May Suffice." *Spatial Cognition VII*: 234–247. doi:10.1007/978-3-642-14749-4_21.
- Rizzo III J.F., Wyatt J.L., Loewenstein J., Kelly S.K. & Doug S. 2003a. "Perceptual Efficacy of Electrical Stimulation of Human Retina with a Microelectrode Array during Short-Term Surgical Trials." *Investigative Ophthalmology & Visual Science* 44 (12) (December 1): 5362–5369. doi:10.1167/iovs.02-0817.
- Rizzo III J.F., Wyatt J.L., Loewenstein J., Kelly S.K. & Shire D.B. 2003b. "Methods and Perceptual Thresholds for Short-Term Electrical Stimulation of Human Retina with Microelectrode Arrays." *Investigative Ophthalmology & Visual Science* 44 (12): 5355.
- Rizzo J.F. 2011. "Update on Retinal Prosthetic Research: The Boston Retinal Implant Project." *Journal of Neuro-Ophthalmology: The Official Journal of the North American Neuro-Ophthalmology Society* 31 (2) (June): 160–8. doi:10.1097/WNO.0b013e31821eb79e.
- Roentgen U.R., Gelderblom G.J., Soede M. & de Witte L.P. 2008. "Inventory of Electronic Mobility Aids for Persons with Visual Impairments: A Literature Review." *Journal of Visual Impairment & Blindness* 102 (11): 702–724.
- Roessler G., Laube T., Brockmann C., Kirschkamp T., Mazinani B., Goertz M., Koch C., et al. 2009. "Implantation and Explantation of a Wireless Epiretinal Retina Implant Device:

Observations during the EPIRET3 Prospective Clinical Trial.” *Investigative Ophthalmology & Visual Science* 50 (6) (June): 3003–8. doi:10.1167/iovs.08-2752.

Ruddle R.A., Volkova E., Mohler B. & Bühlhoff H.H. 2011. “The Effect of Landmark and Body-Based Sensory Information on Route Knowledge.” *Memory & Cognition* 39 (4) (May): 686–699. doi:10.3758/s13421-010-0054-z.

Ruotolo F., Ruggiero G., Vinciguerra M. & Iachini T. 2012. “Sequential vs Simultaneous Encoding of Spatial Information: A Comparison between the Blind and the Sighted.” *Acta Psychologica* 139 (2): 382–389. doi:10.1016/j.actpsy.2011.11.011.

Sabbah N., Authie C.N., Sanda N., Mohand-Said S., Sahel J. -a. & Safran a. B. 2014. “Importance of Eye Position on Spatial Localization in Blind Subjects Wearing an Argus II Retinal Prosthesis.” *Investigative Ophthalmology & Visual Science* 55: 8259–8266. doi:10.1167/iovs.14-15392.

Sakaguchi H., Fujikado T., Kanda H., Osanai M., Fang X., Nakauchi K., Ikuno Y., et al. 2004. “Electrical Stimulation with a Needle-Type Electrode Inserted into the Optic Nerve in Rabbit Eyes.” *Japanese Journal of Ophthalmology* 48: 552–557. doi:10.1007/s10384-004-0114-7.

Sakaguchi H., Kamei M., Fujikado T., Yonezawa E., Ozawa M., Cecilia-Gonzalez C., Ustari-Gonzalez O., Quiroz-Mercado H. & Tano Y. 2009. “Artificial Vision by Direct Optic Nerve Electrode (AV-DONE) Implantation in a Blind Patient with Retinitis Pigmentosa.” *Artificial Organs* 12 (3) (January 1): 206–9. doi:10.1007/s10047-009-0467-2.

Sanchez J., Tadres A., Pascual-Leone A. & Merabet L. 2009. “Blind Children Navigation through Gaming and Associated Brain Plasticity.” In *2009 Virtual Rehabilitation International Conference*, 29–36. IEEE. doi:10.1109/ICVR.2009.5174201.

Sanchez-Vives M. V & Slater M. 2005. “From Presence to Consciousness through Virtual Reality.” *Nature Reviews. Neuroscience* 6 (4): 332–339. doi:10.1038/nrn1651.

Scharwächter T. & Franke U. 2015. “Low-Level Fusion of Color , Texture and Depth for Robust Road Scene Understanding.” In *Intelligent Vehicles Symposium (IV)*, 599–604.

Schmidt E.M., Bak M.J., Hambrecht F.T., Kufta C. V, O’Rourke D.K. & Vallabhanath P. 1996. “Feasibility of a Visual Prosthesis for the Blind Based on Intracortical Microstimulation of the Visual Cortex.” *Brain* 119 (Pt 2 (April): 507–22.

Shepard R.N. & Metzler J. 1971. “Mental Rotation of Three-Dimensional Objects Abstract . The Time Required to Recognize That Two Perspective Drawings Portray.” *Science (New York, N.Y.)* 171 (FEBRUARY): 701–703.

Shire D.B., Kelly S.K., Chen J., Doyle P., Gingerich M.D., Cogan S.F., Drohan W.A., et al. 2009. “Development and Implantation of a Minimally Invasive Wireless Subretinal Neurostimulator.” *IEEE Transactions on Bio-Medical Engineering* 56 (10) (October): 2502–11. doi:10.1109/TBME.2009.2021401.

Siegel A.W. & White S. 1975. “The Development of Spatial Representations of Large-Scale Environments.” In *Advances in Child Development and Behavior*, edited by H Reese, 10:10–55. New York: Academic Press. doi:http://dx.doi.org/10.1016/S0065-2407(08)60007-5.

Simson H., Gold D. & Zuvella B. 2005. "Inégalité Des Chances : Rapport Sur Les Besoins Des Personnes Aveugles Ou Handicapées Visuelles Vivant Au Canada."

Sommerhalder J., Oueghlani E., Bagnoud M., Leonards U., Safran A.B. & Pelizzone M. 2003. "Simulation of Artificial Vision: I. Eccentric Reading of Isolated Words, and Perceptual Learning." *Vision Research* 43 (3) (February): 269–83.

Sommerhalder J., Rappaz B., de Haller R., Pérez Fornos A., Safran A.B. & Pelizzone M. 2004. "Simulation of Artificial Vision: II. Eccentric Reading of Full-Page Text and the Learning of This Task." *Vision Research* 44 (14) (January): 1693–706. doi:10.1016/j.visres.2004.01.017.

Souman J.L., Giordano P.R., Schwaiger M., Frissen I., Thümmel T., Ulbrich H., Luca a. De, Bülthoff H.H. & Ernst M.O. 2011. "CyberWalk: Enabling Unconstrained Omnidirectional Walking through Virtual Environments." *ACM Transactions on Applied Perception* 8 (4) (November 1): 1–22. doi:10.1145/2043603.2043607.

Srivastava N.R. & Troyk P.R. 2005. "A Proposed Intracortical Visual Prosthesis Image Processing System." In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5:5264–7. doi:10.1109/IEMBS.2005.1615667.

Srivastava N.R., Troyk P.R. & Dagnelie G. 2009. "Detection, Eye-Hand Coordination and Virtual Mobility Performance in Simulated Vision for a Cortical Visual Prosthesis Device." *Journal of Neural Engineering* 6 (3) (June): 035008. doi:10.1088/1741-2560/6/3/035008.

Steck S.D. & Mallot H. a. 2000. "The Role of Global and Local Landmarks in Virtual Environment Navigation." *Presence: Teleoperators and Virtual Environments* 9 (1): 69–83. doi:10.1162/105474600566628.

Stingl K., Bartz-Schmidt K.U., Besch D., Braun A., Bruckmann A., Gekeler F., Greppmaier U., et al. 2013. "Artificial Vision with Wirelessly Powered Subretinal Electronic Implant Alpha-IMS." *Proceedings. Biological Sciences / The Royal Society* 280 (1757): 20130077. doi:10.1098/rspb.2013.0077.

Stingl K., Bartz-Schmidt K.U., Besch D., Chee C.K., Cottrill C.L., Gekeler F., Groppe M., et al. 2015. "Subretinal Visual Implant Alpha IMS – Clinical Trial Interim Report." *Vision Research*: 1–12. doi:10.1016/j.visres.2015.03.001.

Stronks H.C. & Dagnelie G. 2014. "The Functional Performance of the Argus II Retinal Prosthesis." *Expert Review of Medical Devices* 11 (1): 23–30. doi:10.1586/17434440.2014.862494.

Sun J., Lu Y., Cao P., Li X., Cai C., Chai X., Ren Q. & Li L. 2011. "Spatiotemporal Properties of Multiphased Electrically Evoked Potentials Elicited by Penetrative Optic Nerve Stimulation in Rabbits." *Investigative Ophthalmology & Visual Science* 52 (1) (January): 146–54. doi:10.1167/iovs.09-4024.

Tatur G. 2011. "Conception D'un Système de Vision Par Phosphènes."

Terziman L. 2012. "Contribution to the Study of 3D Interaction Techniques and Sensory Feedbacks for Improving Navigation and Walking in Virtual Reality." INSA de Rennes.

Thinus-Blanc C. & Gaunet F. 1997. "Representation of Space in Blind Persons: Vision as a Spatial Sense?" *Psychological Bulletin* 121 (1): 20–42. doi:10.1037/0033-2909.121.1.20.

- Thompson R.W., Barnett G.D., Humayun M.S., Dagnelie G. & Barnett D.G. 2003. "Facial Recognition Using Simulated Prosthetic Pixelized Vision." *Investigative Ophthalmology & Visual Science* 44 (11) (November 1): 5035–5042. doi:10.1167/iovs.03-0341.
- Thorndyke P.W. & Hayes-Roth B. 1982. "Differences in Spatial Knowledge Acquired from Maps and Navigation." *Cognitive Psychology* 14 (4) (October): 560–589. doi:10.1016/0010-0285(82)90019-6.
- Tinti C., Adenzato M., Tamietto M. & Cornoldi C. 2006. "Visual Experience Is Not Necessary for Efficient Survey Spatial Cognition: Evidence from Blindness." *Quarterly Journal of Experimental Psychology* 59 (7): 1306–1328. doi:10.1080/17470210500214275.
- Tolman E.C. 1948. "Cognitive Maps in Rats and Men." *The Psychological Review* 55 (4) (July): 189–208.
- Troyk P.R., Bak M.J., Berg J.A., Bradley D.C., Cogan S.F., Erickson R., Kufta C. V, McCreery D., Schmidt E.M. & Towle V.L. 2003. "A Model for Intracortical Visual Prosthesis Research." *Artificial Organs* 27 (11) (November): 1005–15.
- Vaidya A., Borgonovi E., Taylor R.S., Sahel J.-A., Rizzo S., Stanga P.E., Kukreja A. & Walter P. 2014. "The Cost-Effectiveness of the Argus II Retinal Prosthesis in Retinitis Pigmentosa Patients." *BMC Ophthalmology* 14: 49. doi:10.1186/1471-2415-14-49.
- Vandenberg S.G. & Kuse A.R. 1978. "Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization." *Perceptual and Motor Skills* 47 (2): 599–604.
- Vanlierde A. & Wanet-Defalque M.C. 2005. "The Role of Visual Experience in Mental Imagery." *Journal of Visual Impairment and Blindness* 99 (3): 165–178.
- Vanlierde A. & Wanet-Defalque M.-C. 2004. "Abilities and Strategies of Blind and Sighted Subjects in Visuo-Spatial Imagery." *Acta Psychologica* 116 (2): 205–222. doi:10.1016/j.actpsy.2004.03.001.
- Veraart C., Raftopoulos C., Mortimer J.T., Delbeke J., Pins D., Michaux G., Vanlierde A., Parrini S. & Wanet-Defalque M.-C. 1998. "Visual Sensations Produced by Optic Nerve Stimulation Using an Implanted Self-Sizing Spiral Cuff Electrode." *Brain Research* 813 (1) (November): 181–186. doi:10.1016/S0006-8993(98)00977-9.
- Veraart C. & Wanet-Defalque M.C. 1987. "Representation of Locomotor Space by the Blind." *Perception & Psychophysics* 42 (2): 132–139. doi:10.3758/BF03210501.
- Villalobos J., Fallon J.B., Nayagam D. a X., Shivdasani M.N., Luu C.D., Allen P.J., Shepherd R.K. & Williams C.E. 2014. "Cortical Activation Following Chronic Passive Implantation of a Wide-Field Suprachoroidal Retinal Prosthesis." *Journal of Neural Engineering* 11: 046017. doi:10.1088/1741-2560/11/4/046017.
- Villalobos J., Nayagam D.A.X., Allen P.J., McKelvie P., Luu C.D., Ayton L.N., Freemantle A.L., et al. 2013. "A Wide-Field Suprachoroidal Retinal Prosthesis Is Stable and Well Tolerated Following Chronic Implantation." *Investigative Ophthalmology & Visual Science* 54 (5) (May): 3751–62. doi:10.1167/iovs.12-10843.

Vingerhoets G., Lannoo E. & Bauwens S. 1996. "Analysis of the Money Road-Map Test Performance in Normal and Brain-Damaged Subjects." *Archives of Clinical Neuropsychology* 11 (1): 1–9. doi:10.1016/0887-6177(95)00055-0.

