



**HAL**  
open science

# Conception et réalisation d'un système multi-fonctionnel d'aide à la mobilité pour personnes mavyantes et aveugles.

Jean Connier

## ► To cite this version:

Jean Connier. Conception et réalisation d'un système multi-fonctionnel d'aide à la mobilité pour personnes mavyantes et aveugles.. Intelligence artificielle [cs.AI]. Université Clermont Auvergne [2017-2020], 2019. Français. NNT : 2019CLFAC097 . tel-02926122

**HAL Id: tel-02926122**

**<https://theses.hal.science/tel-02926122>**

Submitted on 31 Aug 2020

**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.

# Université Clermont Auvergne

École Doctorale Sciences pour l'Ingénieur de Clermont-Ferrand

Thèse présentée par

Jean Connier

pour obtenir le grade de Docteur d'Université  
Spécialité informatique

Conception et réalisation d'un système multi-fonctionnel d'aide à la  
mobilité pour personnes malvoyantes et aveugles

Soutenue publiquement le 13 décembre 2019 devant le jury composé de

M. HOU Kun Mean	Professeur des Universités	Directeur de thèse
M. JAQUET Daniel	Président de l'association GAIPAR	(invité)
Mme LI Jian-Jin	Maître de conférences	Co-encadrante
M. MIGUET Serge	Professeur des Universités	Examineur
Mme PISSALOUX Edwige	Professeur des Universités	Rapporteur
M. QUILLIOT Alain	Professeur des Universités	Examineur
M. VASLIN Philippe	Maître de conférences	Co-encadrant (invité)
M. de VAULX Christophe	Maître de conférences	Co-encadrant (invité)
M. ZHOU Haiying	Professeur des Universités	Rapporteur



# Résumé

Malvoyance et cécité sont sources d'importantes difficultés de mobilité chez les personnes qu'elles touchent. Pour tenter d'alléger la charge que font peser ces difficultés, des dispositifs d'assistance variés ont été imaginés, conçus, testés, et parfois adoptés. La conception de tels dispositifs d'assistance à la mobilité se heurte à l'ampleur de la problématique, située à l'intersection de trois domaines déjà individuellement complexes : malvoyance, mobilité, et perception.

L'analyse de plusieurs types de dispositifs d'assistance – cannes classiques, électroniques, dispositifs urbains, dispositifs non directement dédiés à la mobilité – nous a aidé à faire ressortir plusieurs angles d'approche, selon les caractéristiques techniques, les fonctionnalités, la forme, et la dépendance à une infrastructure. Pris ensemble, ils ont permis de dresser un portrait général des dispositifs existants. À côté de ces approches assez classiques, nous proposons un modèle d'analyse des dispositifs selon la manière dont ils s'insèrent dans le processus perception / mobilité des personnes. Ce modèle présente l'intérêt de pouvoir s'appliquer, *a priori*, à l'ensemble des dispositifs et d'être à la fois pertinent dans leur évaluation et leur classification.

Nous avons conçu et développé un dispositif d'assistance électronique, le système 2SEES reposant, comme son prédécesseur le SEES, sur trois plates-formes : une canne intelligente, un smartphone, et un serveur. Les problématiques d'autonomie et d'universalité, primordiales dans tout dispositif électronique mobile, sont explicitées et approfondies. Une nouveauté est la prise en compte des problématiques de robustesse : le 2SEES est donc construit autour de l'équilibre complexe entre autonomie, universalité, et robustesse, notions relativement peu visibles dans les travaux existants. Pour augmenter à la fois la robustesse et l'autonomie, nous avons intégré un nombre important de capteurs et de processeurs dans la canne, à la fois par la duplication de composants, de façon à augmenter la tolérance aux pannes, et par l'intégration de composants hétérogènes, afin d'augmenter la robustesse face à la diversité des environnements.

Deux aspects de cet équilibre ont été plus spécifiquement étudiés. D'une part, la nécessité de robustesse a été mise en évidence par une analyse des affinités entre plusieurs types de capteurs d'obstacles et différents matériaux constituant potentiels d'obstacles. D'autre part, nous avons développé un prototype de fonctionnalité de localisation peu dépendante d'infrastructures, et donc déployable rapidement, ainsi qu'économe en énergie. Cette fonctionnalité repose principalement sur des capteurs embarqués (encodeur de roue, capteurs inertiels) et sur un filtre particu-

laire simplifié, qui estime la position de la personne en vérifiant la cohérence de trajectoires dérivées des données capteurs avec la carte du lieu.

Outre la recherche de cet équilibre entre robustesse, autonomie, et universalité, nous avons développé une fonctionnalité novatrice de communication avec les objets intelligents, nommée SO2SEES. Cette fonctionnalité permet à l'utilisateur de poser des questions en langage naturel au système 2SEES, auxquelles ce dernier répond en exploitant les informations mises à disposition par les objets intelligents environnants. Afin de simplifier le système, la personne ne pose pas ses propres questions, mais les sélectionne dans un ensemble proposé en fonction du contexte formé par les objets présents et les données qu'ils mettent à disposition. Ce choix entraîne un basculement d'un système de compréhension du langage naturel, fonctionnant classiquement grâce à l'apprentissage machine, vers un système expert travaillant sur des bases de connaissances distribuées et dynamiques. Dans cette configuration, la problématique principale est l'interopérabilité de SO2SEES avec les autres acteurs impliqués dans la fonctionnalité, les objets intelligents et leurs plates-formes.

# Abstract

Visual impairment and blindness are sources of mobility difficulties for the affected people. In order to lighten the burden of these difficulties, many mobility aids have been imagined, designed, tested, and more or less adopted. Designers of such assistive systems soon run into the complexity of the issue, which stands at the intersection of three domains that are, by themselves, complex : visual impairment, mobility, and perception.

Having decided to design an *electronic* assistive system from the beginning, we tried to step back and analyzed a wide range of blind aids : white canes, mobile electronic devices serving different mobility purposes, urban systems, and systems not primarily designed for mobility. Their diversity helped us analyze assistive systems through several prisms: according to their technical characteristics, their functions, their shape, and their dependence to an infrastructure. Taken individually, each of these approaches quickly shows its limits, but, together, they draw an interesting portrait of the existing devices. Besides these rather classical approaches, we propose a new model for analyzing assistive systems, which relies on the way these systems take place in a person's perception / mobility process. This model has the advantages of being, a priori, relevant for all assistive systems – in spite of their dissimilarity – and meaningful for both evaluation and classification.

We have designed and built an electronic mobility aid, called the 2SEES system. Like its predecessor, the SEES system, the 2SEES relies on three platforms : a smart cane, a smartphone, and a cloud computing back-end. The issues of energy consumption and geographical universality, essential for any mobile device, are made explicit and studied. A novelty of the 2SEES resides in its account of reliability issues ; it is thus designed around the complex equilibrium between energy consumption, universality, and reliability. Despite their importance in a system destined to be adopted by end users, these three notions are scarcely visible in the relevant literature. To concurrently enhance robustness and autonomy, we have integrated several sensors and processors in the smart cane, both by introducing redundancy, for fault tolerance, and by integrating heterogeneous sensors, for robustness against the diversity of environments.

Two aspects of this equilibrium have been further studied. First, the need for robustness has been highlighted by a study of affinities between obstacle sensors and several types of potential obstacle materials. Secondly, we have tried to develop an energy-efficient indoor localization function that is little dependent on infrastructures, and therefore easily scalable. This function works with embedded sensors

(wheel encoder, inertial measurement unit) and a simplified particle filter, which estimates the position by checking the coherence of trajectories derived from sensor data against the map of the location.

In addition to this work on the balance between robustness, energy consumption, and universality, we have developed a novel function, named SO2SEES, which allows communication between users of the 2SEES and smart objects. This function enables users to ask, in natural language, questions to the 2SEES, which are answered using information coming from surrounding smart objects. In order to keep the system simple, users do not formulate their own questions, but are instead invited to pick them in a set of predefined questions, which are proposed by the system according to the nearby objects and the information they offer. This mode of operation shifts the system from a natural language processing artificial intelligence to an expert system working on dynamic and distributed knowledge bases. In this latter configuration, the main issue is the interoperability between the 2SEES and the smart objects and their back-ends that take part in the functionality.

# Remerciements

Je tiens d'abord à remercier mes encadrants Kun Mean Hou, Jian-Jin Li, Philippe Vaslin et Christophe de Vault, pour m'avoir accompagné et conseillé tout au long de cette thèse. Les discussions que nous avons eu ont été très enrichissantes. Merci aussi à Daniel Jacquet pour ses points de vue toujours pertinents.

Je tiens à remercier chaleureusement Serge Miguet, Edwige Pissaloux, Alain Quilliot, et Haiying Zhou pour avoir accepté de faire partie du jury, malgré des délais souvent peu raisonnables, et pour leurs remarques et commentaires.

Merci aux membres de l'équipe SMIR, en particulier à Hongling, pour son aide indispensable, sa connaissance infinie, et bien sûr, sa gentillesse.

Pour leur disponibilité, leur efficacité, leur patience, et surtout leur très grande gentillesse, merci à Béatrice, Séverine, Bastien, Nicolas, Fabrice, Stéphanie, Bruno, William, Aure, Martine, Stéphanie, Annie, Mme. Torrisani et aux autres membres du personnel du laboratoire, de l'ISIMA, et de l'université que j'oublie.

Merci à « mes » étudiants pour leur implication, leur patience, leurs questions et retours !

Je remercie l'ensemble de mes collègues, amis, et entourage pour ces quelques années. Dans un ordre en partie géographico-historique, je remercie donc Sahar, Marina, Théo, Caroline, Karima, Khaled, Amina, Antoine, Vanel, Jinhua, Danh, Suan, Jocelyn, Ousmane, Franklin, Toan, Benoit, Ruben, Mamadou, Libo, Xiongmin, Yibo, Tian Bin, Sun Jie, Xiaojing, Zhao Yan, Xu Shu, Gao Hui, Ye Lihua, Zhengze, Julian, Gil, Duy, Miao, Di, Qing, Hua, Denis, Chao, Isabelle, Pascaline, mes parents, Soline, Jeff. Alexis et Carlos, pour les trajets à vélo et les expéditions nocturnes, les repas de midi et la cuisine du soir, la vaste demeure, les discussions. Merci à Luxuan, pour sa sensibilité, sa révolte feutrée, sa singularité.



# Table des matières

<b>Résumé</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Remerciements</b>	<b>vii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Mobilité, malvoyance, et perception</b>	<b>3</b>
1.1 Mobilité . . . . .	3
1.1.1 Modèle de base . . . . .	4
1.1.2 Modèle étendu . . . . .	5
1.2 Malvoyance et cécité . . . . .	6
1.2.1 Définition et symptômes . . . . .	8
1.2.2 Pathologies et prévalences respectives . . . . .	8
1.2.3 Ciblage . . . . .	10
1.2.4 De la problématique à la perception . . . . .	12
1.3 Perception . . . . .	13
1.3.1 Progression historique . . . . .	14
1.3.2 Problématiques centrales pour la perception des PMV . . . . .	20
<b>2 Dispositifs artificiels, état de l'art</b>	<b>27</b>
2.1 Introduction . . . . .	27
2.1.1 Historique de l'assistance aux personnes déficientes visuelles	27
2.2 Aides électroniques à la mobilité . . . . .	30
2.2.1 Cannes électroniques « simples » . . . . .	30
2.2.2 Supplémentation, substitution sensorielle . . . . .	34
2.2.3 Implants rétiniens . . . . .	41
2.2.4 Aides à la navigation . . . . .	43
2.2.5 Aides basées sur la vision . . . . .	46
2.2.6 Aides fonctionnellement basées sur des objets dispersés dans l'environnement . . . . .	51
2.2.7 Aides multifonctionnelles . . . . .	55
2.3 Discussion . . . . .	56
2.3.1 Analyse par rapport aux fonctionnalités . . . . .	57
2.3.2 Technologie des capteurs . . . . .	58
2.3.3 Forme et port . . . . .	62

2.3.4	Interface utilisateur . . . . .	64
2.3.5	Fiabilité . . . . .	64
2.3.6	Infrastructures . . . . .	65
2.3.7	Analyse par rapport au positionnement perceptif . . . . .	68
2.3.8	Résumé par rapport à la problématique de la thèse . . . . .	74
<b>3</b>	<b>Développement de la canne intelligente 2SEES</b>	<b>79</b>
3.1	Développement du SEES . . . . .	79
3.1.1	Motivation et concept . . . . .	79
3.1.2	Plate-forme . . . . .	82
3.1.3	Fonctionnalités et expérimentations . . . . .	85
3.2	Système 2SEES et architecture . . . . .	92
3.3	Conception et réalisation de la plate-forme matérielle . . . . .	97
3.3.1	Forme et conséquences . . . . .	97
3.3.2	Conception de la plate-forme mécanique . . . . .	98
3.3.3	Conception de la plate-forme électronique . . . . .	102
3.3.4	Système d'exploitation . . . . .	108
3.4	Conception et implémentation de la fonctionnalité de navigation . . . . .	115
3.4.1	Problématique . . . . .	115
3.4.2	Filtrage particulière . . . . .	118
3.4.3	Filtrage particulière simplifié . . . . .	122
3.5	Conception, implémentation et test de la fonctionnalité de détection d'obstacles . . . . .	125
3.5.1	Test des affinités entre capteurs et matériaux . . . . .	127
3.5.2	Discussion . . . . .	130
3.5.3	Fonctionnement global du 2SEES . . . . .	132
<b>4</b>	<b>Objets intelligents pour la perception : SO2SEES</b>	<b>135</b>
4.1	Introduction et concepts préliminaires . . . . .	135
4.1.1	Assistance à la perception de l'environnement . . . . .	135
4.1.2	Objets connectés . . . . .	137
4.2	SO2SEES . . . . .	143
4.2.1	Concept . . . . .	143
4.2.2	Architecture globale et fonctionnement . . . . .	144
4.2.3	Fonctionnalités, requêtes, et règles . . . . .	146
4.3	Scénarios de test . . . . .	157
4.3.1	Matériel utilisé . . . . .	158
4.3.2	Implémentation logicielle . . . . .	159
4.3.3	Scénarios de test . . . . .	162
4.4	Discussion . . . . .	169
4.4.1	Analyse du contexte . . . . .	169
4.4.2	Bases de connaissances . . . . .	170
4.4.3	Requêtes SQWRL . . . . .	170
4.4.4	Gestion des erreurs . . . . .	172
4.4.5	Problématiques d'interopérabilité . . . . .	175
4.4.6	Apports perceptifs . . . . .	180
	<b>Conclusion</b>	<b>183</b>