Völkel T., Kühn R. & Weber G. 2008. "Mobility Impaired Pedestrians Are Not Cars: Requirements for the Annotation of Geographical Data." In *Computers Helping People with Special*, 5105:1085–1092.

Waller D., Loomis J.M. & Steck S.D. 2003. "Inertial Cues Do Not Enhance Knowledge of Environmental Layout." *Psychonomic Bulletin & Review* 10 (4): 987–993. doi:10.3758/BF03196563.

Wang J., Lu Y., Gu L., Zhou C. & Chai X. 2014a. "Moving Objects Recognition under Simulated Prosthetic Vision Using Background-Subtraction-Based Image Processing Strategies." *Information Sciences* (March). doi:10.1016/j.ins.2014.02.136.

Wang J., Wu X., Lu Y., Wu H., Kan H. & Chai X. 2014b. "Face Recognition in Simulated Prosthetic Vision: Face Detection-Based Image Processing Strategies." *Journal of Neural Engineering* 11: 046009. doi:10.1088/1741-2560/11/4/046009.

Wang J., Wu X., Lu Y., Wu H., Kan H. & Chai X. 2014c. "Face Recognition in Simulated Prosthetic Vision: Face Detection-Based Image Processing Strategies." *Journal of Neural Engineering* 11 (4) (June 12): 046009. doi:10.1088/1741-2560/11/4/046009.

Wang L., Yang L. & Dagnelie G. 2008. "Virtual Wayfinding Using Simulated Prosthetic Vision in Gaze-Locked Viewing." *Optometry and Vision Science* 85 (11): 1057–1063. doi:10.1097/OPX.0b013e31818b9f36.Virtual.

Ward J. & Meijer P. 2010. "Visual Experiences in the Blind Induced by an Auditory Sensory Substitution Device." *Consciousness and Cognition* 19 (1): 492–500. doi:10.1016/j.concog.2009.10.006.

Weiland J.D., Yanai D., Mahadevappa M., Williamson R., Mech B. V, Fujii G.Y., Little J., Greenberg R.J., de Juan E. & Humayun M.S. 2003. "Electrical Stimulation of Retina in Blind Humans." In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, 2021–2022. IEEE. doi:10.1109/IEMBS.2003.1280131.

Weiland J.D., Yanai D., Mahadevappa M., Williamson R., Mech B. V, Fujii G.Y., Little J., Greenberg R.J., de Juan E. & Humayun M.S. 2004. "Visual Task Performance in Blind Humans with Retinal Prosthetic Implants." In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6:4172–3. doi:10.1109/IEMBS.2004.1404164.

WHO. 2014. "Visual Impairment and Blindness Fact Sheet N° 282." World Health Organization.

Willis S. & Helal S. 2005. "RFID Information Grid for Blind Navigation and Wayfinding." In *Proceedings - International Symposium on Wearable Computers, ISWC*, 34–37. doi:10.1109/ISWC.2005.46.

Wu K.J., Zhang C., Huang W.C., Li L.M. & Ren Q.S. 2010. "Current Research of C-Sight Visual Prosthesis for the Blind." *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*: 5875–5878. doi:10.1109/IEMBS.2010.5627521.

Yang Y., Guo H., Tong S., Zhu Y. & Qiu Y. 2009. "Neurophysiology Study of Early Visual Processing of Face and Non-Face Recognition under Simulated Prosthetic Vision." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2009* (January): 3952–5. doi:10.1109/IEMBS.2009.5333672.

Zacharias J. 2006. "Exploratory Spatial Behaviour in Real and Virtual Environments." *Landscape and Urban Planning* 78 (1-2) (October): 1–13. doi:10.1016/j.landurbplan.2005.05.002.

Zapf M.P., Member P.B.M., Lovell N.H., Zheng S., Suaning G.J. & Member S. 2014. "Towards Photorealistic and Immersive Virtual-Reality Environments for Simulated Prosthetic Vision: Integrating Recent Breakthroughs in Consumer Hardware and Software." In *Engineering in Medicine and Biology Society (EMBC), 2014 Annual International Conference of the IEEE.*, 2597–2600.

Zapf M.P.H., Boon M.-Y., Matteucci P.B., Lovell N.H. & Suaning G.J. 2015. "Towards an Assistive Peripheral Visual Prosthesis for Long-Term Treatment of Retinitis Pigmentosa: Evaluating Mobility Performance in Immersive Simulations." *Journal of Neural Engineering* 12 (3): 036001. doi:10.1088/1741-2560/12/3/036001.

Zhao Y., Lu Y., Tian Y., Li L., Ren Q. & Chai X. 2010. "Image Processing Based Recognition of Images with a Limited Number of Pixels Using Simulated Prosthetic Vision." *Information Sciences* 180 (16) (August): 2915–2924. doi:10.1016/j.ins.2010.04.021.

Zhao Y., Lu Y., Zhou C., Chen Y., Ren Q. & Chai X. 2011. "Chinese Character Recognition Using Simulated Phosphene Maps." *Investigative Ophthalmology & Visual Science* 52 (6) (May): 3404–12. doi:10.1167/iovs.09-4234.

Zhou D.D., Dorn J.D. & Greenberg R.J. 2013a. "The Argus[®] II Retinal Prosthesis System: An Overview." *Multimedia and Expo* doi:10.1109/ICMEW.2013.6618428.

Zhou D.D., Dorn J.D. & Greenberg R.J. 2013b. "The Argus[®] II Retinal Prosthesis System: An Overview." *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference*: 1–6. doi:10.1109/ICMEW.2013.6618428.

Zrenner E. 2013. "Fighting Blindness with Microelectronics." *Science Translational Medicine* 5 (210) (November 6): 210ps16. doi:10.1126/scitranslmed.3007399.

Zrenner E., Wilke R.G.H., Sachs H., Gekeler F., Besch D., Benav H., Wilhelm B., Wrobel W.-G. & Group S.S. 2009. "Subretinal Microelectrode Arrays Implanted Into Blind Retinitis Pigmentosa Patients Allow Recognition of Letters and Direction of Thin Stripes." In *International Federation for Medical and Biological Engineering*, edited by Olaf Dössel and Wolfgang C. Schlegel, 25/9:444–447. IFMBE Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-03889-1.

Table des figures

Introduction	15
Figure 1 : Dispositif utilisé par Bach-y-Rita. La grille de stimulation est fixée sur le dossier de la chaise. La caméra est disponible à proximité de la chaise. Source : [Bach-y-Rita et al. 1969].....	18
Figure 2 : Le Tongue Unit Display évalué en tâche de navigation. Le sujet porte une caméra dont le signal est traité pour fournir une séquence d'électrostimulation à la matrice d'électrodes. Source : [Chebat et al. 2011].....	18
PARTIE 1 ÉTAT DE L'ART.....	21
Figure 3 : Schéma de l'œil. Source : Jérôme Dupire.....	23
Figure 4 : Les différentes couches cellulaires de la rétine. Source : http://ophtalmologie.pro	24
Figure 5 : Évolution de la densité des photorécepteurs sur l'ensemble de la rétine. La fovéa ainsi que le point aveugle sont visibles. Source : cours de R. Puiol.....	25
Figure 6 : Chemin parcouru par l'information issue du champ visuel gauche et droit pour chaque œil. Source : http://lecerveau.mcgill.ca	26
Figure 7 : Positions des aires corticales visuelles. Source : académie de Dijon	27
Figure 8 : Différents lieux d'implantation pour les neuroprothèses visuelles. A. Cortex visuel et nerf optique. B. Implants épitrétiens, au contact des cellules ganglionnaires. C. Implants sous-rétiens, insérés à la place des photorécepteurs. D. Les implants supra-choroïdiens ont comme avantage d'être moins difficiles à implanter. Source : [Zrenner 2013].....	30
Figure 9 : Système en développement par le Monash Vision Group. Des lunettes contiennent une caméra, ce signal est traité par un processeur avec lequel les sujets peuvent interagir, l'information et l'énergie sont envoyées à l'implant grâce à une antenne située au niveau du lobe occipital. Source : [Monash Vision Group 2013].....	31
Figure 10 : Implant du projet Intracortical Visual Prosthesis. Les modules utilisés sont visibles à droite sur fond jaune. Source : Institut Technologique de l'Illinois.....	32
Figure 11 : À gauche : Utah Electrode Array, implant d'électrodes de microstimulations développé par Norman. À droite le système initialement prévu pour restaurer la vision chez les aveugles à l'aide de cet implant. Source : [Normann 2007].....	33

Figure 12 : Schéma de fonctionnement de l’implant développé dans le cadre du projet MiViP. Un connecteur transcutané assure la liaison entre l’implant et le dispositif externe de stimulation. Source : [Veraart et al. 1998]34

Figure 13 : Carte des phosphènes utilisés pour transmettre l’information visuelle d’une caméra au nerf optique. Chaque point représente la position d’un phosphène dans le champ visuel du sujet. Source : [Brelén et al. 2005]..... 35

Figure 14 : Projet C-Sight d’implant du nerf optique. Une caméra implantée envoie l’information visuelle à un processeur qui transmet ensuite le signal adapté aux électrodes situés sur le nerf optique. Source : [Chai et al. 2008]36

Figure 15 : Implant développé par le projet AV-DONE. Les électrodes sont insérées dans la tête du nerf optique. Sept électrodes de stimulation et une électrode de retour sont présentes. Source : [Nishida et al. 2015]..... 37

Figure 16 : Système complet de l’Argus I. (A) Caméra montée sur lunette; (B) Antenne de transmission; (C) Boîtier récepteur; (D) implant intraoculaire. Source : [Humayun et al. 2003].38

Figure 17 : Matrice d’électrodes située à l’intérieur de l’œil au contact de la rétine. Source : [Humayun et al. 2003].38

Figure 18 : L’implant Argus II. L’antenne située au niveau de la tempe permet de récupérer l’information provenant du dispositif externe. Source : Second Sight38

Figure 19 : Tâche de mobilité effectuée par les premiers patients ayant reçu l’Argus II. À gauche, les sujets doivent se diriger vers une porte sombre située sur un mur clair. À droite, les sujets doivent marcher le long d’une ligne blanche jusqu’au bout. Source : [Humayun et al. 2009].....39

Figure 20 : Gauche : Pourcentages de succès pour l’orientation et la marche vers la porte. Droite : Pourcentages de succès pour le suivi de ligne. Source : [Humayun et al. 2012]..... 40

Figure 21 : Implant Alpha IMS. L’énergie nécessaire ainsi que le contrôle de la luminosité passent par une antenne située derrière l’oreille. Source : [Stingl et al. 2015]..... 41

Figure 22 : Images de l’implant développé dans le projet Boston Retinal Implant Project. À gauche : les lunettes prévues pour l’implant, une caméra est située sur la branche des lunettes (flèche rouge). À droite : Le transfert d’information se fait via une bobine située sur la sclère (flèche jaune). Le dispositif de traitement de l’information est situé dans un boîtier en titane contre l’œil (flèche blanche). Seule la matrice d’électrodes est implantée à l’intérieur de l’œil (flèche rouge). Source : [Rizzo 2011]42

Figure 23 : À droite : Schéma de fonctionnement de l’implant IMI. À gauche : Implant IMI, la matrice d’électrodes est visible sur l’extrémité gauche du dispositif. Source : [Hornig et al. 2007].....42