# Table des figures

1.1	Lois de groupage de la <i>Gestaltpsychologie</i> [Rock and Palmer, 1990].	17
2.1	Reproduction d'une planche de la <i>Dioptrique</i> de Descartes dans la <i>Lettre sur les aveugles</i> de Diderot; la planche daterait de 1724 [Diderot, 2011] (image tirée de l'édition en ligne disponible sur le site de la BNF : <a href="https://gallica.bnf.fr/essentiels/diderot/lettre-sur-les-aveugles">https://gallica.bnf.fr/essentiels/diderot/lettre-sur-les-aveugles</a> ). . . . .	28
2.2	Le capteur d'obstacles <i>Pathsounder</i> , à gauche, se porte en collier sur la poitrine; l'appareil visible à droite est son chargeur [Russel, 1965].	31
2.3	Le miniguide [Research, 2005]. . . . .	32
2.4	Un Teletact (capteur laser) monté sur une canne blanche [Farcy et al., 2006]. . . . .	33
2.5	Caméra et fauteuil instrumentés utilisés pour le système de substitution sensorielle de Bach-y-Rita <i>et al.</i> [Bach-y Rita et al., 1969]. . .	35
2.6	Hsu <i>et al.</i> utilisent un algorithme de filtrage d'images pour faire ressortir les objets (à gauche, l'image en résolution diminuée et niveaux de gris; à droite, la sortie de l'algorithme) [Hsu et al., 2013]. . . . .	36
2.7	Le système de Kajimoto, Suzuki et Kanno. En haut, une photo du dispositif en cours d'utilisation; en bas, une représentation schématique de son fonctionnement matériel et logiciel [Kajimoto et al., 2014]. . . . .	37
2.8	Les différents éléments développés dans le projet Sound of Vision [Sound of Vision Consortium, 2018]. . . . .	39
2.9	Emplacements potentiels des implants rétiniens. A : implant épi-rétinal; B : implant sous-rétinien; C : implant supra-choroïdal; D : implant épi-scléral. Source [?]. . . . .	42
2.10	Le <i>Personal Guidance System</i> de Loomis <i>et al.</i> (photo prise sur le site du système : <a href="http://www.geog.ucsb.edu/pgs/main.htm">http://www.geog.ucsb.edu/pgs/main.htm</a> (accès le 13 septembre 2019)). . . . .	44
2.11	<i>Seeing Eye GPS</i> est une application mobile de navigation « virage par virage ». Image tirée de <a href="https://www.senderogroup.com/products/shopseeingeyegps.html">https://www.senderogroup.com/products/shopseeingeyegps.html</a> (accès le 13 septembre 2019). .	45
2.12	Le dispositif de Filipe <i>et al.</i> analyse le profil de profondeur sur des lignes pour identifier les obstacles et escaliers. . . . .	47
2.13	Étapes d'analyse des images pour les obstacles de Tapu <i>et al.</i> [Tapu et al., 2013]. . . . .	48

2.14	L'application de reconnaissance Microsoft Seeing permet notamment de décrire des scènes (images prises sur le site de l'application <a href="https://www.microsoft.com/en-us/ai/seeing-ai">https://www.microsoft.com/en-us/ai/seeing-ai</a> ). . . . .	50
2.15	La canne <i>iCane</i> de Chang <i>et al.</i> (image tirée de [Chang et al., 2005])	53
2.16	Le robot de Kulyukin <i>et al.</i> (image tirée de [Kulyukin et al., 2004, 2006]). . . . .	54
2.17	Le dispositif GoSense Rango (image tirée de [GoSense SAS]). . . . .	57
2.18	Modèle de la chaîne perceptive . . . . .	69
2.19	Modèle du système $\{dispositif\ d'assistance + personne\}$ . . . . .	72
2.20	Chemins possibles de l'information dans le système $\{dispositif\ d'assistance + personne\}$ . . . . .	72
3.1	L'architecture globale du projet SEES. . . . .	81
3.2	Le prototype de SEE-stick. À gauche, une vue d'ensemble du prototype, à droite, une vue plus détaillée sur la carte à laquelle sont connectés la plupart des capteurs. . . . .	84
3.3	Composants électroniques constitutifs du prototype de SEE-stick et leurs interconnexions. . . . .	84
3.4	Répartition des composants électroniques sur la plate-forme mécanique du prototype de SEE-stick. . . . .	85
3.5	Trajet en extérieur réalisé avec le SEES. L'image est tirée de l'interface web du système [Yusro, 2017]. . . . .	86
3.6	Fonctionnement de la navigation à l'estime : une intégration de déplacements infinitésimaux permet de calculer un déplacement. . . . .	87
3.7	La reconstruction de la trajectoire avec les données de distance parcourue et de d'orientation (notée $\alpha$ sur la figure). . . . .	87
3.8	Résultat du calcul d'un parcours en intérieur visant à mettre en évidence la dérive en effectuant plusieurs tours à l'intérieur d'une pièce : l'estimation est assez bonne. . . . .	89
3.9	Fonctionnement global de l'application de détection visuelle de l'état des feux et de poteaux ( [Yusro, 2017]). . . . .	90
3.10	Détection de l'état d'un feu tricolore ([Yusro, 2017]). . . . .	91
3.11	Fonctionnement de l'algorithme de détection de poteaux (figure de [Yusro, 2017]). . . . .	91
3.12	Vue 3D éclatée de la tête de canne du LMSC. . . . .	99
3.13	Détail du positionnement de l'encodeur magnétique et de l'axe de roue. . . . .	100
3.14	Fonctionnement des trois groupes de capteurs d'obstacles du LMSC en fonction de l'inclinaison de la canne : (a) normale, (b) forte. . . . .	101
3.15	Anneau de stabilisation des capteurs d'obstacles. . . . .	101
3.16	Deuxième version du LMSC, avec support de capteurs fixe et capteurs et roues additionnelles. . . . .	101
3.17	Architecture de nœud de réseau de capteurs à simple cœur. . . . .	103
3.18	Architecture de nœud de réseau de capteurs multi-cœur. . . . .	104
3.19	Architecture multi-cœurs prévue initialement pour le LMSC, avec l'affectation des tâches pour chaque microcontrôleur (MCU sur la figure) et pour les Raspberry Pi. . . . .	106

3.20	Architecture redondante symétrique. . . . .	107
3.21	La carte uSu-Edu. . . . .	108
3.22	Architecture initialement adoptée pour le LMSC. . . . .	109
3.23	Architecture du LMSC après ajout de capteurs. . . . .	109
3.24	Connexions entre l'application, le système, les pilotes, et le matériel. . . . .	111
3.25	Liens entre l'application, le système et le simulateur ; ce dernier simule le matériel et les pilotes. . . . .	112
3.26	Le système d'exploitation de la carte uSu-Edu représente les tâches par des automates. . . . .	114
3.27	L'automate de gauche est la simplification de l'automate de droite, dans lequel la tâche X peut être répétée au plus $k$ fois alors que dans la version simplifiée, cette limite disparaît. . . . .	114
3.28	Trajectoire estimée par le filtre particulaire sur des données pré-enregistrées avec le SEE-stick. À gauche, la trajectoire sans ré-échantillonnage des particules, permettant de mieux voir l'effet des détections de capteurs ultrasoniques ; à droite, la trajectoire ré-échantillonnée. . . . .	121
3.29	Trajectoire estimée par le filtre particulaire sur des données de navigation à l'estime pré-enregistrées avec le SEE-stick et des données de capteurs de distance fixées artificiellement à des valeurs correspondant à la réalité pour mieux voir le fonctionnement du filtre. À gauche, la trajectoire sans ré-échantillonnage des particules, permettant de voir l'effet des détections simulées de capteurs ultrasoniques ; à droite, la trajectoire ré-échantillonnée. . . . .	122
3.30	Fonctionnement visuel de la méthode de correction des trajectoires par la détection des collisions avec les murs (cartographiés). . . . .	123
3.31	Estimation de trajectoire par la méthode particulaire simplifiée – trajectoires. À gauche, toutes les particules sont représentées ; à droite, l'estimation résultante de la moyenne des particules. . . . .	124
3.32	Une cellule de détection composée de deux capteurs infrarouges et deux capteurs ultrasoniques, pour assurer à la fois redondance et adaptabilité à l'environnement. . . . .	126
3.33	Configuration pour le test des affinités capteur – matériau. . . . .	128
3.34	Configuration du test visant à déterminer les effets des interférences sur les capteurs de distance. . . . .	129
3.35	L'application smartphone du 2SEES. . . . .	133
4.1	Aperçu de la gamme d'interactions possibles par la fonction d'assistance à la perception. . . . .	137
4.2	Pile « classique » des standards et langages du web sémantique du W3C. Source : Marobil (CC0) ( <a href="https://en.wikipedia.org/wiki/File:Semantic_web_stack.svg">https://en.wikipedia.org/wiki/File:Semantic_web_stack.svg</a> ) . . . . .	141
4.3	Standards du web sémantique utilisés par la fonctionnalité SO2SEES. . . . .	142
4.4	Intégration de WoT-TD avec une ontologie plus large (source : [Kae-bisch et al., 2019a] ; Copyright © [2019] World Wide Web Consortium, (MIT, ERCIM, Keio, Beihang). <a href="http://www.w3.org/Consortium/Legal/2015/doc-license">http://www.w3.org/Consortium/Legal/2015/doc-license</a> ). . . . .	143
4.5	Acteurs et plates-formes impliquées dans la fonctionnalité SO2SEES. . . . .	145

4.6	Modèle des interactions entre les composantes intervenant dans la fonctionnalité SO2SEES. . . . .	146
4.7	Cheminement simplifié d'une requête et de la réponse délivrée. . . . .	153
4.8	Cheminement d'une requête prenant en compte le contexte. . . . .	155
4.9	Maquettes d'objets : à gauche, la fenêtre intelligente ; à droite, la porte intelligente. . . . .	159
4.10	Exemple de serveur local, boîtier ouvert : (1) carte uSu-Edu, branchée sur la carte Raspberry Pi 3 (dont on voit les ports USB), (2) antenne IEEE 802.15.4, (3) alimentation secteur de la carte uSu-Edu, (4) passage des câbles sortant du boîtier, et (5) carte 4G USB. . . . .	160
4.11	Capture d'écran de l'application mobile 2SEES et l'interface SO2SEES proposant deux requêtes. . . . .	162
4.12	Plan de la pièce dans le premier scénario d'expérimentation montrant les objets inclus. . . . .	163
4.13	Canaux de communication dans le premier scénario. . . . .	164
4.14	Relations d'influence entre les objets et la température. . . . .	164
4.15	Plan de la pièce dans le second scénario d'expérimentation montrant les objets inclus. . . . .	168
4.16	Graphe des requêtes, indices de contexte et descriptions d'objets. . . . .	170
4.17	Capture d'écran de l'application web de gestion des requêtes et des indices de contexte. La requête « Pourquoi fait-il froid ? » est ici liée à quatre indices contextuels (deux, dupliqués en français et anglais) : l'état d'ouverture de la fenêtre, et le capteur de température LM75B. . . . .	171
4.18	Ensemble des bases de connaissances potentiellement mobilisées pour la réponse à une requête. . . . .	171

# Liste des tableaux

2.1	Combinaisons possibles des caractéristiques d'une infrastructure. . .	66
2.2	Tableau décrivant les différents cheminements possibles suivant la numérotation de la figure 2.20. . . . .	73
2.3	Tableau récapitulatif des dispositifs analysés, classés suivant leurs fonctionnalités et leurs caractéristiques ( <i>SV</i> désigne la synthèse vocale, <i>SS</i> la substitution ou supplémentation sensorielle, <i>3D</i> les caméras de profondeur (double capteur, temps-de-vol ou à lumière structurée), <i>IMU</i> désigne une centrale inertielle ( <i>Inertial Measurement Unit</i> en anglais) et <i>rec.</i> est l'abréviation de <i>reconnaissance</i> ). . . . .	76
3.1	Matrice des sources potentielles de défaillance. . . . .	97
3.2	Comparaison des qualités respectives des dispositifs d'assistance à la mobilité des personnes malvoyantes. . . . .	98
3.3	Affinités capteurs – matériaux à 50 cm, 100 cm, 150 cm. Les valeurs dans le tableau sont les niveau moyens de sortie des capteurs dans l'unité native de la sortie, soit des millivolts pour les capteurs infrarouges et ultrasoniques, et directement des centimètres pour le Lidar Lite v3, avec entre parenthèses les écarts-types. Les valeurs en gras sont remarquables car la sortie du capteur y est clairement faussée.	129
3.4	Effet des interférences sur les capteurs ultrasoniques. Les valeurs sont en millivolts, avec l'écart-type entre parenthèses. . . . .	129
3.5	Évaluation du taux de détection selon le capteur avec la formule 3.9. Les taux de positifs corrects (PC) et de négatifs corrects (NC) sont comparés. . . . .	130



# Introduction

La question au cœur des travaux présentés dans cette thèse est l'amélioration de la mobilité des personnes malvoyantes. La mobilité et la lecture<sup>1</sup> sont deux des plus importants défis pour les personnes déficientes visuelles [Giudice and Legge, 2008]. Comment marcher quand on ne voit pas ? Comment traverser une route ? Comment acheter ce dont on a besoin pour vivre ? Comment retrouver un objet ? La mobilité est peut-être le domaine des activités humaines dans lequel les difficultés liées à la malvoyance sont les plus intuitives. Peut-être parce que le contexte dans lequel les personnes valides rencontrent – et remarquent – le plus les personnes malvoyantes est la rue, au cours de leurs déplacements. Peut-être aussi parce que chacun a déjà fait au moins une fois l'expérience de se déplacer dans le noir absolu, la nuit, ou en fermant les yeux. C'est très difficile : après quelques pas seulement, on ne sait plus où l'on se trouve dans un espace que l'on vient de « quitter » visuellement. Les singularités de l'environnement, pourtant connues car vues il y a un instant, deviennent soudain menaçantes : murs, chaises, escaliers, poteaux, portes deviennent autant de risques de chocs et de chutes. Être plongé dans un espace que l'on perçoit vide alors que nous le savons peuplé d'obstacles est gênant, et la crainte met souvent rapidement fin à « l'expérimentation » de la cécité. Accéder à un niveau de compréhension plus poussé est cependant difficile : d'une part, parce que l'expérience de la cécité est globalement inaccessible aux voyants, d'autre part, parce que les notions de mobilité et de malvoyance sont intrinsèquement complexes. Le chapitre 1 éclaire les concepts centraux utilisés tout au long de cette thèse et la manière dont nous nous positionnons par rapport à eux : déficience visuelle, mobilité, et perception.

De nombreux dispositifs électroniques d'aide à la mobilité ont été développés depuis les années 1960. Ils sont très variés, allant du capteur de distance monté sur une canne blanche classique à des robots mobiles d'assistance en passant par des applications smartphones d'analyse de scène. Malheureusement, ces aides sont généralement très peu utilisées par les personnes déficientes visuelles [Loomis et al., 2012]. Cette non-adoption constitue la problématique centrale des dispositifs électronique d'assistance à la mobilité. Elle est liée aux complexités de la mobilité et des handicaps visuels qui sous-tendent les nombreux modes d'assistance possibles. La conception d'un dispositif électronique d'assistance repose sur deux piliers : d'une part, l'acquisition et le traitement des informations qui seront fournies à la personne, et, d'autre part, l'interface personne-machine par laquelle les informations

---

1. Il est à noter que la lecture est une tâche courante dans les activités de mobilité, ne serait-ce que pour prendre connaissance du contenu des panneaux indicateurs.

sont fournies. Pour qu'un système soit fonctionnel et utilisable, il est indispensable que ces deux aspects soient performants. Le chapitre 2 présente un panorama des dispositifs d'assistance à la mobilité pour les personnes déficientes visuelles ; la discussion qui accompagne cet état-de-l'art vise à expliciter les possibilités dans lesquelles viennent s'insérer les développements effectués pendant cette thèse.

Les travaux entrepris dans cette thèse se sont concentrés sur le premier aspect de la conception des dispositifs d'assistance, c'est-à-dire la collecte et le traitement des informations, plus que sur le second, c'est-à-dire l'interface. Le chapitre 3 décrit le système conçu, nommé 2SEES, qui intègre de nombreux capteurs et réunit trois familles d'assistance dans un même système : détection d'obstacles, navigation, et perception de l'environnement. Dans le cadre de la navigation, nous avons cherché à développer un algorithme de localisation léger, économe en énergie, et indépendant de toute infrastructure externe. Nous avons étudié la robustesse d'une fonctionnalité de détection d'obstacles en fonction des capteurs utilisés et du contexte. Ces deux fonctionnalités s'inscrivent dans les problématiques de réduction de la consommation d'énergie et d'adaptation au contexte qui sont communes à de nombreux systèmes embarqués. Elles sont également liées à la problématique de la tolérance aux fautes et de la robustesse du système.

L'assistance à la perception de l'environnement a fait l'objet d'une attention particulière : une fonctionnalité novatrice permettant aux personnes déficientes visuelles d'interagir avec les objets intelligents environnants par le biais de requêtes de haut niveau a été conceptualisée et implémentée (chapitre 4). Cette fonctionnalité, que nous avons nommée SO2SEES, repose sur les standards du web sémantique du W3C pour faciliter son interopérabilité et son intégration avec les écosystèmes d'objets intelligents existants et futurs. Elle rejoint les deux autres fonctionnalités car elle est soumise aux mêmes contraintes : consommation énergétique, adaptation au contexte, tolérance aux fautes et robustesse.

# Chapitre 1

## Mobilité, malvoyance, et perception

### 1.1 Mobilité

Nous devons décrire avec un peu plus de détails ce que nous entendrons ici par le mot *mobilité*. La mobilité, dans sa définition la plus large, est la « capacité à se mouvoir ou à être mû. » [Académie Française, 2019]. Nous devons resserrer un peu cette définition. Notre mobilité concerne des personnes; leur capacité à être mues nous intéresse peu, contrairement à celle à se mouvoir. Celle-ci ne représente pas seulement la faculté physique nécessaire, mais surtout l’aptitude générale à aller dans des lieux choisis. En résumé, la mobilité représente la capacité des personnes à se déplacer là où elles le souhaitent. C’est ainsi que la décrit Giudice : « *The ability to navigate from place to place [...]* » [Giudice and Legge, 2008].

Cette définition de la mobilité (ainsi que toutes ses variantes dans la formulation) est juste mais insuffisante. En effet, elle contient tout mais n’explique rien : elle ne décrit la mobilité qu’à son plus haut niveau, c’est-à-dire celui de son objectif ou de la fonction qu’elle remplit. Ce niveau peut paraître naturel. Cependant, pour améliorer la mobilité, ce qui est notre problématique, il faut nous intéresser à la manière dont cet objectif de déplacement d’un endroit à un autre est réalisé. La complexité apparaît ici : même pour des personnes dotées d’une bonne vision, certains déplacements peuvent se révéler complexes, nécessiter des réflexions préalables, des qualités motrices particulières, et comporter des risques. Aller chercher un objet dans une pièce voisine est presque toujours simple; alors qu’aller dans un bureau de poste peut être légèrement complexe – en raison de la variabilité des horaires; traverser un canyon sans pont ou se rendre à un endroit éloigné du globe sera, presque nécessairement, difficile. La difficulté est intrinsèque à certains déplacements. Pour les aides à la mobilité, la diversité des situations possibles dont on doit tenir compte est une autre difficulté. La différence fondamentale entre la mobilité des personnes déficientes visuelles et celle des personnes voyantes tient au fait que bon nombre des déplacements qui semblent triviaux pour les secondes sont loin de l’être pour les premières. Cela conduit à une augmentation générale de la

difficulté des trajets.

### 1.1.1 Modèle de base

Pour mieux comprendre les problèmes liés à la mobilité, nous devons l'analyser et examiner les fonctions qu'elle implique. Le nombre de fonctions nécessaires varie selon les déplacements, mais on peut commencer par trouver un ensemble minimal. Pissaloux a identifié quatre fonctions principales liées à des niveaux de consciences différents et faisant intervenir plusieurs types de représentations spatiales [Pissaloux et al., 2017] :

1. La marche qui consiste à se déplacer en évitant les obstacles, et qui comprend les sous-fonctions suivantes :
  - (a) La détection d'obstacles ;
  - (b) La localisation (direction, distance) des obstacles par rapport à son propre corps ;
  - (c) L'estimation de la forme des obstacles.
2. L'orientation, c'est-à-dire la conscience ou la connaissance de sa position dans un trajet, par rapport à son point de départ et à sa destination. L'orientation est une fonction qui présuppose (dans ce cadre) l'existence d'un trajet et dont les sous-fonctions sont :
  - (a) L'identification de sa propre position ;
  - (b) L'identification de la position de la destination ;
  - (c) La détermination de la direction à prendre pour atteindre sa destination.
3. La conscience ou connaissance spatiale (*spatial awareness* et *spatial knowledge* en anglais dans le texte), la connaissance mémorisée des lieux, ce qui inclut :
  - (a) La connaissance spatiale des lieux ;
  - (b) La connaissance « sémantique » des lieux (fonction sociale et urbaine) ;
  - (c) La connaissance des stimuli utiles potentiels, qui permettent par exemple d'identifier les points de repères.
4. Enfin, la navigation est le plus haut niveau, et intègre l'utilisation des fonctions précédentes.

La mise en œuvre de la navigation est faite par une stratégie de navigation [Pissaloux et al., 2017], c'est-à-dire la méthode que va suivre une personne pour arriver à sa destination. La stratégie de navigation la plus citée dans le cas des personnes déficientes visuelles est l'intégration de chemins, dans laquelle un trajet complexe est divisé en une séquence de trajets simples. Un trajet simple est un trajet sur lequel la principale activité est de mesurer la distance ; les transitions entre les trajets simples sont alors les changements de direction et les vérifications de position par l'intermédiaire de points de repères (« je vais jusqu'à la poubelle qui se trouve à 50 pas, puis je tourne à droite à l'angle suivant, etc. »).

### 1.1.2 Modèle étendu

Ce modèle est pertinent d'une part parce que ses fonctions sont nécessaires pour toute mobilité humaine, et d'autre part parce que le découpage en fonction qu'il met en oeuvre est adapté à notre cas. On peut l'étendre vers le bas, c'est-à-dire expliciter ses fonctions afin de rendre plus naturelle la description de certains trajets. Premièrement, son niveau le plus bas est la marche, car on y considère que la personne dispose d'une motricité adéquate [Pissaloux et al., 2017]. Mais rien ne nous empêche de décomposer les fonctions pour faire apparaître les problématiques de motricité. Ce niveau de décomposition est-il pertinent ? Plusieurs éléments peuvent inciter à en tenir compte : le poly-handicap moteur/déficience visuelle, les surfaces intrinsèquement difficiles à franchir sans vision, et la plus grande universalité du modèle, qui pourrait alors servir dans le cadre de travaux sur la mobilité des personnes avec mono-handicaps moteurs – on peut réciproquement s'inspirer des travaux dans ce domaine pour construire notre modèle.

On peut également étendre ce modèle par le haut. Une structuration intuitive apparaît quand on examine ces exemples, on voit qu'il y a une hiérarchie dans la mobilité. Un déplacement, selon l'échelle, peut être un déplacement local, par exemple, en marchant, ou à plus large échelle, comme un trajet en train ; l'implémentation du déplacement le plus large sera réalisée en le découpant en sous-problèmes de mobilité à plus petite échelle. On a ainsi une division du trajet en plusieurs segments. Par exemple, pour aller à l'université, quelqu'un partira de son domicile jusqu'à l'arrêt de tramway le plus proche, puis prendra un tramway jusqu'à l'arrêt le plus proche de l'université, et enfin, marchera jusqu'à l'université. Les segments sont simplifiés dans leur description, et ne pas les utiliser revient à expliciter le trajet dans tous ses détails : la personne va de la pièce où elle se trouve dans son domicile jusqu'à la porte de sortie, puis, de là, se dirige vers l'arrêt de tramway, monte dans le premier tramway qui passe et va dans la bonne direction, se dirige vers une place libre dans celui-ci, attend son arrêt en prêtant attention aux signes indiquant la proximité de l'arrêt désiré, puis sort du tramway et marche jusqu'à la porte du bâtiment universitaire qui l'intéresse particulièrement, etc. Une segmentation dont le nombre d'itérations ou de niveaux hiérarchiques est en adéquation avec la complexité du trajet a l'avantage de regrouper les phases corrélées en étapes, ce qui rend l'intégralité du trajet intelligible en elle-même.

Enfin, le modèle de base est extensible horizontalement, dans la nature des fonctions impliquées. En effet, la diversité des sous-tâches qui sont potentiellement à effectuer est lourde. Quelques exemples :

1. Pour se déplacer dans un bâtiment destiné à recevoir du public, il faut généralement être capable de lire à distance les panneaux indicateurs pour savoir, par exemple, la fonction d'un lieu, ou la direction à suivre pour atteindre un autre point. Les cas les plus communs sont certainement les indications désignant les toilettes et celles désignant les sorties ;
2. Prendre un transport en commun nécessite généralement un ensemble d'actions non-triviales, générant une charge cognitive importante même chez beaucoup de personnes ayant une bonne vision. En effet, cette action suppose de :

- (a) Disposer d'un titre de transport, obtenu soit dans une machine, soit auprès d'un vendeur physique, soit dans le véhicule lui-même – chacune de ces possibilités apportant elle-même son cortège de sous-fonctions ;
  - (b) Connaître l'emplacement des arrêts ;
  - (c) Savoir quelle(s) ligne(s) desser(ven)t les arrêts désirés ;
  - (d) Parvenir à composer ou valider son titre de transport.
3. Prendre un avion est également une opération complexe, qui nécessite de :
- (a) Aller à l'aéroport, souvent accessible par des transports spécifiques ;
  - (b) Disposer d'un billet imprimé ou bien le faire imprimer à un guichet de l'aéroport ; ce qui nécessite de trouver le guichet de sa compagnie aérienne ;
  - (c) Faire enregistrer ses bagages si on en a ;
  - (d) Passer les zones de contrôle de sécurité ;
  - (e) Avoir connaissance des horaires, retards éventuels, porte d'embarquement de son avion ;
  - (f) Parvenir ensuite à trouver la bonne porte d'embarquement.