Figure 24 : Epi-Ret 3, À gauche : Schéma de l’implant, l’information et l’énergie sont transmises à l’implant via une bobine située à l’intérieur de l’œil. À droite : Photographie

de l'implant, la bobine est reliée aux électrodes de stimulation visibles à gauche et placées contre la rétine. Source : [Mokwa et al. 2008].....	44
Figure 25 : Principe de la Vision Prothétique Simulée. L'image originale est traitée par un Simulateur de Vision Prothétique qui dispose des Paramètres de la Matrice (taille, disposition, <i>dropout</i>) ainsi que des caractéristiques des phosphènes (forme, nombre de niveaux de gris). Le simulateur crée alors un rendu qui peut être affiché ensuite pour les sujets via un moniteur ou un visiocasque.....	47
Figure 26 : Simulateur de vision prothétique de Cha. Une caméra couplée à une lentille (C) envoie le signal vidéo directement à un écran qui contient un masque perforé pour simuler la vision phosphénique (M). La caméra et le moniteur sont alimentés par une batterie portée à la ceinture (B). Source [Cha et al. 1992a].....	49
Figure 27 : Dispositif utilisé par Cha dans une expérience de navigation en vision prothétique simulée. Les traits pleins et traits pointillés indiquent la présence ou l'absence de parois en tissu. Les disques noirs sont des obstacles cylindriques suspendus à des rails. Les trajets démarraient à la porte 1 ou 2 pour se terminer à la porte 3. Un exemple de trajet est représenté en pointillé. Source : [Cha et al. 1992b].....	49
Figure 28 : Tâche visuo-motrice en vision prothétique. Les sujets doivent regarder la forme imprimée sur l'objet qu'ils ont en main puis déposer l'objet sur la forme en question présente sur le tableau devant eux. Source : [Pérez Fornos et al. 2008].....	55
Figure 29 : Le système mis en place par Rheede et ses collaborateurs comporte un oculomètre ainsi qu'un visiocasque. Des tissus sombres étaient utilisés en plus pour masquer la vision périphérique des sujets. Source : [Rheede et al. 2010].....	55
Figure 30 : Casque utilisé pour la tâche de mobilité dans l'expérience de Dowling. Source : [Dowling et al. 2005].....	59
Figure 31 : Carte du trajet à parcourir dans l'expérience de Dowling. Les sujets devaient aller du point A au point B en effectuant différentes actions sur le trajet. Source : [Dowling et al. 2005].....	59
Figure 32 : À gauche : L'environnement virtuel utilisé par Dagnelie et ses collaborateurs avec le rendu en vision prothétique simulée à droite. Source : [Dagnelie et al. 2007].....	60
Figure 33 : Environnement virtuel utilisé par Rheede et ses collaborateurs vu du dessus. Les sujets devaient parcourir deux trajets dans cet environnement. Pour le premier ils devaient traverser le tunnel 1 (1), passer sous un pont (3) puis retourner sous le tunnel 2 (2). Pour le deuxième ils devaient traverser le tunnel 2 (2), passer sous un pont (3), tourner à droite, suivre la structure et passer à côté d'un conteneur (4) et enfin tourner de nouveau à droite pour passer sous un autre pont (5). Source : [Rheede et al. 2010].....	61
Figure 34 : Un obstacle et sa représentation en vision prothétique simulée par Parikh et collaborateurs. Les bords de la matrice de phosphène servent à indiquer la direction des obstacles immédiats. Dans cet exemple, le cône est détecté et sa direction est indiquée par un phosphène blanc situé dans la partie supérieure de la matrice. Source : [Parikh et al. 2013].....	62

Figure 35 : Rendus développés par McCarthy. (a) Photo de départ en nuances de gris. Le faible contraste entre le sol et l'obstacle le rend difficile à voir sans traitement (b). En appliquant seulement un rendu de distance, l'obstacle n'est pas visible car trop près du sol (c). L'algorithme qu'ils ont développé leur permet d'isoler l'objet en le distinguant du sol (d) ce qui permet de le mettre en surbrillance dans le rendu de distance augmentée qu'ils ont développé (e). Source : [McCarthy et al. 2014].....63

Figure 36 : Rendus utilisés par Zapf lors de ses travaux de thèse, deux tâches de mobilités ont été effectuées avec deux tailles de matrice différentes (10x10 phosphènes ou 20x20 phosphènes). Source : [Zapf et al. 2014].....64

Figure 37 : Tâches effectuées par les sujets dans l'expérimentation de Zapf. Les sujets devaient d'abord traverser un couloir en évitant des plots (a) puis éviter des piétons statiques (b) avant de suivre une ligne blanche au sol (d) pour finalement longer un trottoir en évitant des piétons en mouvement (c). Source : [Zapf et al. 2015].....64

Figure 38 : Item du questionnaire de Vandenberg. Le sujet doit trouver parmi les quatre derniers dessins celui représentant l'objet dessiné dans le premier. Source : [Peters et al. 1995].....68

Figure 39 : Money Road Map Test. À chaque intersection, le sujet doit indiquer rapidement dans quelle direction il doit tourner pour suivre le trajet. Source : [Rainville et al. 2002].....68

Figure 40 : Exemples d'items du test de Guilford et Zimmerman. Pour chacun des trois items, les sujets doivent choisir parmi les cinq propositions celle qui correspond au mouvement du bateau. Source : [Dejos 2012]......69

Figure 41 : À gauche le modèle de Siegel et White considère que des informations de distance ne sont mémorisées qu'à partir du niveau de représentation Survey. À droite : Ishikawa et Montello proposent que l'information métrique soit mémorisée dès le début du processus et que les trois niveaux de représentations ne soient plus successifs dans le temps. Source : [Ishikawa & Montello 2006].....73

Figure 42 : À gauche : Représentation en graphe des connaissances spatiales, ce qui correspond au niveau route de Siegel et White, les points de repère sont reliés entre eux par des routes. À droite : Représentation des connaissances spatiales sous forme d'une carte cognitive intégrée (Survey). Source : [Meilinger 2008a].....75

Figure 43 : Modèle de Meilinger, les vista spaces contiennent chacun un référentiel et sont reliés entre eux par des *perspective shifts*. Source : [Meilinger 2008a].....76

PARTIE 2 CONTRIBUTIONS.....81

Figure 1 : Exemple de rendu par localisation testé dans notre équipe. Le traitement de l'image permet de détecter les visages pour n'en présenter que la position aux sujets dans l'approche par localisation (Extrait de [Denis 2014]).....83

Figure 2 : Conditions et points de repère de l'expérience de navigation par localisation. Chaque ligne correspond à un point de repère donné. Dans l'ordre : une porte, un distributeur d'eau et un poster. Chaque colonne correspond à une condition : a) Irrlicht,

b) Brut et c) Localisation. Dans l'approche par localisation, les portes sont rendues par trois points blancs, les distributeurs d'eau par deux points gris clair et les posters par quatre points gris foncé. Pour aider les sujets à fixer leur regard, une croix rouge fixe était présente au centre de l'écran.....	86
Figure 3 : Asservissement des percepts. a) Image de l'environnement générée par le moteur 3D Irrlicht. b) Quand le sujet dirige son regard (cercle jaune) sur la croix de fixation, la matrice de phosphènes est placée au centre de l'écran. La flèche jaune indique un futur déplacement du regard vers le haut et la droite. c) La position de la matrice de phosphènes est modifiée pour correspondre au déplacement du regard. Le sujet ne peut donc pas explorer la matrice de phosphènes. La croix rouge encourage le sujet à garder le regard droit devant.....	87
Figure 4 : Schéma fonctionnel du dispositif d'expérimentation utilisé. Quatre agents sont en fonctionnement simultanément. Tous sont connectés à un bus logiciel afin d'échanger des informations. Les flèches représentent les différents messages IVY réceptionnés par les agents ou envoyés au bus.....	88
Figure 5 : Fond de carte utilisé par les sujets pour dessiner les trajets effectués.....	90
Figure 6 : Temps moyen mis par les sujets pour accomplir la tâche. Les barres d'erreurs sont les intervalles de confiance à 95 %.....	91
Figure 7 : Nombre moyen d'erreurs présentes dans les dessins pour chaque condition. En condition contrôle Irrlicht aucun sujet n'a fait d'erreur.	92
Figure 8 : Effet de superposition des items en VPS par localisation. a) Capture d'un couloir rendu par le moteur 3D Irrlicht. b) Rendu Brut. c) Rendu par localisation : seul le poster de droite est reconnaissable, la porte étant trop écrasée, elle peut facilement être confondue avec un distributeur d'eau. Le code correspondant au poster de gauche n'est pas entièrement visible non plus car l'angle supérieur droit se confond avec un des deux autres angles.	94
Figure 9 : Les trois environnements utilisés dans le cadre de cette expérimentation. Les bijoux sont représentés en bleu, la base de départ en Jaune. La flèche noire indique l'orientation de départ.....	99
Figure 10 : Présentation des trois rendus de la deuxième expérimentation. a) Le rendu Irrlicht utilisé ici comme condition contrôle. b) Rendu brut classique. c) Rendu brut augmenté. La base de départ est affichée sur la première ligne. Un angle dans l'environnement est représenté dans la deuxième ligne. La troisième ligne montre la perception d'un joyau.	100
Figure 11 : Schéma fonctionnel des différents programmes utilisés dans la deuxième expérimentation. Les caméras de motion capture permettent de mesurer les déplacements du casque pour les transmettre au moteur 3D Irrlicht.....	102
Figure 12 : Indice de performance des trois conditions. Les sujets ont effectué la tâche plus rapidement en condition augmentée qu'en condition brut simple. Les barres d'erreurs sont les intervalles de confiance à 95 %.....	104

Figure 13 : Moyenne des dessins de carte par condition. Le fait d'améliorer la perception des contrastes entre le sol et les murs améliore l'image mentale des sujets. Les barres d'erreurs sont les intervalles de confiance à 95 %.....105

Figure 14 : Exemple de rendu basé sur la distance entre la caméra et les surfaces environnantes. (a) Un couloir contient un obstacle au sol. (b) Rendu brut classique (c) Rendu de distance (d) Le sol est détecté ce qui permet d'isoler l'obstacle qui ne lui appartient pas (e) Le rendu de distance est augmenté par l'isolation du sol et l'ajout de l'obstacle en surbrillance. Source : [McCarthy et al. 2014] 110

Figure 16 : Rendus prothétiques du même point de vue que Figure 15. a) Rendu VPS-LimInf. b) Rendu VPS-Lim3. c) Rendu VPS-Lim6. d) Rendu VPS-Lim9. Ces rendus sont générés en temps réel à partir des rendus de vision Irrlicht présentés dans la Figure 15.113

Figure 17 : Rendus prothétiques augmentant la structure de l'environnement. a) Rendu Irrlicht original (IRR-LimInf). b) Rendu brut classique (VPS-LimInf). c) Rendu prothétique basé sur la distance existant entre les éléments et la caméra (VPS-Distance). d) Rendu VPS-FilDeFer.....114

Figure 18 : Schéma fonctionnel des agents utilisés pour cette troisième expérimentation. Même si certains agents ont été recodés en grande partie, les interactions entre agents sont similaires entre cette expérimentation et la précédente.....116

Figure 19 : Moyennes de l'Indice de Performance des conditions de vision Irrlicht. Les différences significatives sont indiquées avec * pour $p < 0,05$; ** pour $p < 0,01$ et *** pour $p < 0,001$. Les barres d'erreurs sont les intervalles de confiance à 95 % 119

Figure 20 : Moyennes de l'Indice de Performance sur les conditions de vision prothétique. Les principaux résultats statistiques sont indiqués : ** pour $p < 0,01$; *** pour $p < 0,001$. Les barres d'erreurs sont les intervalles de confiance à 95 %.....120

Figure 21 : Moyennes du Nasa-TLX pour les conditions de rendus Irrlicht. Les barres d'erreurs sont les intervalles de confiance à 95 %.....120

Figure 22 : Moyennes du Nasa-TLX pour les conditions de vision prothétique. Les barres d'erreur sont les intervalles de confiance à 95 %.....121

Figure 23 : Exemples de cartes dessinées par les sujets. La carte de l'environnement est représentée en haut à gauche. Les trois autres images sont des représentations produites par des sujets après exploration. Les juges disposaient de la carte de départ pour évaluer tous les dessins.....122

Figure 24 : Moyenne des notes attribuées par les juges sur les dessins pour chacune des conditions Irrlicht. Les barres d'erreurs sont les intervalles de confiance à 95 %.....122

Figure 25 : Moyenne des scores des dessins pour chaque Conditions VPS. Les barres d'erreurs sont les intervalles de confiance à 95 %.....123

Figure 26 : Le score de Santa Barbara des sujets est affiché en relation avec les trois variables dépendantes de l'expérimentation. Pour chaque graphique, chaque élément correspond à un sujet, en abscisse se trouve l'échelle du questionnaire de Santa-Barbara

et en ordonnée se trouve a) l'Indice de Performance b) les scores des dessins c) le Nasa-TLX.....	124
Figure 27 : Comparaison des moyennes des indices de performances entre le groupe des non-joueurs et le groupe des joueurs. En bleu sont représentées les conditions de rendus Irrlicht. En jaune sont les conditions en VPS. Les barres d'erreur sont les intervalles de confiance à 95 %.....	125
Figure 28 : Moyennes du Nasa-TLX des groupes de joueurs et non-joueurs. Les conditions Irrlicht sont représentées en bleu, les conditions VPS le sont en jaune.....	125
Figure 29 : Moyennes des scores des dessins obtenus par les groupes de joueurs et non- joueurs sur chacune des conditions. Les conditions Irrlicht sont représentées en bleu, celles VPS le sont en jaune.....	126
Figure 30 : Exemples de dessins produits par des sujets après exploration avec le rendu VPS-LimInf. Seules la base de départ et la position des joyaux proches sont indiquées.	128
Figure 31 : Différents jeux vidéo classiques limitant la distance de vision. a) Les jeux Silent Hill et b) Amnesia : The Dark Descent utilise la limitation de la profondeur pour amener de l'angoisse aux joueurs. c) Le jeu Tomb Raider l'utilise pour donner une ambiance propice à l'exploration et pour diriger le joueur vers des points lumineux. d) Le mod Dynamic Light développé pour Minecraft permet de rajouter une touche de réalisme au jeu de départ.	131
Figure 32 : Jeux vidéo utilisant un rendu en fil de fer. a) Akalabeth : World of Doom est un des premiers jeux de rôle dans lequel des donjons sont à explorer. b) Battlezone a connu un grand succès et proposait des combats entre tanks. c) Star Wars : The Arcade Game est un simulateur spatial également très populaire. d) REZ est un railshooter en 3D utilisant des rendus en fil de fer pour la plupart de ses éléments.	132
Figure 33 : Différences entre les groupes de Joueurs et Non-joueurs sur les variables dépendantes et le score de Santa-Barbara. Les moyennes ont été obtenues en moyennant sur toutes les conditions. Les barres d'erreurs sont les intervalles de confiance à 95 %.....	134
Figure 34 : Place du Capitole. Les éléments cités par les sujets sont représentés. Les points rouges, bleus et violet correspondent aux positions prises par les sujets pour décrire cet environnement. Certains sujets ont décrit l'environnement depuis une position allocentrée et ne sont donc pas représentés.....	138
Figure 35 : Place Esquirol. Les éléments cités par les participants sont représentés sur ce plan.	139
PARTIE 3 DISCUSSION ET PERSPECTIVES.....	141
Figure 36 : Prévision faite en 2009 sur le développement des implants. L'échelle de gauche illustre les capacités obtenues avec ces résolutions et l'échelle de droite, le nombre potentiel de personnes implantées. La flèche rouge qui a été ajoutée sur cette image indique la situation actuelle. Source : [Chader et al. 2009].....	142