Toutes ces opérations s'effectuent dans des zones potentiellement éloignées, en particulier dans les grands aéroports, dont les portes d'embarquements sont typiquement réparties entre plusieurs terminaux éloignés, entre lesquels la circulation s'effectue par un service de navettes ou de train léger automatique. Pour limiter la complexité de leur usage, les aéroports font d'importants efforts de signalisation, d'organisation et de canalisation des flux. Ces efforts ne suffisent pas à rendre leur utilisation aisée pour les usagers voyants. Pour les personnes malvoyantes, prendre un avion dans un aéroport international relèverait probablement de l'exploit si des membres du personnel n'étaient pas disponibles pour les accompagner.

## 1.2 Malvoyance et cécité

La connaissance « naturelle » des problèmes de mobilité exposés plus haut tend à se heurter, dans une représentation intuitive de la situation, à un autre aspect du handicap, menant à une représentation incomplète sinon faussée. En effet, des capacités spécifiques peuvent être acquises par certaines personnes malvoyantes ; ces capacités, impressionnantes pour des voyants, viennent en partie compenser la perte de la vision. La plus impressionnante est sans doute l'écholocalisation : certaines personnes malvoyantes sont capables de « sentir » la distance des objets [Kellogg, 1962]. Si on y réfléchit, cette technique n'est pas tant impressionnante en raison sa performance dans l'absolu – la vision est un outil bien plus efficace – qu'en raison de l'impression (d'ailleurs fausse) d'inaccessibilité perçue par les personnes voyantes.

La première partie de la Lettre sur les Aveugles de Diderot (XVII<sup>e</sup> siècle) est consacrée à la description des prodigieuses capacités d'un aveugle de naissance [Diderot,

2011]. Son ouïe est fine : « Il a la mémoire des sons à un degré surprenant, et les visages ne nous offrent pas une diversité plus grande que celle qu’il observe dans les voix », et il sait s’en servir : « Notre aveugle adresse au bruit ou à la voix si sûrement que je ne doute pas qu’un tel exercice ne rendit pas les aveugles très adroits et très dangereux. [...] il saisit le premier objet qui lui tomba sous la main, le lui lança, l’atteignit au milieu du front, et l’étendit par terre. ». Il a développé des techniques pour effectuer des tâches requérant normalement la vue : pour enfiler un fil dans une aiguille, il « [...] dispose l’ouverture de l’aiguille transversalement entre ses lèvres et dans la même direction que celle de sa bouche ; puis, à l’aide de sa langue et de la succion, il attire le fil qui suit sont haleine [...] » : il « [...] estime la proximité du feu aux degrés de la chaleur ; la plénitude des vaisseaux, au bruit que font en tombant les liqueurs qu’il transvase [...] ». Fort de ses capacités, il refuserait la vue si ce n’était pour la curiosité que l’expérience lui inspire : « [...] Si la curiosité ne me dominait pas, dit-il, j’aimerais bien autant avoir de longs bras [...] ». De telles attributions laissent à penser que les personnes déficientes visuelles sont capables de partiellement compenser leur handicap. En réalité, même lorsqu’elle existe, la « compensation » ne permet pas d’arriver à un niveau fonctionnellement équivalent à celui d’une personne voyante. Premièrement, les possibilités de l’écholocation et des techniques basées sur le son et le toucher sont limitées : par exemple, on ne peut pas les utiliser pour la détection d’amers de repérage lointains ou pour lire les plaques des rues. Deuxièmement, la diversité des personnes et des formes de malvoyance fait que cette adaptation n’est pas universelle : tout le monde n’a pas les mêmes capacités. Les stéréotypes dits « positifs » que la population globale peut avoir à l’égard des personnes déficientes visuelles ont été étudiés [Bolt, 2006]. Le génie musical, le succès sans effort et la possession de sens non-visuels extraordinairement sensibles, par exemple, sont trois stéréotypes positifs infondés [Hersh, 2018]. Ils sont potentiellement aussi gênants que les stéréotypes négatifs – infondés eux aussi [Hersh, 2018]. Il est important de prêter une attention particulière à ne pas se laisser influencer par de tels stéréotypes, ce qui n’est pas évident : si certaines affirmations le sont ostensiblement, d’autres, concernant par exemple le niveau d’autonomie, le sont moins clairement. Il est tentant de rassembler toutes les personnes dans une unique catégorie, qu’elle soit positive (« les aveugles savent utiliser une canne et peuvent se déplacer seuls ») ou négative (« les aveugles restent chez eux »), alors qu’en réalité, il existe une très grande variabilité entre les individus. Pour la conception d’un système d’assistance, se pose alors un problème gênant : ou bien la complexité du système augmente car on tente de prendre en compte la diversité des utilisateurs, ou bien on se borne à travailler sur un système qui ne sera utile qu’à une population restreinte. Au vu des différences d’âge, d’éducation, de niveaux de revenus, de lieux de vie, de culture et de goûts personnels, il paraît difficile de s’extraire de cette dualité et de proposer un système pertinent pour le plus grand nombre.

En premier lieu, pour avoir une idée plus précise de la population souffrant de malvoyance, nous devons examiner :

1. La définition et les symptômes cliniques de la malvoyance ;
2. Les différentes pathologies à son origine et leurs prévalences respectives.

### 1.2.1 Définition et symptômes

L'Organisation Mondiale de la Santé (OMS) définit six niveaux de déficience visuelle affectant la vision de loin et un niveau affectant la vision de près [World Health Organisation, 2018]. La déficience visuelle affectant la vision de loin nous intéresse davantage, étant donné son impact beaucoup plus important sur la mobilité. La classification en niveaux repose sur la mesure de l'acuité visuelle avec correction.

1. Malvoyance légère : acuité inférieure strictement inférieure à 6/12 et supérieure ou égale à 6/18 ;
2. Malvoyance modérée : acuité inférieure strictement inférieure à 6/18 et supérieure ou égale à 6/60 ;
3. Malvoyance sévère : acuité inférieure strictement inférieure à 6/60 et supérieure ou égale à 3/60 ;
4. Cécité : acuité inférieure strictement inférieure à 3/60 et supérieure ou égale à 1/60 ;
5. Cécité : acuité inférieure strictement inférieure à 1/60 avec perception de la lumière ;
6. Cécité : pas de perception de la lumière.

Le handicap des personnes dont la vision ne couvre qu'un champ central de  $10^\circ$  (perte de vision périphérique) appartient à la troisième catégorie [Abiose et al., 2003].

### 1.2.2 Pathologies et prévalences respectives

De nombreuses pathologies peuvent causer la déficience visuelle. Une étude en a analysé les causes principales dans le monde [Flaxman et al., 2017]. Sept types de pathologies ont été retenus :

- Cataracte : opacification du cristallin, souvent due à l'âge mais parfois congénitale ou conséquence d'une affection oculaire [World Health Organisation, 2019] ;
- Erreurs de réfraction non corrigées : typiquement myopie, hypermétropie, astigmatisme et presbytie non corrigées ;
- Glaucome : perte de sensibilité liée à la dégradation des cellules ganglionnaires rétiniennes et du nerf optique ;
- Dégénérescence maculaire liée à l'âge (DMLA) : lésions de la macula (zone centrale de la rétine, correspondant à la vision centrale) apparaissant avec l'âge ;
- Rétinopathie diabétique : lésions rétinales causées par des troubles vasculaires ;
- Opacification de la cornée (trachome mis à part) : les traumatismes oculaires et les ulcérations de la cornée, l'onchocercose ou cécité des rivières (maladie due à un parasite), ou encore la lèpre peuvent conduire à une opacification irréversible de la cornée.

- Trachome : maladie infectieuse provoquant une sclérosation de la paupière et une opacification irréversible de la cornée.

Les autres causes, connues ou non, ont été comptabilisées dans une huitième catégorie.

En 2015, la cécité (niveaux 4, 5 et 6) touchait 36 millions de personnes dans le monde (intervalle de confiance à 80 % : 12,9 millions à 65,4 millions) et les prévalences pour les personnes âgées de plus de 50 ans étaient en 2015 dans l'ordre décroissant [Flaxman et al., 2017] :

1. La cataracte avec 35,15 % (IC-80 %<sup>1</sup> : 26,40 % – 44,03 %) ;
2. Les erreurs réfractives non corrigées avec 20,28 % (IC-80 % : 18,23 % – 24,24 %) ;
3. Le glaucome avec 8,49 % des cas (IC-80 % : 2,99 % – 15,66 %) ;
4. La DMLA avec 5,93 % (IC-80 % : 1,46 % – 12,18 %) ;
5. L'opacification de la cornée (trachome mis à part) représentait 3,21 % des cas (IC-80 % : 0,50 % – 7,19 %) ;
6. La rétinopathie diabétique avec 1,06 % des cas (IC-80 % : 0,15 % – 2,38 %) ;
7. Enfin, le trachome représentait 0,97 % des cas (IC-80 % : 0,80 % – 1,15 %) ;

Les autres causes, non discriminées, représentaient 24,92 % des cas (intervalle de confiance à 80 % : 9,58 % – 43,36 %).

La malvoyance modérée à sévère (niveaux 2 et 3) touchait, en 2015, 216,6 millions de personnes (IC-80 % : 98,5 millions à 359,1 millions) [Flaxman et al., 2017] et les prévalences pour les personnes âgées de plus de 50 ans étaient dans l'ordre décroissant :

1. Les erreurs réfractives non corrigées avec 52,34 % (IC-80 % : 18,23 % – 24,24 %) ;
2. La cataracte avec 25,15 % (IC-80 % : 18,83 % – 31,76 %) ;
3. La DMLA avec 4,38 % (IC-80 % : 1,05 % – 9,15 %) ;
4. Le glaucome avec 2,05 % des cas (IC-80 % : 0,62 % – 4,03 %) ;
5. La rétinopathie diabétique avec 1,30 % des cas (IC-80 % : 0,20 % – 2,93 %) ;
6. L'opacification de la cornée (trachome mis à part), représentait 1,14 % des cas (IC-80 % : 0,17 % – 2,48 %) ;
7. Le trachome représentait 0,64 % des cas (IC-80 % : 0,50 % – 0,79 %) ;

Les autres causes, non discriminées, représentaient 13,00 % des cas (IC-80 % : 4,14 % – 24,57 %).

Parmi ces pathologies, certaines sont guérissables : les erreurs de réfraction peuvent être corrigées avec des lunettes, la cataracte peut être opérée en remplaçant le cristallin par un implant artificiel. Pour d'autres, l'apparition peut être diagnostiquée et évitée, ou l'évolution stoppée. C'est le cas du glaucome, qui peut parfois être prévenu, et de la rétinopathie diabétique. L'opacification de la cornée peut être

---

1. Intervalle de confiance à 80 %.

prévenue (dans le cas du trachome, par exemple), mais une fois survenue, seule une greffe permet de recouvrer la vision. La DMLA a deux formes : la DMLA exsudative, qui peut être endiguée et DMLA atrophique, pour laquelle il n'existe pas de traitement.

L'amélioration des systèmes de santé et de l'accès aux soins, au niveau mondial, fait que les situations de malvoyance ou de cécité qui peuvent être évitées le sont de plus en plus. En revanche, l'accroissement et le vieillissement de la population provoquent, eux, une augmentation du nombre de personnes malvoyantes ou aveugles. L'effet de cette croissance est, pour le moment, plus important que la diminution [Flaxman et al., 2017], et il est donc probable que le nombre de personnes malvoyantes ou aveugles continue à augmenter dans les prochaines années.

Qu'en est-il des populations infantiles ? Un article de Foster et Gilbert s'est attaché à analyser la situation des enfants aveugles (catégories 4, 5, 6 de l'OMC données plus haut ou champ visuel inférieur à 10 °) âgés de 0 à 15 ans [Foster and Gilbert, 1992]. Cette étude a montré une corrélation positive entre taux de mortalité infantile et de la cécité infantile : en 1988, la prévalence de cécité infantile dans les pays souffrant d'un taux de mortalité infantile très élevé était 5 fois supérieure à celle des pays où ce taux était bas. Cette corrélation semble logique puisque mortalité et cécité infantile partagent en grande partie les mêmes causes : la faiblesse des systèmes de santé et les problèmes de nutrition [Foster and Gilbert, 1992]. Il était ainsi estimé qu'en 1990, 90 % de la population d'enfants aveugles vivait dans des pays en voie de développement [Foster and Gilbert, 1992].

La cécité est difficile à évaluer et à diagnostiquer chez les enfants avant l'âge de 2 ans [Foster and Gilbert, 1992] et les formes plus légères de malvoyance le sont probablement encore plus.

En conclusion, la population de personnes malvoyantes est pour sa plus large partie constituée de personnes ayant acquis le handicap au cours de leur vie, voire à un âge avancé. On constate que la cécité est bien moins fréquente que la malvoyance sévère ou modérée. Il y a un certain décalage entre la représentation que la plupart des personnes dans les pays occidentaux se font de ce handicap et la situation réelle. Le niveau de handicap qui domine souvent les représentations, le plus discuté, est la cécité totale, en particulier la cécité depuis la naissance de personnes plutôt jeunes vivant en ville dans un pays occidental. Les trois catégories que forment la cécité totale, la jeunesse, et l'occident sont, si l'on se borne à dénombrer les populations, sont individuellement minoritaires. Leur combinaison ou intersection l'est donc encore plus.

### 1.2.3 Ciblage

Cette disparité soulève une question importante pour les concepteurs de systèmes d'assistance : qui doit-on cibler ? Si on suit les statistiques de population de la manière la plus simple possible – attitude soulevant des questions utilitaristes que nous laisserons de côté ici – on se focalisera sur les personnes âgées, vivant dans les pays en voie de développement et souffrant de malvoyance modérée ou sévère

acquise au cours de leur vie. Dans la conception de notre dispositif d'assistance, nous avons cependant ciblé une population plutôt occidentale, d'âge « indéterminé » et souffrant de cécité totale (incapacité à percevoir la lumière), pour les raisons détaillées ci-dessous.