Table des tableaux

Tableau 1 : Principales études portant sur l'évaluation de l'acuité visuelle en vision prothétique simulée.....	50
Tableau 2 : Principales études portant sur l'évaluation de la lecture VPS.....	52
Tableau 3 : Principales études portant sur la reconnaissance et la localisation de visages en vision prothétique simulée.....	53
Tableau 4 : Principales études portant sur la reconnaissance et la saisie d'objets en vision prothétique simulée.....	57
Tableau 5 : Études portant sur la mobilité et la navigation en vision prothétique simulée.....	65

ANNEXES

Table des matières

Questionnaires utilisés.....	191
A. Test CODE	191
B. Vivacité d’Imagerie Mentale	192
C. Nasa-TLX	193
D. Skelton	194
E. Questionnaire de Santa Barbara.....	196
Expérimentations	199
A. Première expérimentation	199
B. Deuxième expérimentation	199
B.1. Environnements utilisés.....	199
C. Troisième expérimentation.....	199
C.1. Environnement d’apprentissage	199
C.2. Environnements d’expérimentation	200
C.3. Consignes fournies aux juges pour l’évaluation des dessins.....	201
D. Quatrième expérimentation	202
D.1. Questionnaire soumis aux sujets	202
D.2. Descriptions faites par les sujets des places et du trajet.....	207
D.3. Résultats du questionnaire.....	214
D.3.a. Sujets	214
D.3.b. Résultats	214
D.3.c. Description de la place du Capitole	215
D.3.d. Description de la place Esquirol.....	215
D.3.e. Navigation dans la rue Saint-Rome.....	216

Questionnaires utilisés

Différents questionnaires ont été utilisés pour les expérimentations.

A. Test CODE

Le test CODE est un test permettant d'évaluer la mémoire de travail des sujets. La passation de ce test est très courte. Nous l'avons principalement utilisé pour s'assurer qu'aucun de nos sujets n'avaient de déficit de mémoire de travail avant de commencer l'expérience. Un symbole est associé à chacun des neuf chiffres (le zéro n'est pas inclus). Les sujets ont ensuite deux minutes pour remplir dans l'ordre une suite de chiffres avec les symboles correspondants, sans en passer et en allant le plus vite possible.

Test CODE

Temps de passation : 2min

Sujet N° ____

Score : ____ /133

1	2	3	4	5	6	7	8	9
—	⊥	□	⊓	⊔	○	△	⊗	≡

Items d'exemple

2	1	3	7	2	4	8	2	1	3	2	1	4	2	3	5	2	3	1	4
5	6	3	1	4	1	5	4	2	7	6	3	5	7	2	8	5	4	6	3
7	2	8	1	9	5	8	4	7	3	6	2	5	1	9	2	8	3	7	4
6	5	9	4	8	3	7	2	6	1	5	4	6	3	7	9	2	8	1	7
9	4	6	8	5	9	7	1	8	5	2	9	4	8	6	3	7	9	8	6
2	7	3	6	5	1	9	8	4	5	7	3	1	4	8	7	9	1	4	5
7	1	8	2	9	3	6	7	2	8	5	2	3	1	4	8	4	2	7	6

B. Vivacité d'Imagerie Mentale

Le test de Vivacité d'Imagerie Mentale évalue les capacités des sujets à se représenter des sensations. Les consignes sont les suivantes :

L'objectif de ce questionnaire est d'évaluer la qualité et la vivacité de votre imagerie mentale. Les questions posées vont vous conduire à évoquer des images mentales. Vous devrez évaluer la vivacité de chaque image à l'aide d'une échelle en 5 points, allant de 1 à 5, où

1 = aucune image n'est présente, vous « savez » seulement que vous êtes en train de penser à l'objet.

2 = l'image est vague et faible.

3 = l'image est moyennement claire et modérément vivace.

4 = l'image est claire et assez vivace.

5 = l'image est parfaitement claire et aussi vivace que la perception ou l'expérience directe de l'objet.

Vous devrez évaluer la vivacité de l'image évoquée par chaque question en entourant une des 5 valeurs de l'échelle. Avant de commencer, assurez-vous que vous êtes à l'aise avec l'échelle d'évaluation. Tout au long du questionnaire, utilisez cette échelle pour évaluer la vivacité de chaque image mentale. Essayez de répondre à chaque item séparément, en faisant abstraction des réponses données précédemment. Si vous le souhaitez, vous pouvez fermer les yeux pendant que vous évoquerez les images.

Imagerie Visuelle

Imaginez-vous en train de regarder les éléments suivants. Considérez attentivement l'image qui vous vient à l'esprit. Évaluez les images évoquées par chaque item sur l'échelle en 5 points.

1 - L'ampleur d'un pas lorsqu'on marche et la posture adoptée

2 - La couleur de vêtements qui vous sont familiers

3 - Le soleil se couchant à l'horizon

Imagerie Cutanée (Tactile)

Imaginez-vous en train de 'ressentir' ou de toucher chacun des items suivants. Considérez attentivement la sensation qui vous vient à l'esprit. Évaluez les images évoquées par chaque item sur l'échelle en 5 points.

4 - Sable

5 - Laine

6 - Fourrure

7 - Piqûre d'une épingle

8 - Tiédeur d'un bain chaud

Imagerie kinesthésique

Imaginez-vous en train de réaliser chacune des actions suivantes. Considérez attentivement l'image qui vous vient à l'esprit. Évaluez les images évoquées par chaque item sur l'échelle en 5 points.

9 - Monter un étage à pied

10 - Descendre d'un trottoir

11 - Tourner la clé dans la serrure d'une porte

12 - Atteindre une étagère haut placée

13 - Donner un coup de pied dans quelque chose pour l'enlever du chemin

C. Nasa-TLX

Le test Nasa-TLX évalue la charge cognitive ressentie par les différents sujets. La version française présente ci-dessous a été traduite et adaptée par Julien Cegarra, maître de conférences au centre universitaire Jean-François Champollion qui a bien voulu nous la fournir.

Pour calculer cet indice, les sujets évaluent six dimensions différentes sur une échelle allant de 0 à 10. Ensuite, pour chacun des quinze couples possibles formés par deux de ces dimensions les sujets doivent choisir laquelle leur paraît être la plus importante pour la tâche. Pour chacune de ces comparaisons, un point est attribué à la comparaison choisie. Au final, quinze points sont répartis sur six dimensions avec un maximum de cinq points par dimension. Il faut calculer ensuite pour chacune des dimensions le produit de son score par le nombre de points obtenus et faire la somme de ces produits. Une somme pondérée est obtenue, il suffit ensuite de la diviser par quinze pour obtenir le score du Nasa-TLX. Voici les six dimensions en question et la description qui a été donnée aux sujets. Pour nos expérimentations, les sujets disposaient d'une échelle allant de 0 à 100 par paliers de 10.

EXIGENCE MENTALE (0: Faible / 100: Forte) : Dans quelle mesure des opérations mentales et perceptives ont-elles été requises (par ex.: penser, décider, calculer, se rappeler, regarder, chercher, etc.) ? Ont-elles conduit à une tâche plutôt facile ou difficile, simple ou complexe, abordable ou exigeante ?

EXIGENCE PHYSIQUE (0: Faible / 100: Forte) : Dans quelle mesure des opérations physiques ont-elles été requises (par ex.: pousser, tirer, tourner, superviser, activer, etc.) ? Ont-elles conduit à une tâche plutôt facile ou difficile, lente ou rapide, lâche ou vigoureuse, reposante ou ardue ?

EXIGENCE TEMPORELLE (0: Faible / 100: Forte) : Quelle était la pression temporelle, que ce soit à cause de la cadence rythme ou de l'allure des tâches ou de l'apparition des éléments de la tâche avez-vous ressenti ? L'allure était-elle lente et tranquille ou rapide et frénétique ?

EFFORT (0: Faible / 100: Forte) : Quelle a été la difficulté d'accomplir (mentalement et physiquement) la tâche avec un niveau de performance tel que le vôtre ?

PERFORMANCE (0: Bonne / 100: Faible) : Quelle réussite vous attribuez-vous en ce qui concerne l'atteinte des buts de la tâche fixés par l'expérimentateur (ou par vous-même) ? Dans quelle mesure êtes-vous satisfait de votre performance dans l'accomplissement de ces buts ?

FRUSTRATION (0: Faible / 100: Forte) : Au cours de la tâche, quel sentiment d'être peu sûr de vous, découragé, irrité, stressé et agacé avez-vous ressenti contrairement au fait d'être sûr de vous, satisfait, content, détendu et complaisant ?

D. Skelton

Le questionnaire de Skelton a été utilisé dans nos premières expériences et est essentiellement utilisé pour détecter des déficits. Pour calculer cet indice, il faut faire la moyenne des scores de chaque question en prenant garde d'inverser l'échelle pour la question 9. Pour le test original, les sujets répondent sur une barre continue et les réponses sont à mesurer entre 1 et 100. La version francophone nous a été fournie par Hélène Sauzéron et Bernard N'Kaoua, membres de l'équipe Handicap et Système Nerveux de l'Université de Bordeaux. Les consignes à donner aux sujets sont les suivantes :

VERSION ORIGINALE

Everyday Spatial Questionnaire

Place a mark on the line that best describes your recent experience (e.g. in the last month).

1. Do you get lost in small familiar buildings (like a friend's house, corner store, etc.)?

Never (ever) /-----/ Every time.

2. Do you get lost when you go into in large buildings you've been in before (Malls, office buildings)?

Never (ever) /-----/ Every time.

3. *Do you get lost when you go into large buildings for the first time?*
Never (ever) |-----| Every time.
4. *Do you feel disoriented when you come out of an unfamiliar building?*
Never (ever) |-----| Every time.
5. *Do you feel or get lost when you are in familiar parts of town (like your own neighbourhood)?*
Never (ever) |-----| Every time.
6. *Do you feel or get lost when you are in unfamiliar parts of town?*
Never (ever) |-----| Every time.
7. *Do you get people to take you over a route before you'll go that way yourself?*
Never (ever) |-----| Every time.
8. *Do you stop and ask directions when you are on your way to someplace (or want to)?*
Never (ever) |-----| Every time.
9. *Do you make up shortcuts and figure out new routes from one place to another?*
Never (ever) |-----| Every time.
10. *Do you forget where you put things?*
Never (ever) |-----| Every time.
11. *Do you have trouble finding your car in a parking lot?*
Never (ever) |-----| Every time.
12. *Do you put things like keys and your wallet in specific locations so you can find them again?*
Never (ever) |-----| Every time.
13. *Is it important to you that your clothes and food cupboards are arranged in a very specific way?*
Not at all |-----| Absolutely!

VERSION FRANCOPHONE

Ceci est un questionnaire portant sur l'orientation spatiale et les activités quotidiennes. Pour chaque affirmation, il vous est demandé de la noter en fonction de ce qui vous semble décrire au mieux votre expérience récente (le mois dernier). Veuillez noter ces affirmations en vous appuyant sur votre expérience récente. Échelle : 1=Jamais 100= Tout le temps. Pour la dernière question : 1= Pas du tout. 100= Absolument.

1. *Vous perdez-vous dans des emplacements familiers (maison d'un ami, boutique ...)?*
2. *Vous perdez-vous quand vous allez dans des bâtiments que vous avez déjà visité (centre commercial)?*
3. *Vous perdez-vous quand vous allez dans des bâtiments pour la première fois?*
4. *Vous sentez-vous désorienté quand vous sortez d'un bâtiment non familier?*
5. *Est-ce que vous vous perdez quand vous êtes dans une partie familière de la ville (comme dans votre propre voisinage)?*

6. Est-ce que vous vous perdez ou avez l'impression de vous perdre quand vous êtes dans une partie non familière de la ville ?
7. Demandez-vous à des gens de vous accompagner sur la bonne route avant de la trouver tout(e) seul(e) ?
8. Ressentez-vous le besoin d'arrêter les gens pour demander la direction quand vous êtes en chemin pour aller quelque part ?
9. Prenez-vous des raccourcis et inventez-vous de nouvelles routes pour aller d'un endroit à un autre ?
10. Oubliez-vous où vous posez les objets ?
11. Avez-vous des difficultés à trouver votre voiture dans un parking ?
12. Posez-vous des objets spécifiques comme vos clés, portefeuille à des endroits précis de manière à les retrouver plus tard ?
13. Est-il important pour vous que vos vêtements, vos ustensiles de cuisines ... soient arrangés de manière spécifique ?