Tout d'abord, l'âge ne représente pas un facteur en lui-même déterminant mais certaines de ses conséquences possibles le sont. Deux nous intéressent principalement ici : la facilité d'apprentissage et les handicaps supplémentaires. Pour la première, il est admis que plus l'âge avance, plus il est difficile pour une personne d'apprendre l'utilisation ou adopter d'un nouveau système. Cela peut être dû à des raisons biologiques : citant Stokes [Stokes, 1992], Soong *et al.* mentionnent une faible performance du transfert depuis la mémoire à court terme vers la mémoire à long terme et de la mémoire en général comparée à des populations plus jeunes, après avoir constaté la non-amélioration de la mobilité de personnes malvoyantes âgées immédiatement après un apprentissage de techniques d'orientation et mobilité [Soong et al., 2001]. En outre, la probabilité d'être touché par d'autres handicaps augmente sensiblement avec le vieillissement, rendant les tâches de mobilité encore plus complexes. Au final, il semble que la différence principale entre un dispositif ciblant particulièrement les personnes âgées réside dans le *niveau de qualité* plus élevé nécessaire à son adoption plus que dans son fonctionnement intrinsèque. Les principales lignes directrices de développement ne seront donc probablement pas affectées par un ciblage de catégorie. Ce dernier point étant celui sur lequel s'est portée la majeure partie de nos travaux, le choix de cibler une classe d'âge particulière n'était pas de première importance.

Le deuxième critère, le niveau de cécité, rejoint un peu le précédent. En effet, il apparaît que le développement d'un système destiné à des personnes malvoyantes ou aveugles repose sur les mêmes problématiques et fait appel à la collecte et au traitement d'informations similaires. La différence se trouve dans le degré de handicap à compenser. Même si cela reste à prouver, l'idée que la compensation d'une déficience totale permette aussi de compenser une déficience moins sévère paraît au premier abord plus fondée que l'inverse.

Le troisième point, la localisation géographique, est complexe, car il se trouve à la croisée de nombreuses problématiques. La raison majeure pour laquelle les personnes souffrant de malvoyance ou de cécité sont plus nombreuses dans les pays en voie de développement par rapport aux pays développés est la faiblesse de leurs systèmes de santé. Une amélioration du système de santé, action préventive et curative, paraît alors plus pertinente que le développement de systèmes d'assistances, action palliative. L'idéal est évidemment d'avoir les deux à la fois ; mais est-ce possible ? La question de savoir à quel point le développement de l'un influe sur le non-développement de l'autre, est complexe. Une focalisation des efforts paraît bénéfique [Lewallen and Courtright, 2001]. À côté de cela, il est difficile de se convaincre que des systèmes d'assistance qui ont des difficultés à s'imposer dans les pays les plus développés puissent être utilisés à grande échelle dans des pays en voie de développement. Il n'est pas impossible que cela se réalise, les innovations ayant potentiellement plus de facilités d'adoption dans des pays qui se développent, on peut voir les exemples de très forte utilisation des technologies de communication

pour les paiements en Chine et dans beaucoup de pays d’Afrique. Le coût, la facilité d’utilisation, l’utilité, l’efficacité et la robustesse de ces systèmes ont permis leur adoption. Mais les conditions préalables pour que cela se produise, au premier rang desquelles figure le niveau d’accompagnement des systèmes de santé et des annexes (associations, structures de formation), ne semblent pas remplies. Même dans ce cas, de nombreuses différences locales de natures diverses, tels les niveaux de revenus, le niveau des infrastructures, ou le niveau de « standardisation » urbain pour en citer quelques-unes sont probablement plus à même d’être prises en compte par des équipes de recherche locales. Toutes ces raisons représentent autant de contraintes additionnelles complexifiant les travaux ciblant les pays en voie de développement.

### 1.2.4 De la problématique à la perception

Cette section vise à expliciter l’intérêt d’étudier la perception sous plusieurs de ses aspects dans le cadre du développement d’un système d’assistance aux personnes malvoyantes.

On a vu que la mobilité est une tâche complexe qui peut être décomposée en fonctions. Ces fonctions font appel à des grands groupes de fonctions cognitives et physiologiques. La motricité est la plus évidente : pour se déplacer, il faut avoir posséder la capacité physique de le faire. Les membres inférieurs assurent ce rôle chez la plupart des personnes – chez d’autres, ce sont les membres supérieurs qui font tourner les roues d’un fauteuil roulant manuel, voire un fauteuil roulant électrique commandé manuellement ou par le regard. La mémoire, la réflexion, les désirs, les émotions... sont également impliqués. La mémoire à long terme est évidemment indispensable car elle conditionne toutes les autres facultés de haut niveau. La réflexion explicite (plutôt que l’intuition) est au moins nécessaire à la détermination des objectifs des déplacements, et elle est souvent mise à profit pour implémenter la stratégie de navigation voire les étapes intermédiaires complexes elles-mêmes. La plupart des autres facultés (désir, émotions, etc.) sont impliquées dans les phases préliminaires. La faculté qui nous intéresse le plus est la perception. Chacune des sous-fonctions du modèle de mobilité repose sur une forme ou une autre de perception :

1. La motricité fonctionne grâce aux différents récepteurs de proprioception ;
2. La marche nécessite de pouvoir éviter les obstacles, et donc de les détecter ;
3. L’orientation nécessite la reconnaissance de points de repères externes, ou bien, pour des durées courtes, un suivi de la direction via les organes sensoriels internes (oreille interne) ;
4. La connaissance spatiale repose sur un apprentissage sensoriel préalable, très probablement réalisé par l’expérience sensorielle elle-même ou la simulation mentale de l’expérience, ce qui constitue une forme de perception spatiale ;
5. La navigation repose sur les autres sous-fonctions et donc sur la perception indirectement. Il est aisé de voir qu’elle est nécessaire dans l’implémentation de la stratégie d’intégration de chemin, celle-ci reposant sur l’identification de points de repères à certains endroits-clés du trajet ;

6. Les fonctions annexes utilisent également de la perception : lire des panneaux, interagir avec des personnes ou des systèmes électroniques ou informatiques, etc.

Tout repose sur la perception, ce qui n'est pas surprenant étant donné qu'elle est le moyen par lequel les informations environnementales nous parviennent. Une question plus intéressante pour nous est de savoir ce qui repose sur la vision. Ce n'est pas le cas de la motricité, ni du suivi de direction à court terme, ni de l'interaction avec des personnes. Dans notre liste, tout le reste dépend de la vision ou l'utilise de manière préférentielle, c'est-à-dire quand elle est disponible.

On peut ainsi considérer que la mobilité repose sur trois piliers : la motricité, la cognition (au sens de réflexion, mémoire, etc.) et la perception (que nous excluons de la cognition pour mieux la faire ressortir). À l'évidence, chez les personnes malvoyantes, malgré la diversité des lésions et pathologies que nous avons présentées plus haut, c'est d'abord la vision qui présente des déficiences. C'est donc elle que nous devons étudier prioritairement. Elle est à l'intersection de la mobilité et de la déficience visuelle : elle est à la source de l'écart que nous cherchons à combler. De plus, tout système visant à améliorer la mobilité des personnes déficientes visuelles devra comporter un ou plusieurs composants perceptifs. Nous devons donc étudier la perception dans une perspective plus large que la seule la perception « adaptée » consécutive à la perte de vision.

## 1.3 Perception

Forme première, voire unique, de contact avec le monde, la perception est une faculté manifestement importante. À ce titre, il est logique qu'elle ait depuis longtemps fait l'objet d'études. La perception est difficile à définir. Parce qu'elle a nécessairement des liens avec de nombreux objets complexes (en particulier, la conscience) et dépend du contexte, la perception, dans son sens le plus large, est difficile à définir. L'isolation d'une « essence » de la perception, qui serait le sujet d'étude de tous les domaines intéressés à la fois, est la manifestation concrète de cette difficulté. Car la perception a de nombreuses facettes, soulève de nombreuses questions, et est en conséquence étudiée par plusieurs disciplines. En effet, même si on se place dans le cadre d'une définition de la perception restreinte aux sens extéroceptifs et proprioceptifs, on constate que plusieurs disciplines et spécialités sont nécessairement impliquées.

Pour comprendre le fonctionnement de la perception, il est nécessaire d'étudier les organes qui la supportent. Cela inclut, bien entendu, les organes transducteurs extéroceptifs (yeux, oreilles, nez, langue, surfaces sensibles), proprioceptifs (permettant de sentir la position des articulations), et le système vestibulaire qui contribue à la station verticale et à la locomotion grâce à la sensation des accélérations et des rotations. Cela inclut aussi le cerveau, où naissent les représentations et s'effectuent les processus cognitifs, et les canaux nerveux, par lesquels transitent les signaux. Ces organes, différents les uns des autres, sont le sujet d'études de domaines spécialisés de la médecine : ophtalmologie, otologie, rhinologie et neurosciences. Ces dernières,

de par la richesse de leur sujet d'étude, couvrent elles-même de très nombreuses spécialités – dont la diversité et la spécialisation pourraient encore être amenées à se développer [Carandini, 2019]. Une compréhension parfaite et entière de ces organes nous permettrait de comprendre et d'expliquer comment se produisent les phénomènes perceptifs, depuis leur expression physiologique jusqu'à la réalisation des processus cognitifs (approche ascendante). Cependant, le degré nécessaire de perfectionnement dans la modélisation et la compréhension de ces organes n'est pas encore atteint. Une telle approche, que nous pouvons qualifier d'ascendante, n'est pour le moment pas implémentée. Pour preuve, le Human Brain Project, important projet européen associant une centaine d'organismes de recherche et environ cinq cent chercheurs [Human Brain Project], vise un objectif préalable à une simulation complète de cette envergure : celui de la mise en place d'une infrastructure et d'outils informatiques permettant l'intégration des données et modèles provenant de différentes études en neurosciences afin d'obtenir un simulateur généraliste pour les neurosciences et les domaines relatifs au cerveau permettant de tester des hypothèses variées [Einevoll et al., 2019]. L'élaboration d'un modèle, allant du fonctionnement individuel des neurones jusqu'à l'observation des capacités cognitives serait une avancée majeure pour la science. Le niveau de compréhension auquel nous avons pour le moment accès ne va pas beaucoup plus haut que ledit fonctionnement individuel neuronal ; le niveau suivant, c'est-à-dire les réseaux de neurones, n'est pas encore entièrement calculable [Einevoll et al., 2019].

Puisqu'il n'est pas encore possible d'appréhender complètement la perception en suivant uniquement une approche basée sur l'observation neurologique, elle doit également être étudiée par une approche descendante, c'est-à-dire en examinant ses phénomènes observables « de l'extérieur ». C'est l'approche classique qui a été suivie d'abord en philosophie puis en psychologie expérimentale. C'est ainsi qu'ont évolué historiquement la plupart des travaux sur le fonctionnement des processus mentaux jusqu'aux avancées des neurosciences dans les années 1970.

### 1.3.1 Progression historique

Étant donnée l'importance centrale de la perception dans la vie des individus, et sa fonction d'interface entre le monde « extérieur » et la conscience, elle a depuis longtemps intéressé les philosophes. Elle était, et reste, un sujet de traité au travers des problématiques qu'elle soulève, parfois même sans être isolée et identifiée en tant que telle. En effet, elle est en général transparente, remplissant ses fonctions sans se faire remarquer. Les situations où elle apparaît, c'est-à-dire où l'on voit qu'un composant de notre organisme sert à percevoir l'environnement, sont les situations de défaillance. Malvoyance et cécité (mais pas seulement) peuvent aussi être révélatrices de la complexité du système perceptif. L'étude de la perception et l'étude des handicaps perceptifs ont ainsi toujours été liés, et les questions posées par les auteurs restent au moins en partie pertinentes aujourd'hui. Un rapide survol de quelques travaux historiques peut nous aider à mieux cerner les liens entre perception, handicap, et assistance.

Dans le *Discours de la Méthode* (1637), Descartes mentionne l'utilisation d'un bâ-

ton pour percevoir les objets par les personnes aveugles [Descartes, 1637]. Marion Chottin a analysé ce texte pour le rapprocher des développements récents des dispositifs d'assistance [Chottin, 2018]. L'usage qui est fait du bâton soulève la question de sa nature, et par extension celle de la nature des systèmes modernes. Descartes ne considère pas le bâton comme un remplacement de la vision, car il ne permet pas de voir, mais ne considère pas non plus la perception avec le bâton comme étant de même nature que la perception tactile commune [Chottin, 2018]. Le bâton est donc une nouvelle forme de perception, un « sixième sens ». Le bâton n'est pas considéré par Descartes comme un véritable *organe* mais il en est néanmoins très proche, en étant « presque » un. Il n'en est pas un parce que ce n'est pas une partie constitutive du corps, qui a pour Descartes un caractère divin.

Le clivage organe/non-organe n'est pas en lui-même très intéressant pour nous, car il n'apporte pas beaucoup d'informations : presque tous les dispositifs dont nous allons parler se trouvent dans la catégorie non-organique. Par contre, comprendre en quoi un bâton est *très proche* d'un organe est intéressant. Qu'est-ce qui détermine l'appartenance à la catégorie « presque-organique » ? Cette question nous intéresse car elle est liée à celle de la recherche de critères pertinents sur lesquels classer les dispositifs d'assistance. Le jugement de pertinence sur lequel repose la séparation de ces catégories est la base sur laquelle repose le critère. Nous devons donc examiner pourquoi le bâton est considéré comme « presque » organique. C'est, *a priori*, l'absence de raisonnement qui confère ce caractère perceptif. La personne munie du bâton ne *pense* pas à la position de ses mains quand elle le manipule ; la perception avec le bâton ne s'effectue pas par une réflexion explicite. Elle est intuitive, en cela similaire aux autres modalités de perception.