E. Questionnaire de Santa Barbara

Le questionnaire de Santa Barbara ou *Santa Barbara Sense-of-Direction Scale* donne une auto-évaluation du sens de l'orientation par les sujets à partir d'évènements quotidiens. Contrairement au questionnaire de Skelton, cette échelle ne permet pas d'identifier des déficits de l'orientation. Pour calculer l'indice de ce questionnaire, il faut moyenner les scores donnés par les sujets pour chaque question en inversant l'échelle pour les questions 1, 3, 4, 5, 7, 9 et 14. Mary Hegarty qui a participé à l'élaboration de ce test a mis en ligne le test lui-même ainsi que son mode d'utilisation¹. Les consignes à donner aux sujets sont les suivantes :

VERSION ORIGINALE

Sex: F M

Today's Date: _____

Age: _____

V. 2

This questionnaire consists of several statements about your spatial and navigational abilities, preferences, and experiences. After each statement, you should circle a number to indicate your level of agreement with the statement. Circle "1" if you strongly agree that the statement applies to you, "7" if you strongly disagree, or some number in between if your agreement is intermediate. Circle "4" if you neither agree nor disagree.

1. *I am very good at giving directions.*

strongly agree 1 2 3 4 5 6 7 strongly disagree

¹ <https://labs.psych.ucsb.edu/hegarty/mary/content/santa-barbara-sense-direction-scale> (Consulté le 12/09/2015).

2. *I have a poor memory for where I left things.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
3. *I am very good at judging distances.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
4. *My “sense of direction” is very good.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
5. *I tend to think of my environment in terms of cardinal directions (N, S, E, W).*
strongly agree 1 2 3 4 5 6 7 strongly disagree
6. *I very easily get lost in a new city.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
7. *I enjoy reading maps.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
8. *I have trouble understanding directions.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
9. *I am very good at reading maps.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
10. *I don’t remember routes very well while riding as a passenger in a car.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
11. *I don’t enjoy giving directions.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
12. *It’s not important to me to know where I am.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
13. *I usually let someone else do the navigational planning for long trips.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
14. *I can usually remember a new route after I have traveled it only once.*
strongly agree 1 2 3 4 5 6 7 strongly disagree
15. *I don’t have a very good “mental map” of my environment.*
strongly agree 1 2 3 4 5 6 7 strongly disagree

VERSION FRANCOPHONE

ECHELLE SANTA BARBARA DU SENS DES DIRECTIONS V.2

Sexe: F M

Date de passation : _____

Âge : _____

Ce questionnaire se compose de plusieurs énoncés au sujet de vos expériences, vos préférences et de vos capacités spatiales et de navigation. Après chaque affirmation, vous devez choisir un chiffre pour indiquer votre niveau d'accord avec l'énoncé. Choisissez «1» si vous

êtes fortement d'accord avec l'énoncé, choisissez «7» si vous n'êtes pas du tout d'accord avec l'énoncé, ou un nombre entre les deux selon votre degré d'accord. Choisissez le «4» si vous êtes ni d'accord, ni en désaccord.

1. *Je sais très bien indiquer un itinéraire (ex : de chez vous au magasin).*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

2. *J'ai une mauvaise mémoire des endroits où j'ai laissé des choses.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

3. *Je sais très bien évaluer les distances.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

4. *Mon «sens de l'orientation» est très bon.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

5. *J'ai tendance à penser mon environnement en termes de points cardinaux (Nord, Sud, Est, Ouest).*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

6. *Je me perds très facilement dans une nouvelle ville.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

7. *J'aime lire des cartes.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

8. *J'ai du mal à comprendre les itinéraires.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

9. *Je sais très bien lire les cartes.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

10. *Je ne me souviens pas très bien des routes quand je suis passager d'une voiture.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

11. *Je n'aime pas décrire un itinéraire.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

12. *Ce n'est pas important pour moi de savoir où je suis.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

13. *J'ai l'habitude de laisser quelqu'un d'autre planifier la navigation pour les longs voyages.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

14. *Généralement je me souviens d'un nouveau trajet après l'avoir parcouru une seule fois.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

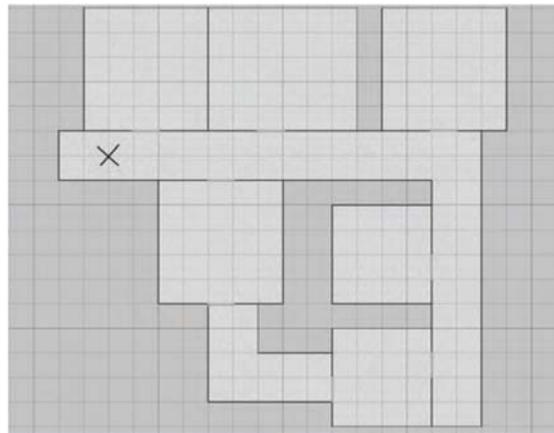
15. *Je n'arrive pas très bien à me représenter mon environnement.*

Tout à fait d'accord 1 2 3 4 5 6 7 pas du tout d'accord

Expérimentations

A. Première expérimentation

Pour la première expérimentation, seul un environnement qui contenait plusieurs trajets fut utilisé. Cet environnement était le suivant :



B. Deuxième expérimentation

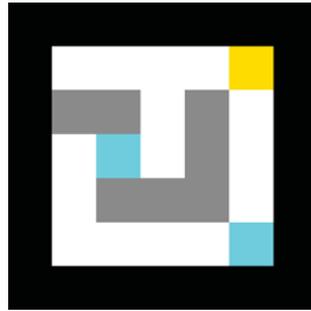
B.1. Environnements utilisés



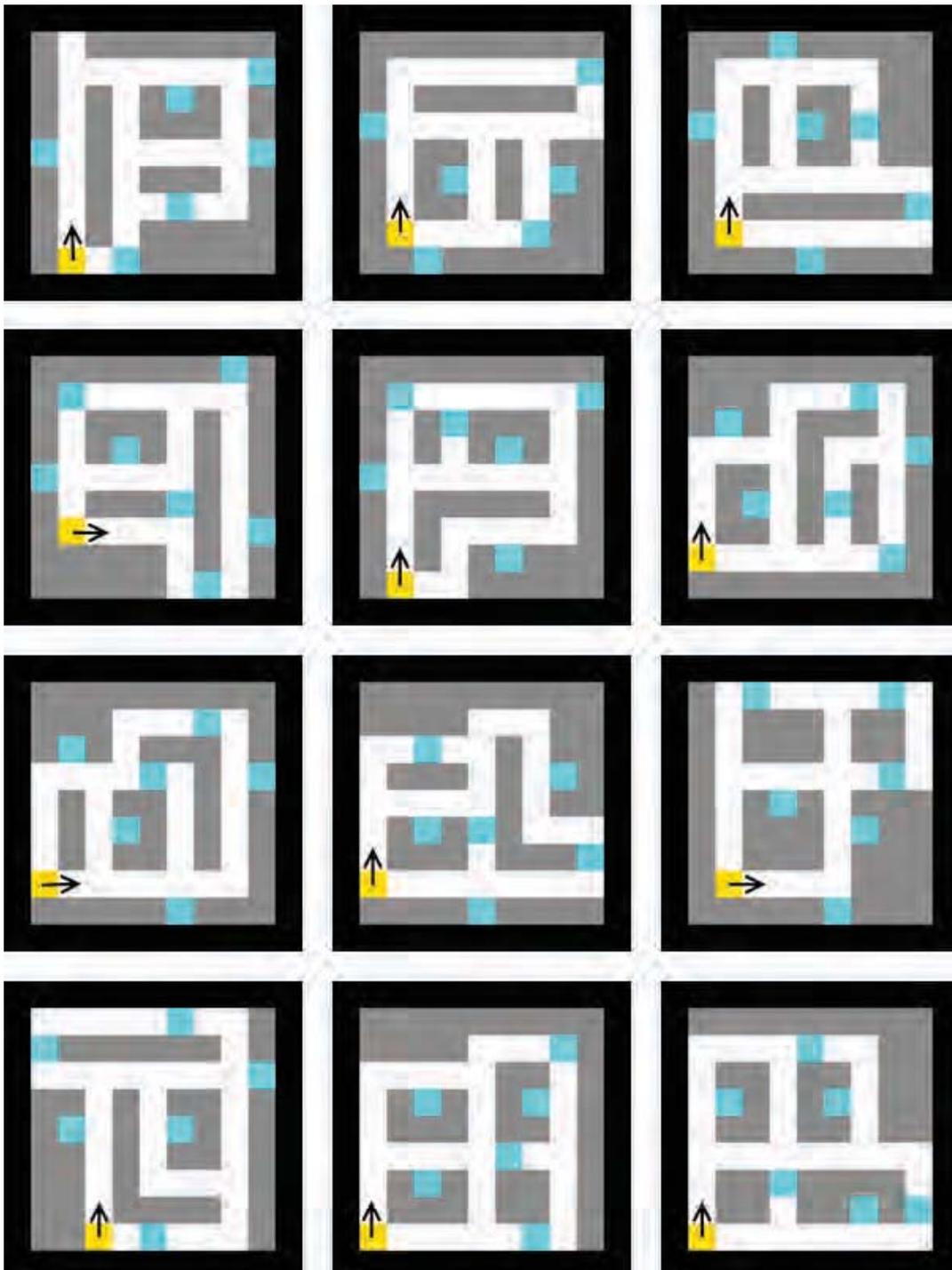
C. Troisième expérimentation

C.1. Environnement d'apprentissage

Tous les environnements ont été dessinés puis créés sous le logiciel Sketchup. Ensuite ils ont été importés sous CopperCube pour les extraire au format de scène Irrlicht. En jaune est représentée la base de départ et en bleu les joyaux à récolter par les sujets.



C.2. Environnements d'expérimentation



C.3. Consignes fournies aux juges pour l'évaluation des dessins

Ce questionnaire porte sur l'évaluation de dessins produits lors d'une expérimentation. Les sujets ont effectué une tâche de navigation dans un certain environnement et ont ensuite dessiné la carte de l'environnement tel qu'ils se le représentaient.

Les sujets ont exploré 12 environnements différents sous forme de jeu durant lequel ils devaient trouver 6 joyaux et les ramener à la base de départ. Après cette exploration, ils ont dessiné sur papier l'environnement. Les différents éléments qu'ils pouvaient indiquer étaient les suivants :

La base de départ (en JAUNE sur la représentation standard de l'environnement)

L'orientation de départ du sujet (correspondant à la FLÈCHE sur la représentation standard de l'environnement)

Les joyaux (en BLEU sur la représentation standard de l'environnement)

La structure de l'environnement (correspondant aux chemins en BLANC dans la représentation standard).

Pour chaque environnement, il vous est demandé d'évaluer la proximité entre les dessins fournis par les sujets et la représentation standard de l'environnement.

Il vous est demandé de noter chacun de ces dessins avec une note pouvant aller de 0 à 10.

0 signifie que le dessin n'a absolument rien à voir avec l'environnement, qu'il est complètement différent.

10 signifie que le dessin est conforme à l'environnement.

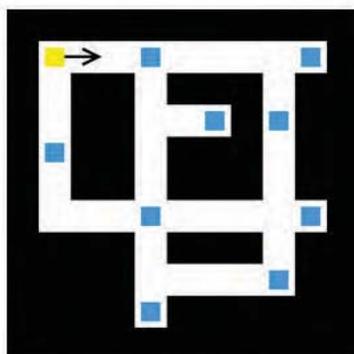
Aide pour la notation :

Différentes personnes ont dessiné, les éléments peuvent donc être représentés de façons très variées suivant les dessins. Les joyaux peuvent être représentés par des croix, des losanges, la lettre J etc.

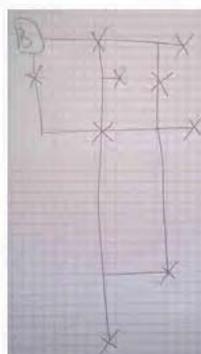
Il vous est demandé de ne pas tenir compte d'éventuelles déformations des distances :

Un dessin qui rend bien compte de la structure de l'environnement mais qui est trop allongé ou qui comporte des parties trop courtes peut avoir la note maximale (exemple 1).

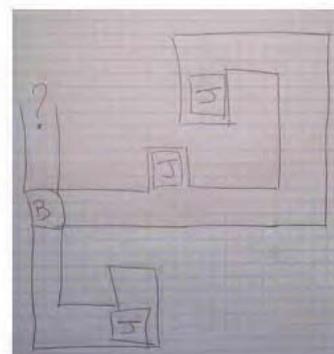
À l'inverse, un dessin qui veut rendre compte de certaines distances mais qui se trompe complètement sur la structure de l'environnement peut avoir la note minimale (exemple 2).