La phénoménologie est un courant philosophique qui s'intéresse à la manière dont les objets du monde physique se manifestent à la conscience d'une personne. Le terme « phénoménologie » vient du concept de *phénomène*, qui désigne justement ce qui apparaît à la conscience. Elle fait partie des approches constructivistes, pour lesquelles les représentations formées dans l'esprit ont la prééminence sur une réalité qui serait, dans une approche opposée, dite réaliste, jugée externe, pré-existante, et objective *a priori*. D'un point de vue historique, les approches réalistes ont généralement précédé les constructivistes. Kant a introduit les approches constructivistes, et à ce titre a permis l'émergence de la phénoménologie [Declerck and Lenay, 2018]. On peut aisément comprendre l'importance qu'ont les processus perceptifs dans un mode de raisonnement phénoménologique : ils sont précisément le sujet d'étude de la phénoménologie dans le domaine de la relation entre la conscience et le monde. Si elle ne couvre pas l'entièreté du champ d'investigation de la phénoménologie – domaine large – la perception a donc intéressé les phénoménologues : Merleau-Ponty y a par exemple consacré une part importante de ses travaux [Merleau-Ponty, 2014].

La deuxième moitié du XIX<sup>e</sup> siècle a vu l'émergence de la psychologie en tant que science distincte de la philosophie, et en particulier de la branche centrée sur l'étude scientifique des processus mentaux : la psychologie expérimentale. Ses premiers développements sont justement liés à l'étude de la perception : la psychophysique est en effet la première incarnation de la psychologie expérimentale [Nicolas, 2002]. Se basant sur les travaux de Ernst-Heinrich Weber, Gustav Theodor Fechner a tenté

de systématiser l'étude de la relation entre les phénomènes physiques et les phénomènes psychologiques [Fechner, 1860]. Pour établir cette relation, la psychophysique observe les activités mentales qui sont engendrées par des stimulations extérieures. Cette discipline avait une vocation universelle, cherchant les liens entre âme et corps, mais la mesure directe des phénomènes psychologiques étant impossible à l'époque, la psychophysique s'est bornée, dans ses expérimentations, au domaine de la perception. Ce contexte permet en effet d'obtenir des mesures : seuils de perception, seuils de saturation, et seuils de différenciation. On en retient généralement uniquement cet aspect perceptif. Une formule générale a été proposée : les sensations varient selon le logarithme des stimulations. Son caractère universel est discutable mais elle fonctionne dans de nombreux cas [Foucault, 1901]. Si le modèle de la relation proposée est utile, il n'atteint pas la portée qu'il visait, n'expliquant pas par exemple les comportements et restant éloigné des questions cognitives. Bien qu'ayant été rapidement critiquée [Foucault, 1901], la psychophysique prise dans un sens restreint, reste utilisée [Nicolas, 2002]. Les techniques expérimentales qu'elle a développées restent en particulier pertinentes et, de plus, elle a ouvert la voie à la psychologie expérimentale.

Une théorie psychologique et philosophique du début du vingtième siècle, la *Gestaltpsychologie* (en allemand) ou « psychologie de la forme », s'est construite autour de réflexions sur la perception. La théorie dominante alors était le behaviorisme ou comportementalisme, qui prolongeait les travaux de Pavlov sur le conditionnement, en avançant la thèse que le comportement des animaux, y compris les humains, est déterminé par les réflexes et les réactions aux circonstances externes, en particulier par les réactions apprises par des mécanismes de punition et de renforcement suite à des rencontres antérieures avec des circonstances similaires. Selon cette théorie, le comportement pouvait donc être expliqué scientifiquement en décomposant le vécu en succession d'évènements et de réactions. La *Gestaltpsychologie*, quant à elle, postulait un fonctionnement de l'être humain différent : laissant de côté le principe de réaction, elle se concentrait sur l'émergence de la perception des objets et « formes » (*Gestalt* en allemand). Les objets ne pouvaient être décomposés sans perdre leur sens, le *tout* revêtant un sens différent de, et supérieur à, la somme de ses *parties*. L'application de ce principe la plus claire et la plus médiatisée – car étant la plus facile à transmettre – était la perception visuelle. Le sens d'une image, immédiatement saisi par un humain, se trouvait dans la forme globale qu'elle prenait plutôt que dans l'accumulation de ses détails. Les psychologues de l'école du Gestalt formulèrent des lois permettant d'expliquer comment des éléments se retrouvent inclus dans un même groupe « sémantique » ou objet par la perception visuelle [Rock and Palmer, 1990] (figure 1.1) :

- La loi de proximité ;
- La loi de similarité ;
- La loi de fermeture ;
- La loi de continuation.

Un autre principe important, nommé *Prägnanz* était le choix de la solution la plus simple en cas d'ambiguïté, comme par exemple quand une forme est partiellement occultée : on s'attend alors que la partie qui n'est pas visible ait une forme continuant de manière simple celle de la partie visible. La *Gestaltpsychologie* comprenait

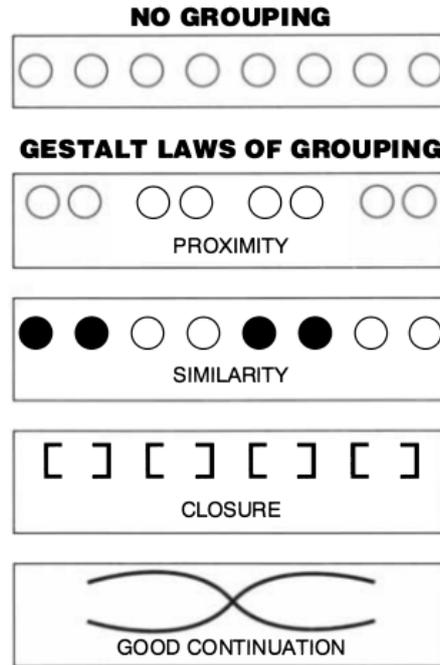


FIGURE 1.1 – Lois de groupage de la *Gestaltpsychologie* [Rock and Palmer, 1990].

également des tentatives de justifications physiologiques de ces lois, avançant une structuration similaire des régions physiologiques impliquées et des phénomènes observés. Ces justifications se sont révélées hasardeuses, de même que les hypothèses sur les localisations neurologiques précises où la perception de formes se réalise, mais les lois ont été globalement vérifiées par la suite [Rock and Palmer, 1990].

La psychologie cognitive a également une visée large, mais contrairement aux théories du Gestalt et du behaviorisme, elle ne fonctionne pas par l'extension d'un principe observé dans un domaine pour expliquer toute l'étendue de la pensée et du comportement. Inspirée par les développements de l'informatique après la seconde guerre mondiale, la psychologie cognitive repose sur la conception des fonctions mentales comme des processus cognitifs traitant des informations. Elle cherche à découpler les fonctions cognitives (langage, perception, mémoire...), à étudier comment elles travaillent et la manière dont elles sont connectées. Afin de répondre au problème, toujours prégnant, de la presque impossible observation directe des phénomènes de pensée, plusieurs approches lui permettent d'acquérir des connaissances. Eysenck en cite quatre [Eysenck, 2001] :

1. La psychologie cognitive expérimentale, qui repose sur des expériences d'observations indirectes traditionnelles (typiquement, mesures différentielles des temps de réactions sur certaines tâches en faisant varier un paramètre significatif) ;
2. La neuropsychologie cognitive, qui étudie des personnes souffrant de lésions cérébrales afin de comprendre le fonctionnement du cerveau ;
3. Les neurosciences computationnelles qui développent des modèles computationnels (Eysenck nomme les neurosciences computationnelles « *Cognitive*

*science* » dans son texte) ;

4. Les neurosciences cognitives, qui étudient comment les structures physiologiques du cerveau supportent la cognition.

La psychologie cognitive, grâce à la souplesse de son principe central – la modélisation des phénomènes mentaux par des processus manipulant des informations, définit un large cadre dans lequel peuvent s’inscrire des travaux variés sur les fonctions du cerveau et du système nerveux. C’est le cas de l’étude de la perception, mais aussi, par exemple, de celles du langage et de la mémoire. Penser ces fonctions dans un cadre commun nous est utile, car il existe des liens entre elles. Ainsi, la mobilité des déficients visuels est-elle très dépendante de leur mémoire : mémoire de travail, mémoire épisodique [Harrar et al., 2018]. Elle est aussi dépendante du langage : d’abord, parce que le guidage par un accompagnateur passe en grande partie par des indications verbales. Ensuite, parce qu’il est probable que la mémoire spatiale est en partie structurée par le langage : Hermer-Vazquez a montré que l’orientation spatiale est liée, dans certaines circonstances, sur les fonctions du langage : ainsi, des sujets testés parvenaient à mémoriser certaines informations par l’intermédiaire de la mémorisation d’un encodage verbal de l’information [Hermer-Vazquez et al., 2001]. De plus, en raison de sa structuration similaire à celle des systèmes informatisés, la psychologie cognitive s’interface bien avec les problématiques liées aux dispositifs électroniques d’assistance.

Dans les années 1960, Gibson a proposé une approche plaçant la perception dans un cadre plus large, en postulant qu’on doit la comprendre en analysant comment elle réalise son « but » [Gibson, 2014] ; le mot *but* étant compris ici non pas dans le sens d’une volonté mais dans celui que la théorie de l’évolution et de la sélection naturelle lui donnent. La perception est l’organe des êtres qui leur permet d’interagir avec leur environnement ; elle doit donc être étudiée en tenant compte de cet environnement. Gibson nomme cela une approche *écologique*, dont découle tout un système de réflexion. Concernant la vision, il en définit quatre types : la vision instantanée d’une image, la vision plus prolongée d’une image avec la tête immobilisée mais les yeux mobiles, la vision *ambiante*, où le sujet peut regarder autour, et enfin, la vision *ambulatoire*, où le sujet peut se déplacer par rapport à la scène. Avant Gibson, les seuls types de vision étudiés par les psychologues étaient les deux premiers, car ils sont beaucoup plus simples à gérer. Or, la plus significative *écologiquement* est la dernière, qui correspond à l’immense majorité des situations rencontrées dans la vie. Pour Gibson, le mouvement est ce qui permet prioritairement à un être d’extraire des *invariants* de l’environnement. Ces invariants concernent des propriétés particulières de l’environnement, et leur détection est à l’origine des perceptions. C’est donc l’étude de la vision ambulatoire qui peut apporter les véritables clés du fonctionnement de la perception visuelle, alors que la vision instantanée n’est qu’une forme dégradée de la vision, ni la plus riche, ni la plus simple. Le plus intéressant pour nous dans la thèse de Gibson est certainement l’importance accordée à l’environnement. D’une part, c’est lui qui génère la perception, et d’autre part, ce sont les interactions possibles avec lui qui le définissent : une grenouille ne perçoit pas autant de choses qu’un humain car elle n’en a pas besoin, et réciproquement, nous ne percevons pas les mouches aussi clairement qu’une grenouille. Ces possibilités d’interactions, que Gibson nomme *affordances*,

ont d'ailleurs été reprises par la suite, y compris dans des domaines éloignés de la psychologie cognitive (le brouillon du standard du W3C Web of Things – Things Description que nous examinerons dans la section 4 utilise le terme [Kaebisch et al., 2019a]). Une approche basée sur les possibilités d'interactions nous est utile car elle met en lumière les différences qui peuvent exister entre la vision humaine normale, la perception en l'absence de vision et, plus encore, la perception avec l'aide d'un dispositif d'assistance (voir section 2).

Après ceux de Gibson, les travaux de Marr sur la vision marquent une étape importante [Marr, 1982]. Particulièrement liés au développement de l'informatique et du mode de pensée qu'elle engendre, ils concernent l'identification de l'essence des informations manipulées par la vision. Marr cherche à caractériser le plus précisément possible ce qu'est la vision en passant par l'analyse des données qu'elle manipule, c'est-à-dire les informations qu'elle fournit. Il propose que l'essence de l'information fournie par la vue est la forme et la configuration spatiale des objets, c'est-à-dire que la vision est avant tout un sens spatial. Marr divise le travail de la vision en plusieurs tâches : représentation des images, détection de contours, de textures, puis de surfaces, puis reconstruction tridimensionnelle. Il cherche pour chacune de ces tâches trois niveaux de caractérisation ou de description. Le premier niveau est dit *computational*. Il correspond, d'une part, à l'utilité de la tâche, ce qui donne son intégration dans le système global, et d'autre part, à la logique de la stratégie qui permet de l'accomplir. Le deuxième niveau correspond à la description du mécanisme qui réalise la tâche, c'est-à-dire à l'algorithme qui transforme les informations. La représentation des données en entrée et en sortie de cet algorithme est également d'une très grande importance. Enfin, le troisième niveau, le plus bas, est celui où s'effectue l'implémentation physique des algorithmes. Ce niveau est réalisé par les neurones et réseaux de neurones du système nerveux, ou bien – et c'est là l'un des intérêts de cette approche – il peut être réalisé par un système électronique ou informatique : il suffit d'implémenter les algorithmes. Cette souplesse dans l'approche de Marr est d'ailleurs une manière pour lui de vérifier la pertinence de ses hypothèses : en comparant les résultats du système visuel humain avec celui de programmes, il juge de la pertinence de ses théories computationnelles. Si les résultats sont similaires, cela indique que le modèle construit et implémenté est une bonne représentation du phénomène réel. En cas de succès et si les performances de l'implémentation artificielle sont suffisantes, l'avantage supplémentaire de cette approche est que l'on dispose d'un algorithme utilisable par les technologies qui pourraient en avoir besoin. Les travaux de Marr en ont inspiré d'autres, et son approche, dite « computationnelle », est devenue une branche à part entière de la psychologie cognitive, située à la frontière entre psychologies et intelligence artificielle.