Représentation Standard



Exemple 1



Exemple 2

D. Quatrième expérimentation

D.1. Questionnaire soumis aux sujets

Ce questionnaire est une étude portant sur la navigation dans la ville de Toulouse.

Bonjour et bienvenue dans ce questionnaire portant sur la représentation des espaces et la navigation.

Cette étude est dirigée par Victor Vergnieux, Marc Macé et Christophe Jouffrais. Nous souhaitons étudier les représentations mentales construites lors de la navigation en environnement urbain. Plus précisément nous nous intéressons à la manière dont vous vous représentez les espaces dans lesquels vous naviguez. Ce questionnaire se concentre sur le trajet allant de la place Esquirol à la place du Capitole. Lors de ce questionnaire ce sont les représentations d'un trajet qui nous évaluons et non les participants ! Il est très important que les réponses apportées soient sincères et spontanées.

Pour participer à cette étude, il est donc nécessaire de connaître ces deux places et le trajet qui va de l'une à l'autre en passant par la rue Saint Rome. La participation à cette étude est libre et non rémunérée.

Ce questionnaire est composé de deux parties distinctes.

- La première partie contient 3 petits questionnaires. Le premier porte sur vous et vos habitudes. Le deuxième porte sur votre capacité d'imagerie mentale. Le troisième se centre sur votre sens de l'orientation.

- La seconde partie contient le cœur du questionnaire et porte sur la place Esquirol, la place du Capitole ainsi que le trajet entre ces deux places.

L'ensemble du questionnaire devrait vous prendre une trentaine de minutes.

Si vous avez des questions ou des remarques, merci de me contacter à l'adresse suivante : Victor.Vergnieux@irit.fr

En vous remerciant très chaleureusement pour votre participation.

Il y a 31 questions dans ce questionnaire

Informations sur le sujet.

1. Veuillez indiquer vos initiales dans le cadre ci-dessous.

2. Vous êtes de sexe :

Veuillez sélectionner une seule des propositions suivantes :

Féminin

Masculin

3. Quel âge avez-vous ? ans

Seuls des nombres peuvent être entrés dans ce champ.

4. Vous êtes :

Veuillez sélectionner une seule des propositions suivantes :

Non-voyant-e

Mal-voyant-e

Voyant

5. Depuis combien de temps êtes-vous non-voyant-e ?

Veillez indiquer la réponse en années.

Répondre à cette question seulement si les conditions suivantes sont réunies :

La réponse était 'Non-voyant-e' à la question '4 [Handicap]' (Vous êtes :)

Seuls des nombres peuvent être entrés dans ce champ.

6. Depuis combien de temps êtes vous mal-voyant-e ?

Veillez indiquer la réponse en années.

Répondre à cette question seulement si les conditions suivantes sont réunies :

La réponse était 'Mal-voyant-e' à la question '4 [Handicap]' (Vous êtes :)

Seuls des nombres peuvent être entrés dans ce champ.

7. Quel est le dernier diplôme que vous avez obtenu ?

Veillez sélectionner une seule des propositions suivantes :

Pas de diplôme

CEP ou CAP + CEP

Brevet des collèges

Baccalauréat

Deug ou Licence

DEA ou Master

Doctorat

Questionnaire Lieu : La place du Capitole

Les questions qui vont suivre porteront sur la place du Capitole. Il est important de noter que cette place est envisagée dans le cadre d'une navigation à pied.

- Si vous êtes non-voyant-e (respectivement mal-voyant-e), décrivez la place du Capitole pour une personne non-voyante (resp. mal-voyante) qui n'y a jamais été et qui a besoin d'y passer dans le cadre d'un trajet. Si vous décrivez cette place depuis une position précise, merci d'indiquer où se situe cette position ou comment vous y accédez.

- Si vous êtes voyant-e, merci de décrire la place du Capitole pour une personne voyante qui n'y a jamais été et qui a besoin d'y passer dans le cadre d'un trajet. Si vous décrivez cette place depuis une position précise, merci d'indiquer où se situe cette position ou comment vous y accédez.

Veillez écrire votre réponse ici :

Quel indice vous fait savoir à coup sûr que vous êtes bien arrivé sur la place du Capitole ?

Veillez écrire votre réponse ici :

Quels sont pour vous les éléments importants pour se repérer sur la place du Capitole ?

Entrez le nombre d'éléments que vous voulez (maximum : 12).

Veillez écrire votre(vos) réponse(s) ici :

Élément 1 :

Élément 2 :

Élément 3 :

Élément 4 :

Élément 5 :

Élément 6 :

Élément 7 :

Élément 8 :

Élément 9 :

Élément 10 :

Élément 11 :

Élément 12 :

Veillez indiquer par où vous accédez généralement à la place du Capitole quand vous vous y rendez à pied.

Veillez écrire votre réponse ici :

Sur une échelle de 1 à 7, à quel point pensez-vous connaître ce lieu ?

Choisissez la réponse appropriée pour chaque élément :

1 : je connais très mal ce lieu

7 : je connais très bien ce lieu

	1	2	3	4	5	6	7
Familiarité	<input type="radio"/>						

Depuis la première fois que vous êtes passé sur la place du Capitole, vous y êtes déjà retourné-e :

Veillez sélectionner une seule des propositions suivantes :

1-5 fois

6-10 fois

10-20 fois

20-50 fois

+ de 50 fois

Indiquer la fréquence à laquelle vous passez sur la place du Capitole. Si vous avez eu une période durant laquelle vous y passiez souvent, indiquer la fréquence lors de votre pic de fréquentation.

Veillez sélectionner une seule des propositions suivantes :

- Annuellement
- Plusieurs fois par an
- Mensuellement
- Plusieurs fois par mois
- Chaque semaine
- Quotidiennement

Si vous avez des précisions à faire sur votre représentation de la place du Capitole, ou des remarques sur les questions, merci de les indiquer dans ce cadre.

Veillez écrire votre réponse ici :

Questionnaire Lieu : La place Esquirol

Les questions qui vont suivre portent sur la place Esquirol. Il est important de noter que cette place est envisagée dans le cadre d'une navigation à pied.

- Si vous êtes non-voyant-e (respectivement mal-voyant-e), décrivez la place Esquirol pour une personne non-voyante (resp. mal-voyante) qui n'y a jamais été et qui a besoin d'y passer dans le cadre d'un trajet. Si vous décrivez cette place depuis une position précise, merci d'indiquer où se situe cette position ou comment vous y accédez.

- Si vous êtes voyant-e, merci de décrire la place Esquirol pour une personne voyante qui n'y a jamais été et qui a besoin d'y passer dans le cadre d'un trajet. Si vous décrivez cette place depuis une position précise, merci d'indiquer où se situe cette position ou comment vous y accédez.

Veillez écrire votre réponse ici :

Quel indice vous fait savoir à coup sûr que vous êtes bien sur la place Esquirol ?

Veillez écrire votre réponse ici :

Quels sont pour vous les éléments importants pour se repérer sur la place Esquirol ?

Vous pouvez entrer le nombre d'éléments que vous voulez (maximum : 12).

Veillez écrire votre(vos) réponse(s) ici :

- Élément 1 :
- Élément 2 :
- Élément 3 :
- Élément 4 :
- Élément 5 :
- Élément 6 :
- Élément 7 :
- Élément 8 :
- Élément 9 :
- Élément 10 :

Élément 11 :

Élément 12 :

Veillez indiquer par où vous accédez généralement à la place Esquirol quand vous vous y rendez à pied.

Veillez écrire votre réponse ici :

Sur une échelle de 1 à 7, à quel point pensez-vous connaître ce lieu ?

Choisissez la réponse appropriée pour chaque élément :

1 : je connais très mal ce lieu

7 : je connais très bien ce lieu

	1	2	3	4	5	6	7
Familiarité	<input type="radio"/>						

Depuis la découverte de ce la place Esquirol, vous y êtes déjà retourné-e :

Veillez sélectionner une seule des propositions suivantes :

- 1-5 fois
- 6-10 fois
- 10-20 fois
- 20-50 fois
- + de 50 fois

Indiquer la fréquence à laquelle vous passez sur la place Esquirol. Si vous avez eu une période durant laquelle vous y passiez souvent, indiquer la fréquence lors de votre pic de fréquentation.

Veillez sélectionner une seule des propositions suivantes :

- Annuellement
- Plusieurs fois par an
- Mensuellement
- Plusieurs fois par mois
- Chaque semaine
- Quotidiennement

Si vous avez des précisions à faire sur votre représentation de la place Esquirol, ou des remarques sur les questions, merci de les indiquer dans ce cadre.

Veillez écrire votre réponse ici :

Questionnaire Trajet

Nous allons maintenant vous demander de vous projeter mentalement sur la place Esquirol afin de faire le trajet qui vous mène jusqu'à la place du Capitole en passant par la rue Saint-Rome.

Merci d'indiquer les différentes étapes qui vous permettent d'effectuer ce trajet en étant le plus précis possible. Les éléments que vous inscrivez peuvent être de différentes nature :

- Les actions que vous menez pour effectuer le trajet
- Les éléments qui confirment que vous êtes sur la bonne voie.
- Les repères qui signalent une décision à prendre.
- Les décisions prises.

Le plus simple pour cet exercice est de s'imaginer au début du trajet et de s'imaginer le faire en notant au fur et à mesure ce qui vous vient.

Veillez écrire votre réponse ici :

Envoyer votre questionnaire.

Merci d'avoir complété ce questionnaire.

D.2. Descriptions faites par les sujets des places et du trajet.

Place du Capitole :

NV1 : C'est une grande place rectangulaire, le bâtiment de la mairie occupe tout le côté droit (largeur) les deux longueurs sont délimitées par la rue du poids de l'huile et la rue la faillette. Le côté gauche (largeur) est constitué d'un bâtiment avec des arcades où se trouvent des terrasses de bars et glaciers.

Indice permettant d'identifier le lieu : le bruit

Éléments importants pour se repérer : la sortie du parking souterrain, le bruit des terrasses, l'odeur du maco

NV2 : La place du capitole se situe entre Esquirol et la place Wilson.

Si tu arrives de la rue St Rome, il te faut aller tout droit. Tu arriveras alors en face de la place côté mairie. La mairie se trouvera sur ta droite. Attention, il y a le marché sur cette place. Si tu continues tout droit, tu peux la contourner... Aussi, en allant tout droit, tu arrives en face du Magdo, tu prendras la prochaine à droite, et monteras la rue.

Si tu arrives de la place Wilson, la place se situera sur ta gauche.

Je ne saurais te décrire précisément où se trouvent les différentes choses sur cette place. D'autant plus qu'elle a été modifiée. Il y a un jet d'eau, des jeux pour les enfants... des animations etc.

Indice permettant d'identifier le lieu : Je sais que je suis sur la place où à côté à cause de sa taille, les animations, les voix, etc.

Éléments importants pour se repérer : Le jet d'eau, les gens, les animations, la mairie

MV1 : je sais juste aller de la rue sant-rome au métro sur la gauche

Indice permettant d'identifier le lieu : la fin de la rue pavée

Éléments importants pour se repérer :

V1 : C'est une énorme place à peu près carrée avec un gros bâtiment très joli d'un côté qui doit faire 100m de long. C'est la mairie avec ses drapeaux français et occitan. On peut traverser dans son centre. En face il y a tout un tas de restaurants sous une allée d'Arcades. Il y aura peut-être un marché sur la place ou une manifestation, tout dépend de l'heure.

Indice permettant d'identifier le lieu : La mairie

Éléments importants pour se repérer : La mairie, Les Arcades, Le McDonalds, Le glacier Faur, Le Bibent

V2 : à partir de la fin de la rue St Rome, de gauche à droite première rue, une longue rue avec kebab puis (on entre dans la place) des banques et au loin d'autres magasins, et au bout, le nouveau zara à la place du virgin à côté de la station de métro «Capitole». La place en elle-même, vaste avec les 12 signes du zodiac au milieu dans l'«étoile» symbole de Midi-Pyrénées. il y a des cubes de granit tout autour de la place sauf pour l'entrée du parking souterrain qui se trouve devant vous (fin rue st Rome). la «facade» gauche, restaurant, marchand de glaces, immeuble de 5 ou 6 étages et des tables pour le café sous une allée couverte (arcades en pierre bien décorées). Façade droite, la mairie. 2 étages, c'est beau, il y a une grande porte au centre pour entrer. Façade en face, ancienne librairie et à droite un MC Do. Au font à droite, il y a une grande rue menant à la place Wilson (et qui passe à côté du zara). Et à l'angle au fond à gauche, il y a un magasin de jouet, Toys «R» Us ou la grande récréée ... je ne sais plus.