### 1.3.2 Problématiques centrales pour la perception des PMV

#### Définition de la perception

On a vu certains des angles sous lesquels la perception a été étudiée au cours de l'histoire ; une autre perspective qu'il est important de mentionner est celle de la psychologie clinique, voire de la neurologie, qui s'attache à étudier les maladies du système nerveux, parmi lesquelles celles de la perception. Quelques exemples intéressants sont exposés dans [Bauer, 2006], [Reed et al., 1996], [Efron, 1969]. Une conséquence de la multiplicité de ces approches est qu'il n'y a pas d'acceptation universelle de ce que le concept de perception recouvre exactement. Pour Efron, elle est « [...] la forme principale de contact cognitif de l'Homme avec le monde autour de lui »<sup>2</sup> [Efron, 1969]. Dans un livre de Schacter *et al.* à destination des étudiants en psychologie, on trouve la définition suivante de la perception : « l'organisation, l'identification, et l'interprétation d'une sensation afin de former une représentation mentale »<sup>3</sup> [Schacter et al., 2014]. La première définition indique principalement la place qu'elle accorde à la perception (l'intermédiaire cognitif principal entre le monde extérieur et la personne). La seconde est plus restrictive ; elle décrit – et donc, si on la prend en tant que définition, impose – une certaine structuration des concepts : d'abord, la perception repose sur des *sensations* ; ensuite, elle sert à créer des *représentations mentales* ; enfin, pour ce faire, elle *organise, identifie, et interprète*.

Les problématiques de définition de la perception sont en partie liées à l'établissement de frontières. Est-ce que la sensation fait partie de la perception ? Qu'en est-il de la mémoire spatiale, ou du raisonnement spatial ? Les progrès dans les neurosciences ont permis d'obtenir de nouvelles connaissances sur ces problématiques. Ainsi, par exemple, on sait que la perception visuelle est très étroitement liée aux facultés motrices [Ahissar and Assa, 2016]. Ces progrès ne sont toutefois pas encore parvenus à établir une définition et un périmètre universellement acceptés de la perception : sa relation avec la cognition fait ainsi débat [Pylyshyn, 1999, Firestone and Scholl, 2014, Teufel and Nanay, 2017, Schnall, 2017, Firestone and Scholl, 2017].

En ce qui nous concerne, n'étant pas spécialistes du domaine, nous adopterons une définition large de la perception, qui englobe les « fonctionnalités » dont nous avons besoin, et permet de laisser de côté certains détails qui sont décrits plus haut. En particulier, nous incluons la perception spatiale comme étant une forme de perception nourrissant une *représentation spatiale*, qui est en fait ce qui nous intéresse en premier lieu. Cette représentation spatiale peut être construite à partir de stimuli sensoriels (i.e. perception spatiale), mais aussi de fonctions et d'objets n'émergeant pas directement de *stimuli* : mémoire spatiale, réflexion et imagination spatiales.

---

2. Notre traduction de l'anglais original « [...] *man's primary form of cognitive contact with the world around him* » [Efron, 1969]

3. Notre traduction de l'anglais original « *The organization, identification, and interpretation of a sensation in order to form a mental representation.* » [Schacter et al., 2014]

### Processus cognitifs et niveaux de conscience

La mobilité englobe des tâches conscientes et inconscientes. On peut remarquer cela en l'analysant sous un angle légèrement différent de celui choisi précédemment. Elle n'est pas uniquement constituée d'un déplacement, elle comporte également une phase de détermination de ce déplacement. La première étape est donc, souvent <sup>4</sup>, le choix de l'endroit où l'on va aller. Cette détermination de la destination est toujours, ou presque toujours, consciente.

Une fois la destination choisie, il faut savoir comment y aller. Ici, les tâches peuvent être conscientes ou inconscientes. Selon l'échelle de distance et de complexité de l'environnement séparant les deux positions, différentes fonctions seront utilisées : pour une personne valide, atteindre une porte à une distance de quelques mètres se fait sans réflexion explicite, contrairement à un déplacement pour aller dans une autre ville ou, peut-être, un autre pays, composé d'un premier déplacement en transport en commun vers une gare, puis l'achat de billets de trains, des correspondances, etc. Malgré le fossé qui les sépare, ces exemples sont tous deux des tâches de mobilité. Avec une définition fonctionnellement large de la mobilité, des problématiques sous-jacentes de natures diverses peuvent être rencontrées, comme on l'a vu au-dessus.

Le modèle de mobilité sur lequel nous nous basons prend en compte pour chacune de ses fonctions de base le niveau de conscience [Pissaloux et al., 2017]. Le niveau de conscience requis pour effectuer les tâches y est utilisé pour estimer le niveau de charge cognitive de chaque fonction.

De manière générale, les travaux touchant à la notion de conscience sont très délicats. Depuis Descartes, la conscience a été étudiée par le biais de l'introspection. Cette étude est cependant délicate et éloignée du champ scientifique, en raison de sa fragilité [Efron, 1967]. En effet, la conscience apparaît n'être observable que par le biais de l'introspection, qui est imprécise et semble difficilement à même de fournir des données fiables, puisqu'elle passe elle-même par le prisme de son objet d'étude. La biologie et la psychologie doivent et préfèrent se reposer sur des indicateurs objectifs dont les données ne sont pas générées par le processus qui les fournit : d'où la psychophysique, la mesure des temps de réaction pour la reconnaissance d'objets ou pour la réaction à des stimuli, etc. La définition de la conscience est, biologiquement, globalement floue, car elle est éloignée des notions physiologiques. En effet, ce concept est avant tout philosophique, et il est difficile de le relier précisément à une fonction définie du corps humain. Elle reste néanmoins un objet d'études en neurosciences [Owen, 2019]. Dans nos travaux, nous utilisons également la conscience sans la définir entièrement. Nous nous en servons lorsque nous nous demandons à quel niveau de conscience une tâche effectuée prend place. Cette question est, comme nous l'avons entrevu plus haut, intéressante en raison du lien qui semble exister entre niveau de conscience et charge cognitive – la mémoire semble particulièrement liée à la charge cognitive [Paas et al., 2003]. Le niveau de charge cognitive nous intéresse parce qu'il est source de fatigue mentale

---

4. Nous ne considérons pas ici les cas où l'on se déplace sans but ou sans but précis, par exemple lors de promenades au hasard.

[Mizuno et al., 2011], et la limitation de la fatigue mentale a également au moins deux conséquences souhaitables : premièrement, elle incite les gens qui le possèdent à l'utiliser dans plus de situations ; deuxièmement, elle améliore « l'adoptabilité » du système, c'est-à-dire sa capacité à être choisi et utilisé par plus de personnes.

## Perception spatiale

Pour Marr, la fonction première de la vision est de fournir des informations sur la forme et la configuration spatiale des objets. Les autres fonctions, comme la reconnaissance de couleurs, ou l'illumination sont secondaires [Marr, 1982]. Fournir des informations sur la forme et la configuration spatiale des objets relève de la *perception spatiale*. Les contours de la notion de perception spatiale sont relativement flous, et, comme précédemment avec la perception en général, nous allons utiliser le concept dans une définition large. La vision n'est pas utilisée seulement pour la perception spatiale : par exemple, la vision sert également à la lecture.

Définir la perception spatiale n'est pas très aisé. Le texte de Marr semble en exclure couleurs et illumination. Cela n'est pas si trivial : par leurs variations, elles permettent d'analyser les surfaces et les volumes. Marr ne les voit que comme des intermédiaires permettant d'extraire l'information la plus importante. Mais un autre aspect de la perception spatiale semble délaissé par la description qu'en fait Marr : la *sémantique* des objets, leur identification en tant qu'entités ayant un sens. Là aussi, illumination et couleurs constituent un paramètre important pour les caractériser. Il nous semble également qu'elles soient des éléments sémantiques à elles seules. Se pose alors la question de l'inclusion d'un niveau « sémantique » dans la pile des informations traitées par la perception. La plupart des objets sont identifiables sans réflexion. Le processus aboutissant à leur « disponibilité mentale » se déroule à un niveau de conscience bas. Par disponibilité mentale, nous entendons la possibilité de décision : par exemple, si une table est en vue, on sait (on le *voit bien*) que c'est une table, et que l'on peut l'utiliser. On n'a pas besoin de réflexion particulière pour interagir avec la table – même si les interactions peuvent, elles, engendrer des réflexions. La prise en compte de la table se fait à un niveau de conscience bas. Au contraire, un élément inconnu ou ayant des caractéristiques particulières, par exemple, menaçant, pourra ne pas se fondre naturellement dans la représentation mentale. Il sera alors un point générateur de réflexion ou en tout cas source de processus mentaux de niveaux plus élevés. Si ce n'est pas le cas, un objet inconnu pourra être, par exemple, simplement ignoré.

Dans une approche de psychologie cognitive, la perception spatiale est un processus ou un ensemble de processus manipulant des informations spatiales – qui permettent les interactions nécessaires avec l'environnement. Les informations manipulées forment l'élément central de la perception spatiale. Il est alors logique de s'intéresser à ces informations. Tolman a avancé l'hypothèse d'une organisation en « cartes mentales » des représentations mentales de l'espace [Tolman, 1948]. On a vu que pour Marr, leur nature est, au niveau le plus abstrait, la configuration spatiale et la forme des éléments présents dans l'environnement. Le modèle géométrique qui est commun à ces deux approches est pertinent dans certains cas mais

apparaît inadapté dans d'autres : des erreurs systématiques dans certaines interactions avec l'espace [Tversky, 1992] montrent que la représentation interne n'est pas purement construite sur des objets contenus dans un espace métrique. Si cela ne remet pas en cause la nature à un niveau théorique ou *computationnel* des informations stockées, cela force à examiner plus en détails la véritable manière dont ces informations sont stockées et utilisées, qui forme en quelque sorte leur nature « de fait ».

Tversky a beaucoup travaillé sur les représentations spatiales mentales, et distingue quatre catégories formant les représentations mentales [Tversky, 2005] :

1. L'espace du corps qui est lié à la perception interne du corps, aux sensations ;
2. L'espace autour du corps qui est dimensionné selon la portée de la perception sensorielle<sup>5</sup> et des possibilités d'action directe. Il est généralement structuré par une sorte de repère déterminé par le corps : en haut et en bas, devant et derrière, à gauche et à droite<sup>6</sup>. Les informations qui viennent construire cet espace peuvent venir de la perception sensorielle, de la mémoire, ou bien de représentations externes comme des cartes.
3. L'espace de la navigation sert à gérer les espaces plus grands. Il est structuré en *lieux* et *chemins* qui les relient. Il est reconstruit à partir de morceaux disparates : échelles différentes, sources des informations variées (perception, cartes, mémoire). Il peut être exploré mentalement et exprimé par le biais de deux perspectives principales : la perspective *trajet* et la perspective *relevé* (*survey* en anglais dans le texte). La perspective trajet correspond au point de vue subjectif d'un déplacement imaginaire, localisant les objets par des termes relatifs au sujet. Dans le point de vue relevé, les objets sont agencés selon un repère externe au sujet, objectif. Les deux perspectives semblent conduire, en interne, à une représentation identique, car les performances de sujets utilisant les deux types sont similaires. C'est à cette échelle que se produisent systématiquement des erreurs. Ces erreurs sont liés à la construction de l'espace de navigation à partir d'éléments de plus petite échelle et de règles de simplification : groupement hiérarchique, d'alignement ;
4. L'espace des représentations externes, c'est-à-dire les outils artificiels fabriqués pour faciliter la réflexion : cartes, schémas.

Cette analyse des représentations mentales de l'espace a été utilisée par Pissaloux dans son modèle : chaque tâche ou sous-fonction est liée à une ou plusieurs catégories de l'espace [Pissaloux et al., 2017]. La *marche* prend place dans l'espace autour du corps. L'*orientation* repose à la fois sur l'espace autour du corps et l'espace de la navigation. La *connaissance spatiale* relève de ces deux mêmes catégories, auxquelles il faut ajouter l'espace des représentations externes. La *navigation* est la composition des autres tâches pour atteindre un but et repose donc sur toutes les catégories d'espace.

Declerck et Lenay apportent un point de vue phénoménologique sur la représentation spatiale mentale [Declerck and Lenay, 2018], construit sur les travaux de Heidegger, Merleau-Ponty, Husserl et des autres philosophes phénoménologues. Deux

5. Donc, *a priori*, variant selon les capacités perceptives.

6. Toutes les directions ne sont pas équivalentes d'un point de vue cognitif.

catégories d'actes reposent sur cette représentation spatiale : les actes présents et les actes représentationnels. Les premiers correspondent aux actes où la personne est présente dans les lieux – ou plutôt, où les lieux sont présents autour de la personne – dont la représentation est active. Cela correspond à une interaction directe entre la représentation et l'espace, c'est-à-dire aux phénomènes perceptifs. La seconde catégorie correspond aux actes où, au contraire, l'espace manipulé n'est pas accessible directement. C'est le cas des situations où l'espace est imaginé, extrait de souvenirs ou bien de cartes. En pratique, les connaissances et croyances interviennent en amont et influencent la perception ; dans ce cas, le processus est à la fois présentiel et représentationnel. La représentation elle-même est structurée en quatre modalités.