Indice permettant d'identifier le lieu : grande place vaste, 4 façades, les zodiacs, la mairie

Éléments importants pour se repérer : place vaste, carrée, 4 façades d'immeubles, zodiac, mairie, grande porte, parking, les blocs en granits, les cafés en face de la mairie, le coin bordel au fond à gauche (pt de vue fin ruest Rome).

V3 : Position devant le glacier Amorino en arrivant de la place Saint Pierre:

La place du Capitole est juste devant vous. Il s'agit d'une large place qui est entouré des 4 côtés par des bâtiments.

Sur votre côté, à l'Ouest de la place, se trouve une galerie avec des restaurants et bars. Le plafond de cette galerie démontre des peintures. Chaque restaurant et bar a une petite terrasse qui donne sur la place du Capitole. Juste après il y a deux routes pour des voitures. La deuxième mène au parking du Capitole qui se situe du côté droite sur la place.

Du coté Sud de la place se trouve à l'Ouest le magasin C&A, ensuite en allant vers le Est la rue Saint Rome, ensuite plusieurs bars et magasins.

Côté Est, il y a le bâtiment du Capitole dans lequel se situe la mairie de Toulouse. Il s'agit d'un grand bâtiment historique qui prend presque toute la largeur de la place. A droite et à gauche il y a une rue qui partent vers l'Est.

Finalement, du côté Nord, il y a deux rues qui partent vers le nord juste à gauche de vous et ensuite plusieurs magasins (Nescafé, Desigual) et ensuite un Mc Donalds.

Indice permettant d'identifier le lieu : Le bâtiment historique, le dessin du croix Occitan sur le sol de la place

Éléments importants pour se repérer : Le bâtiment historique, et son passage au milieu, Mc Donalds, C&A, La croix Occitane, Amorino

V4 : En haut de la rue Saint-Rome, on arrive sur la place du Capitole. Sur le côté droit de la place, il y a l'hôtel de ville. Au centre de l'hôtel de ville, un portail permet d'accéder à la place derrière l'hôtel de ville, qui donne sur le cours Alsace Lorraine. Sur le côté gauche, il y a des arcades sous lesquels on trouve des bars et des restaurants. Au sud, il y a des magasins. Au nord aussi, et notamment un MacDo qui sert souvent de point de rendez-vous. La rue qui part de suite à gauche permet de rejoindre la librairie Ombres blanches puis les quais. Au coin nord-ouest de la place, on peut soit aller à gauche rue pargaminières, soit aller tout droit et remonter la rue du Taur jusqu'à la basilique saint sernin.

Indice permettant d'identifier le lieu : Les drapeaux de l'hôtel de ville (on est du bon côté de la façade)

Éléments importants pour se repérer : L'hôtel de ville, les arcades, le MacDo, L'entrée de la rue Saint Rome

V5 : La place du Capitole est un grand espace dégagé, de forme rectangulaire, entouré par des immeubles de 2 étages (pas très hauts), ayant la façade du Capitole qui occupe toute une longueur d'un côté et des passages couverts avec des magasins et cafés entourant les autres 3 côtés. Si on est face au Capitole, on a à droite (sud) une entrée pour un parking souterrain, puis un grand hôtel et des magasins. A gauche (nord), il y a un McDo, un magasin Desigual, la rue du Taur qui suit en direction de l'église Saint Sernin. Derrière, au long d'un immeuble aussi long que la façade du Capitole, plusieurs restaurants et cafés avec des tables en terrasse et sous le passage couvert.

Indice permettant d'identifier le lieu : La façade du Capitole, la croix occitane marqué par terre, l'espace dégagé.

Éléments importants pour se repérer : Façade du Capitole, Arrivée via la rue du Taur, Arrivée via rue Lafayette, Vue dégagée du ciel, façade de l'immeuble face au Capitole et passage couvert, entrée du parking, Glacier Armorino

V6 : La place du capitole est une immense place carrée bordées de magasins, de cafés et d'un côté de la mairie de Toulouse. On peut y accéder par au moins 6 rues ou accès donnant sur la place. Une route à sens unique longe la place sur 3 des 4 côtés ; à un des coins de cette route se trouve l'accès à un parking sous-terrain

Indice permettant d'identifier le lieu : la taille de l'espace «vide», et son aspect carré

Éléments importants pour se repérer : La mairie, la rue du taur, la rue st rom, le Mcdo

V7 : Quand tu arrives par la petite rue a droite, si tu viens du métro, tu arrives sur la droite de la place du capitole. C'est une grande place rectangulaire, et y a rien au centre. A droite, y a des magasins, en face des terrasses de cafés.

A ta gauche y aura le capitole et une entrée de parking.

Indice permettant d'identifier le lieu : La vue du capitole

Éléments importants pour se repérer : La place, Les terrasses en face, La sortie de parking au centre de la place, Les rues sortant de la place

Place Esquirol :

NV1 : la place esquirol est un grand rectangle. Elle est traversée sur sa longueur par une voie de circulation. Sur le côté gauche se trouvent un kiosque, une terrasse de café, une petite rue pavée, le magasin midika, une boulangerie avec des tables, et la bouche de métro.

Le long du trottoire de gauche se trouvent des vlibs et des abris-bus.

Indice permettant d'identifier le lieu : la rigole au sol

Éléments importants pour se repérer : la présence de la rigole au milieu du trottoire, midica, darty

NV2 : -

Indice permettant d'identifier le lieu : -

Éléments importants pour se repérer : -

MV1 : je sais juste aller de la rue des changes à la place de la trinité

Indice permettant d'identifier le lieu : le bruit de la fontaine

Éléments importants pour se repérer : -

V1 : La place Esquirol ne ressemble pas vraiment à une place, elle est traversée par une route au milieu. Tu verras deux bouches de métro et une entrée de parking souterrain. Il y a quelques cafés, des gros platanes, un kiosque à journaux et des grosses boutiques: Monop', Darty et Midica essentiellement.

Indice permettant d'identifier le lieu : L'agencement de la place.

Éléments importants pour se repérer : Les bouches de métro, Midica, Le Monop', Le kiosque à journaux, Le crédit mutuel, Le sens de circulation des voitures

V2 : ... bon c'est long. Alors arrivée par les Carmes, à droite un café, en face une intersection, à gauche Medicca, mie caline, tisséo & station de métro, + bus, pont. en face, orange, darty, mc do, café, métro. au fond à gauche, la rue st rome (place du Cap), et à droite une autre rue qui mène sans doute vers le capitole (métro). Rue bruyante, dangereuse et je pense qu'il y a une boulangerie à droite, après la route, où on peut acheter des sandwichs

Indice permettant d'identifier le lieu : médicca, orange

Éléments importants pour se repérer : grande circulation, bus, medicca, orange, métro, mie caline, cafés, un pont dir. ouest

V3 : Devant Mc Donalds (en face de Midica):

La place Esquirol est bizarre car elle ne ressemble pas à une place. Effectivement au milieu de la place il y a la rue de Metz, une route avec une assez large fréquentation de voitures et bus qui traverse de gauche à droite. Il y a des magasins et restaurants toute autour de cette place. Il est important de savoir qu'il y a une station de métro qui arrive ici. L'entrée se situe de l'autre côté de la rue, à droite des bâtiments. Il y a plusieurs routes qui partent de cette place. Sur votre droite il y a la rue saint Rome, rue piétonne, qui part vers le Capitole. De l'autre côté de la rue de Metz, il y a le prolongement de cette rue piétonne vers les Carmes. A votre gauche, il y a également

plusieurs rues qui ramènent direction capitole. En face de vous, à gauche du magasin Midica, il y a une rue qui amène vers la station de métro Carmes.

Indice permettant d'identifier le lieu : Midica

Éléments importants pour se repérer : Midica, le musée des augustins, le métro

V4 : La place Esquirol ressemble davantage à un quai où prendre le bus qu'à une place : elle est assez petite et il y a souvent beaucoup de monde. Elle est bordée au nord par la rue de Metz, une rue assez passante, et au Sud par Midica, un magasin de bricolage/décoration. On trouve sur la place 2 ou 3 arrêts de bus desservis par de nombreuses lignes, ainsi qu'une station de Vélô Toulouse et une bouche de métro. En arrivant des Carmes, on peut traverser la place pour rejoindre la rue des Changes ou tourner à droite sur la place pour retrouver la rue commerçante Alsace Lorraine.

Indice permettant d'identifier le lieu : Le magasin Midica

Éléments importants pour se repérer : Le magasin Midica, Le début de la rue des changes, La bouche de métro, La station de vélib

V5 : En arrivant par la rue des Filatiers, on est à coté du grand magasin Midica, à droite. L'entrée du metro Esquirol est à droite, face à des cafés, et en suivant cette direction on arrive à François Verdier. Si on suit à gauche, on a le pont Neuf. En traversant la rue de la Place Esquirol, il y a un arrêt de bus, un café, un magasin Darty, un McDo. Tout droit après la rue des filatiers, il y a la Rue des changes qui suit jusqu'à la place du Capitole. De l'autre coté, la rue des tourneurs aussi. Il y a beaucoup de circulation de voitures, bus, vélos et piétons.

Indice permettant d'identifier le lieu : le croisement de la rue des Filatiers avec la rue de Metz

Éléments importants pour se repérer : le sens unique de la rue de Metz, la rentrée du métro, le rdc de Midica, la vue sur le musée des Augustins, la vue sur le pont Neuf, l'arrêt de bus Esquirol, le croisement avec la rue des tourneurs, le croisement avec la rue des changes,

V6 : La place esquirol porte mal son nom puisque c'est plus un lieu qu'une place : elle est traversée par une avenue à sens unique qui comporte (au moins dans mes souvenirs) 3 voies de circulation. De part et d'autre de la voie on retrouve un accès à la station de métro, des grandes enseignes commerciales et des arbres. On accède à la place principalement par le croisement avec la rue saint rom qui coupe perpendiculairement la place, mais de nombreuses petites rues permettent d'y accéder également.

Indice permettant d'identifier le lieu : Darty et midica

Éléments importants pour se repérer : les accès au métro, Midica, rue saint rom, rue des filatiers, darty

V7 : C'est une sorte d'artere un peu élargie ou y a plein de magasins et de gens qui passent, et y a une station de métro,

Y a un macdo je pense dans le coin

Indice permettant d'identifier le lieu : Le métro

Éléments importants pour se repérer : L'arret de métro

Trajet entre la place Esquirol et la place du Capitole :

NV1 : je longe la place esquirol sur son côté gauche en suivant la rigole au sol

Je contourne la bouche de métro par la gauche

je tourne à droite pour aller chercher le passage piéton

j'appuie sur la télécommande pour activer le feu sonore

je traverse et continue tout droit

sous mes pieds la rue est pavée il n'y a pas de trottoir

très vite il n'y a plus de bruit de circulation de voitures

la rue n'est pas très large j'entend les bruits se répercutés sur les murs de droite et gauche.

Je fais attention parce qu'à un moment ça s'élargit un peu sur la gauche alors je prend garde à conserver le mur sur ma droite.

je me fais bousculer et je bouscule en retour ces connars qui ne regardent pas où ils marchent.

au bout de la rue saint-rome l'écho est différent,

j'entends la circulation des véhicules de la droite vers la gauche,

je me retrouve en plein sur la sortie du parking souterrain.

je sais que je suis arrivé.

NV2 :

MV1 : quand je pars des carmes, je prends la rue des filatiers, je suis sur la place Esquirol au croisement après la fontaine, puis je dois continuer tout droit pour arriver sur la rue st-rome au bout de laquelle il y a la place du capitole

V1 : Je suis devant Midica sur la place Esquirol et je décide d'aller sur la place du Capitole. Je repère l'entrée de la rue Saint Rome grâce au Crédit mutuel.

Je me dirige vers le passage clouté qui permet de traverser la rue sans se faire écraser. Je passe devant la bouche de métro et arrive au passage clouté. Quand il n'y a ni voiture, ni bus, ni vélo je traverse et j'arrive devant les distributeurs du crédit mutuel à ma gauche, et les tables d'une brasserie louche à ma droite. Je continue tout droit sur rue Saint Rome. Je passe devant quelques boutiques de vêtements pas chers ou de grandes chaînes. Je continue sur les pavés en évitant le flux de personnes très nombreux avec des cabas, et les vélos qui osent s'aventurer dans cette rue. Les bâtiments revêtent leur plus beau rose de la brique Toulousaine, et beaucoup de Toulousains revêtent quant à eux leurs habits les plus fashion.

Je croise ensuite des boutiques de percings, de décoration, de chaussures. La rue est étroite et tout le monde est serré.

V2 : Métro Carmes, tout droit, puis traverser le carrefour à Esquirol, puis à gauche, passer devant orange et darty, puis tourner à la première rue à droite, c'est la rue st rome. Tout droit, ça pue, il y a plein de boutique, de gens et de pavés, et à la fin, vous arrivez à la place du Capitole

V3 : Je suis donc devant le Mc Donalds, je tourne à droite et je suis la rue de Metz jusqu'à ce que j'arrive au croisement avec la rue Saint Rome. A gauche il y a un feu piéton. A droite donc la route piétonne qui s'appelle Rue Saint Rome. Je tourne à droite et je passe dans cette rue. Il s'agit d'une rue bien animée avec des magasins (essentiellement vêtements, chaussures) à

droite et à gauche. Une fois sur cette rue, il n'y a pas grande chose à faire que d'avancer droit. Le trajet dure environ 7 minutes. A la fin de la rue juste avant la place du Capitole, il y a à gauche le magasin Pylones (reconnaissable par sa façade verte) et ensuite à gauche également C&A. Quand on a passé ce magasin, on est sur la place du Capitole.

V4 : J'arrive sur la place Esquirol via la rue des Tourneurs. Un coup d'oeil au magasin Midica me confirme que je ne me suis pas trompée de rue. Je longe le magasin, et passe entre le magasin et la station de Velib. Je traverse la rue au passage piéton devant la boutique «A m'en donné». Je traverse la rue et vérifie éventuellement que je suis bien sur la bonne rue (rue des changes). Il faut alors remonter la rue jusqu'à ce qu'on arrive sur la place. Au milieu de la rue il y a une sorte de «virage», un «décalage» mais il n'y a pas de risque de se tromper de rue. La rue est facilement reconnaissable : des pavés, beaucoup de magasins, beaucoup de monde, pas du tout agréable. En haut de la rue il y a sur la gauche le magasin C&A et la Grande Récré (ou Joué Club ?). Il suffit alors de traverser la route pour arriver sur la place du Capitole.

V5 : Depuis la place Esquirol, en laissant Midica derrière et de l'autre côté de la rue, suivre la rue des Changes, plus proche de la Garonne que la rue des tourneurs, parallèles. La rue des Changes est une rue piétonne, étroite, avec beaucoup de boutiques. A la moitié du chemin entre Esquirol et la place du Capitole, la rue des Changes change de nom. Après le croisement avec la rue Peyras, à droite, la rue est nommée rue Saint Rome. On suit cette rue, aussi étroite et avec des boutiques des deux côtés, encore tout droit et on arrive à la place du Capitole par le sud, ayant la façade du Capitole à droite. Compter environ 10 min de marche.

V6 : En arrivant depuis la rue des filatiers : c'est tout droit ...

Au bout de la rue on arrive à un feu rouge piéton qui permet de traverser l'avenue coupant la place esquirol

En face on doit faire un petit droite=>gauche pour arriver sur la rue piétonne (il faut passer les plots escamotables qui interdisent le passage aux voitures).

Ensuite on va tout droit, on passe quelques magasins. On rencontre un croisement ; là il faut faire un petit gauche=>droite et y'a un magasins de vêtements un peu style cowboy (je crois qu'il a fermé il y a quelques temps par contre).

Ensuite c'est toujours tout droit jusqu'à la place du capitole ; on sait qu'on est sur la bonne voie avec les pavés de la rue qui forme un V si on regardait une coupe transversale de la rue

V7 : A partir de la rue des changes, on va tout droit, dans une petite rue sur des pavés, la rue est assez fréquentée, y a des magasins a gauche et a droite,

Y a une affiche pour le musée de l'accompagnement (ou un nom comme ca)

Y a des ruelles a gauche qui sortent, et une rue plus large vers la droite, devant laquelle y a souvent un petit stand qui fait des crepes

On continue tout droit et on arrive éventuellement au capitole via la rue saint-rome, y a pas de décision a prendre parce qu'on va tout droit.

D.3. Résultats du questionnaire

Ce questionnaire était composé de deux parties principales. La première partie portait sur les capacités des sujets et se décomposait elle-même en trois sous-parties :

Une première sous-partie contenait des questions relatives à l'âge et la durée de la cécité des participants le cas échéant.

La deuxième sous-partie contenait le questionnaire de vivacité d'imagerie portant sur les capacités d'imagerie mentale des sujets. Les capacités testées ici étaient visuelles autant que cutanées ou kinesthésiques (voir [Lebaz 2011]).

La dernière sous-partie était une version française du test d'auto-évaluation du sens de l'orientation de Santa Barbara [Hegarty et al. 2002].

La seconde partie portait sur la navigation entre deux places de Toulouse et était composée de trois parties :

Dans la première, il était traité de la place du Capitole, les sujets étaient invités à décrire cette place. Il leur était demandé d'indiquer l'élément caractéristique de cette place qui leur permettait de la reconnaître. Il leur était également demandé de lister des éléments de la place qu'ils utilisaient pour se repérer. Les sujets étaient aussi invités à indiquer leur fréquentation de cette place avec l'accès qu'ils utilisaient le plus.

Dans la deuxième sous-partie, il était demandé aux sujets de décrire la place Esquirol qui était également située dans le centre de Toulouse. De la même façon que précédemment, les sujets étaient invités à indiquer l'élément caractéristique de cette place ainsi que des points de repères. La fréquentation et les accès privilégiés étaient également demandés.

Enfin dans la troisième partie, les sujets décrivaient librement les étapes qui leur permettaient de se rendre de la place Esquirol à la place du Capitole.

D.3.a. Sujets

Le questionnaire a été rempli en ligne par tous les sujets sans présence de l'expérimentateur. Trois personnes non-voyantes et malvoyantes de l'Institut des Jeunes Aveugles de Toulouse ont accepté de le remplir. Parmi les personnes non-voyantes une était aveugle de naissance (sujet NV1) tandis que l'autre était non-voyante depuis l'âge de 13 ans (sujet NV2). Le sujet NV2 a eu un souci au moment de passer le questionnaire et n'a pas pu le remplir entièrement. Ainsi ses données manquent pour la description de la place Esquirol et le trajet. La personne malvoyante l'était depuis l'âge de 17 ans (sujet MV1). Sept personnes voyantes ont également rempli le même questionnaire (sujets V1 à V7). Les sujets voyants ont été recrutés via un appel sur les réseaux sociaux. Les sujets avaient entre 23 et 42 ans lors de la passation et tous demeuraient à Toulouse.

D.3.b. Résultats

Les réponses des sujets sont présentées en trois parties. La première donne les descriptions de la place du Capitole qu'ils ont produit, la deuxième traite des descriptions de la place Esquirol. Enfin la troisième partie traite du trajet entre ces deux places et de quelle manière les sujets l'ont décrit.

D.3.c. Description de la place du Capitole

La place du Capitole est la place centrale de la ville de Toulouse, le bâtiment de l'hôtel de ville la borde.

Les deux sujets non-voyants et le sujet malvoyant ont cité en moyenne quatre éléments caractéristiques de la place alors que les sujets voyants en ont cité en moyenne huit. Le sujet MV1 a indiqué qu'il ne connaissait que le trajet allant de la rue Saint-Rome au métro sans plus de détail. Le sujet NV1 a indiqué les routes bornant la place. Les sujets voyants ont tous indiqué qu'il s'agissait d'une large place, puis ont donné des éléments présents autour de la place ainsi que la présence de la croix occitane au centre. De même ils ont tous indiqué que l'élément qui leur permettait de reconnaître la place était la façade de la mairie. Deux des sujets (NV1 et V6) ont décrit la place de manière allocentrée ou depuis un emplacement non explicité par eux. Tous les autres sujets se sont placés sur la place pour la décrire. Les sujets voyants ont décrit la place depuis une position fixe, deux d'entre eux (V1 et V5) se sont placés face à la mairie, deux autres à la sortie de la rue Saint-Rome (V2 et V4), un sujet a décrit la place depuis le glacier Amorino et un dernier s'est placé à l'angle de l'hôtel de ville et de la rue Lafayette. Un sujet non-voyant (NV2) a utilisé un déplacement pour décrire ce qu'il connaissait de la place. Le trajet qu'il a décrit lui permettait de traverser la place en évitant le bâtiment de la mairie. À l'angle de l'hôtel de ville et de la rue Lafayette, ce sujet a indiqué qu'il était possible de continuer tout droit pour se rendre à un McDonald ou de tourner à droite pour se diriger vers la place Wilson (non visible sur la carte). Le sujet malvoyant MV1 a seulement indiqué qu'il ne connaissait pas des éléments précis de cette place mais était capable de se rendre de la rue Saint-Rome au métro Capitole.

Tous les sujets voyants ont indiqué le bâtiment de l'hôtel de ville et la structure de la place comme les éléments permettant de reconnaître la place. Les deux sujets non-voyants ont indiqué reconnaître la place grâce au bruit ambiant et à l'animation présente. Le sujet malvoyant a seulement indiqué que la fin de la rue pavée lui permettait de savoir qu'il était au niveau de la place du Capitole.

D.3.d. Description de la place Esquirol

La place Esquirol a la particularité d'être traversée par une voie routière sur laquelle passe de nombreux bus. Ainsi elle est séparée en deux en son milieu ce qui fait que certaines personnes ne la considèrent pas comme une place. Les sujets V1, V3, V4, V6 et V7 ont indiqué que ce lieu n'était pas réellement une place. Pour la description, une personne non-voyante (NV1) a décrit cette place de manière allocentrée en indiquant qu'une rigole au sol lui servait de repère et lui permettait de traverser cette place. Une personne non-voyante n'est pas parvenue à cette partie du questionnaire (NV2), le sujet malvoyant (MV1) n'a pu indiquer la place en question car il ne la connaissait pas à proprement parler mais savait qu'il devait la traverser pour rejoindre la place des trinités situés au sud de la place Esquirol (non affichée). Pour les sujets voyants, quatre sujets sur sept ne se sont pas placés à un endroit particulier et ont utilisé un point de vue allocentré (V1, V4, V6 et V7). Deux sujets ont décrit la place depuis un point

particulier (V2, V3). Enfin un dernier sujet voyant a utilisé deux points de vue différent pour décrire la place (V5).

Les éléments qui ont été indiqués par les sujets voyants comme caractérisant la place Esquirol sont l'agencement de la place (V1), l'enseigne Midica (V2, V3, V4), la bouche de métro (V7), l'enseigne Darty (V6), la rue des Filatiers (V5). Le sujet non-voyant a caractérisé la place Esquirol par la rigole lui permettant de la traverser et d'y naviguer.

D.3.e. Navigation dans la rue Saint-Rome

Les sujets ont décrit le trajet menant de la place Esquirol à la place du Capitole en passant par la rue des Changes puis la rue Saint-Rome. Comme la place Esquirol est plutôt complexe avec une rue à traverser et de nombreux éléments qui la peuplent comme des arrêts de bus etc., certains sujets ont décrit le trajet au sein de la place Esquirol pour rejoindre la rue des Changes. Cependant tous les sujets n'ont pas fait cela, et ceux qui l'ont fait n'avaient pas forcément tous le même point de départ. Ainsi le trajet décrit ici est le trajet commun décrit par les sujets. Neuf sujets sur dix ont décrit le trajet, un sujet (NV2) n'ayant pas compléter le questionnaire jusqu'au bout suite à un souci technique. Parmi les sujets voyants, quatre sujets sur sept ont décrit la traversée de cette rue en une seule étape (V1, V2, V3, V7), sans indiquer d'étapes intermédiaires. Les trois autres sujets ajoutaient une étape supplémentaire au milieu du trajet. Cette étape est liée à l'endroit où cette rue change de nom, le carrefour entre les deux rues n'étant pas complètement droit. Ainsi ces trois sujets ont parlé d'un décalage (V4), d'un croisement (V6) ou du changement de nom (V5). Pour décrire la fin du trajet, quatre sujets voyants (V2, V5, V6 et V7) indiquaient qu'ils étaient directement rendus au Capitole. Les trois autres sujets voyants décrivaient des repères qui leur permettaient de savoir qu'ils approchaient du Capitole ou qu'ils y étaient arrivés. Ces repères pouvaient être un élargissement de la rue (V1), les magasins Pylônes (V3) ou C&A (V3, V4) ainsi que le magasin de jouet à l'angle de la rue Saint-Rome et du Capitole (V4).

La personne non-voyante qui a complété le questionnaire jusqu'au bout a indiqué des éléments qui lui confirmaient qu'elle était engagée dans la rue des Changes. Ces éléments étaient les pavés au sol, l'absence de trottoirs, la disparition du bruit de la circulation, de plus elle entendait les bruits se répercuter sur les murs de gauche et droite ce qui lui permettait de confirmer que cette rue était étroite. Ensuite la personne indiquait qu'un élargissement sur la gauche de la rue était présent et qu'il était donc nécessaire de suivre le mur droit de la rue. Ensuite, elle détectait la fin de la rue par le changement de l'écho et le bruit de la circulation qui reprenait. Ce sujet considérait qu'il était arrivé à destination lorsqu'il atteignait la sortie du parking souterrain. Le sujet malvoyant MV1 indiquait seulement qu'il avait besoin d'avancer tout droit pour parcourir cette rue et se rendre à proximité de la place du Capitole.

Parmi les sept personnes voyantes interrogées, quatre décrivaient le trajet en une seule étape et trois utilisaient deux étapes en indiquant comme repère intermédiaire le lieu de changement de nom de la rue. La personne non-voyante qui a répondu à cette question a décrit également le trajet en deux étapes et soulignant un « élargissement à gauche » qui semble être le même point intermédiaire que celui des voyants