La première de ces modalités est la structuration en réseau : l'espace est organisé en *lieux* distincts et reliés par des chemins. Le réseau formé peut être regardé à différents niveaux de granularité, avec plus ou moins de détails. Les pièces sont les *lieux* d'un bâtiment, reliés par des couloirs et des portes ; les bâtiments sont les *lieux* d'une ville, reliés par des rues ; les villes sont les *lieux* d'un pays, reliés par des voies ferrées et des routes ; les pays sont les endroits du monde, reliés par des voies terrestres, fluviales, maritimes ou aériennes. La définition du réseau est variable selon les personnes, mais en général, deux nœuds sont qualifiables d'adjacents s'il est possible de passer « directement » de l'un à l'autre : c'est le cas de deux pièces reliées par une porte, mais c'est aussi le cas de Paris et New York, directement connectées par avion. Ce qui constitue un *lieu* dépend de l'usage fait de l'espace, c'est-à-dire, en général, sa fonction sociale : cuisine, aéroport, boutique, etc.

La deuxième modalité gouvernant la représentation de l'espace repose sur les possibilités physiques. L'espace est analysé en fonction des interactions corporelles qui pourraient s'établir avec lui, en particulier, les déplacements qu'il permet. Le premier référentiel pour analyser l'espace est son propre corps. Cela comprend les vides dans lesquels on peut se déplacer, et dans lesquels les objets peuvent prendre place. Ces vides sont déterminés par la matière des objets elle-même, ainsi que, à un autre niveau, par les normes sociales : par exemple, en principe, on n'entre pas dans l'espace dit « de confort » des autres (qui varie selon les cultures), pas plus qu'on ne franchit le comptoir d'une boutique.

La troisième modalité qui règle la représentation est basée sur les différentes possibilités ou hypothèses concernant les objets dans l'espace. Un exemple est la distance d'un point éloigné : la perception de la distance est déterminée par la conscience qu'il est possible ou non d'accéder au point et par les moyens d'y accéder. Les distances perçues ne sont alors pas déterminées par la distance euclidienne mais par les efforts nécessaires pour atteindre ou d'utiliser les objets. Pour certains, la relation avec les objets est conditionnée par l'évaluation de possibilités. Les objets existent en tant que choses qui sont déplaçables, et on les perçoit comme étant dans un endroit sans qu'ils annulent l'espace qu'ils prennent car on considère l'existence de cet espace indépendante des objets qu'il contient. Malgré leur capacité à être déplaçables, les objets bougent peu – et même, rarement d'eux-même – et cela sert de base à leur identité. Si nous trouvons une tasse sur une table, et nous avons déjà vu une tasse identique auparavant au même endroit, nous penserons que ces

deux tasses n'en sont en fait qu'une seule. En revanche, si nous voyons une tasse identique dans un lieu éloigné, nous ne penserons probablement pas qu'il s'agit de notre tasse précédente. Ainsi, le lieu où se trouve un objet participe à l'identité de l'objet.

La quatrième modalité de représentation de l'espace est déterminée par le regard des autres. Les comportements sont en partie contrôlés par la visibilité. L'espace est donc organisé et perçu en fonction de la visibilité : dans certaines zones on est visible, dans d'autres, non. La visibilité (supposée) est jugée dynamiquement à partir de plusieurs paramètres : les personnes présentes, les directions de leurs regards, les orientations de leurs corps, les matériaux dans l'environnement (si on se trouve derrière une vitre fumée par exemple), les objets présents comme des caméras. Cette évaluation de l'espace sous l'angle de la vision peut être étendue à l'audition, qui cependant rend les choses un peu plus complexes. Il est difficile de savoir si d'autres peuvent nous entendre ; cela nécessite typiquement que les autres émettent des sons.

Declerck et Lenay ont ensuite analysé, avec la même approche phénoménologique, en quoi les représentations spatiales mentales de personnes aveugles pouvaient différer de celles de personnes voyantes [Declerck and Lenay, 2018]. Premièrement, ils ont noté un rôle beaucoup plus important de la mémoire chez les personnes déficientes visuelles. En effet, la perception visuelle est, entre autres, utilisée par les personnes voyantes comme « mémoire externe » : typiquement, pour trouver un objet, une personne voyante n'a besoin que de se souvenir vaguement de sa position et ensuite de le chercher visuellement à cet endroit. Ce n'est pas le cas d'une personne handicapée visuellement : elle doit se souvenir de la position exacte de l'objet. Les interactions spatiales sont alors beaucoup plus dépendantes de la mémoire. Comme les personnes déficientes visuelles ne disposent pas d'une mémoire spécialement performante, l'adaptation se fait dans l'autre sens : elles simplifient leur environnement afin de faciliter le travail de mémorisation. Les choses sont donc soigneusement rangées, à une place connue. Un tel niveau d'organisation est délicat à maintenir et facilement perturbé par des personnes extérieures. De plus, il n'est applicable qu'au domicile de la personne ; l'extérieur reste un domaine dans lequel il est difficile d'être autonome. Deuxièmement, se pose la question de la facilité d'intégration du concept d'occlusion spatiale. Ce concept est en effet étroitement lié à la vision ; les manifestations d'occlusion accessibles aux personnes malvoyantes sont moins saisissantes que celles accessibles aux personnes voyantes : sensation du soleil, occlusion auditive (mais un brouillage par une autre source peut avoir des effets similaires). Cela dépend évidemment de la pathologie : les personnes ayant déjà vu ne sont pas affectées par ce problème. L'occlusion est un élément important de la représentation spatiale avancée par Declerck et Lenay : elle est liée à la troisième modalité, qui a trait à l'analyse des possibilités de présence et d'absence, et surtout à la dernière modalité, la perceptibilité par les autres. Avec des possibilités limitées pour percevoir la présence et le regard des autres, la quatrième modalité est peu à même d'enrichir la représentation spatiale.

Malgré les différences entre les approches et les analyses de Tversky et de Declerck et Lenay, une similarité existe entre leurs travaux respectifs. Cette similarité tient

principalement à une division entre *espace environnant* et *espace navigué*, séparés dans les deux analyses des représentations. Chez Tversky, la frontière est claire, puisque l'analyse est en partie au moins basée sur les distances : elle se trouve entre l'espace autour du corps et l'espace de la navigation. Chez Declerck et Lenay, elle se trouve principalement entre la modalité en réseau et les autres modalités. On a vu plus haut la correspondance entre les différents espaces de Tversky et les sous-fonctions de la mobilité du modèle de Pissaloux. La séparation en deux parties de la représentation spatiale n'est pas nette : si la marche apparaît clairement comme reposant sur le domaine environnant, l'orientation fait appel aux deux domaines à la fois. Dans le chapitre suivant, on verra que cette division refait surface sous une autre forme dans l'état de l'art des aides électroniques à la mobilité.

# Chapitre 2

## Dispositifs artificiels, état de l'art

### 2.1 Introduction

De nombreux objets, techniques et principes permettent d'améliorer la vie des personnes déficientes visuelles. Une grande variabilité peut être observée parmi ces dispositifs d'assistance. Leur succès est néanmoins rare : d'une part, le développement d'un dispositif à même de fournir une réelle assistance est difficile, et d'autre part, l'adoption et l'utilisation d'un tel dispositif est soumise à de nombreuses contraintes.

#### 2.1.1 Historique de l'assistance aux personnes déficientes visuelles

L'utilisation du bâton comme outil d'assistance aux personnes déficientes visuelles semble très ancienne. Descartes atteste de son usage dans le *Discours de la Méthode* [Descartes, 1637] (figure 2.1). Et il en parle comme d'une chose établie. En effet, pour une personne malvoyante, utiliser un bâton comme extension tactile semble naturel et on voit mal ce qui pourrait empêcher quelqu'un de saisir un morceau de bois et de s'en servir ainsi. Le bâton, premier dispositif d'assistance, diffère en cela de la plupart des autres dispositifs d'assistance ultérieurs : son utilisation ne dépend pas d'une réflexion élargie préalable. Ce qui est entendu ici par « réflexion élargie », est une forme de réflexion qui examine non pas le sujet individuellement, mais plutôt entouré d'un environnement social et matériel. La praticité directe de l'utilisation du bâton paraît relativement évidente, quel que soit le contexte, dispensant de la nécessité d'une telle réflexion.

L'utilisation de chiens pour aider les personnes malvoyantes paraît également ancienne. Fishman mentionne des représentations graphiques trouvées à Pompéi (I<sup>er</sup> siècle ou avant), en Chine au XIII<sup>e</sup> siècle, et une description en Europe occidentale, également au XIII<sup>e</sup> siècle [Fishman, 2003]. En terme de réflexion, par rapport à l'utilisation d'un bâton qui peut être directe, le dressage d'un chien en vue d'assister une personne handicapée nécessite une anticipation, du fait de la durée du



FIGURE 2.1 – Reproduction d’une planche de la *Dioptrique* de Descartes dans la *Lettre sur les aveugles* de Diderot ; la planche daterait de 1724 [Diderot, 2011] (image tirée de l’édition en ligne disponible sur le site de la BNF : <https://gallica.bnf.fr/essentiels/diderot/lettre-sur-les-aveugles>).

dressage et de l’analyse des cas pouvant être rencontrés. La complexité de cette analyse et du dressage était cependant sans doute moindre dans un contexte historique d’ancien régime par rapport à aujourd’hui : en effet, une bonne partie des structures devant faire l’objet d’un apprentissage n’étaient pas en place. Par exemple, les passages piétons n’existaient pas et leur traversée n’était donc pas une technique à connaître pour le chien. Son rôle pouvait se borner à éviter les obstacles, simplifiant grandement la problématique du dressage.

La création d’une, puis de plusieurs écoles spécialisées, par Valentin Haüy à la fin du XVIII<sup>e</sup> siècle<sup>1</sup> et du braille au XIX<sup>e</sup> siècle [Jiménez et al., 2009] par Louis Braille (passé par l’école de Valentin Haüy) représente un degré de plus à la fois dans l’attention portée à la société pour les personnes déficientes visuelles et dans le renforcement de la réflexion autour de ce handicap. En effet, le braille induit une réflexion globale et non plus limitée à des individus ou à des groupes restreints. D’une part, en effet, le public potentiellement concerné est large ; d’autre part, le braille est lié à la société « extérieure » par l’apport de textes qu’il permet. Par la suite, le braille sera utilisé à plus large échelle dans plusieurs formes de communications : en France, par exemple, les indications sur les boîtes de médicaments sont écrites en braille<sup>2</sup>, les indications dans les établissements recevant du public, qui doivent légalement être accessibles, le sont souvent grâce à l’emploi du braille.

---

1. Institut National des Jeunes Aveugles – Historique – Valentin Haüy : <http://www.inja.fr/Default/valentin-haui.aspx> (accès le 12 septembre 2019)

2. Décision du 7 mai 2008 prise en application de l’article R. 5121-138 du code de la santé publique, <https://www.legifrance.gouv.fr/eli/decision/2008/5/7/SJSM0820334S/jo/texte> (accès le 12 septembre 2019)

Au XX<sup>e</sup> siècle, la canne blanche est standardisée ; elle a deux fonctions. Premièrement, c'est une version améliorée ou, au moins, conçue pour remplir la fonction du bâton d'autrefois, c'est-à-dire pour sentir par le toucher les obstacles environnants : elle est donc très légère et munie d'un embout spécial qui glisse, frotte, ou frappe les objets et aspérités environnantes. Deuxièmement, c'est la nouveauté principale, elle est d'une couleur remarquable et visible de loin – souvent blanche, ou jaune clair selon les pays (la Belgique a un temps imposé le blanc aux personnes aveugles et le jaune à celles malvoyantes sans être aveugles<sup>3</sup>) – afin de permettre l'identification de son porteur comme étant atteint de déficience visuelle. Elle nécessite donc, elle aussi, une réflexion élargie : en fait, son fonctionnement implique autant la personne qui s'en sert que celles et ceux qui l'entourent. Les chiens-guides servent un rôle similaire de facilitation de la reconnaissance.

La seconde moitié du XX<sup>e</sup> siècle a vu de nombreux développements dans le domaine de l'aide à la mobilité des personnes déficientes visuelles. Ces développements, qui ont en partie trouvé leur origine dans la fin de la seconde guerre mondiale et la prise en charge de nombreux vétérans handicapés, ainsi que dans le développement de technologies de détection à distance (radar, sonar), se sont généralement inscrits dans les tendances de standardisation aux échelles nationales puis internationales. Des équipes de recherche dans les hôpitaux militaires américains ont cherché à améliorer les conditions de vie des vétérans ayant perdu la vue pendant le conflit [Sauerburger, 1996]. Deux familles principales de résultats sortirent de ces travaux. D'une part, un fort développement et une régularisation des techniques de mobilité adaptée à la déficience visuelle, appelées *Orientation and Mobility* (orientation et mobilité) accompagnée d'une standardisation de l'utilisation de la canne blanche [Brabyn, 1982]. D'autre part, les chercheurs tentèrent de mettre au point des dispositifs d'assistance électroniques permettant d'améliorer les cannes blanches. Ces dispositifs sont appelés *Electronic Travel Aids (ETA)*, ce que nous pouvons traduire par « Aides Électroniques à la Mobilité » (AEM). Ces dispositifs nécessitent, eux aussi, une réflexion poussée. Ils manipulent en effet des informations de l'environnement pour les transmettre à leurs utilisateurs ; le traitement de ces informations nécessite la prise en compte explicite de nombreux paramètres environnementaux, au minimum lors de la conception du système, si ce n'est directement dans la logique le gouvernant. Une omission dans cette prise en compte aura bien plus de chances d'aboutir à une situation d'échec du système à remplir sa fonction qu'à un succès imprévu. Or, dans ce domaine, l'échec peut avoir de graves conséquences.

Les dispositifs urbains – aménagements visant à faciliter la mobilité – ont également été inventés après la guerre. En 1965 au Japon, Seiichi Miyake conçoit des bandes podotactiles qui permettent de suivre un chemin avec une canne ou les pieds sans le voir [Sakaguchi et al., 2000]. Ce principe a été repris dans beaucoup d'endroits dans le monde et n'a cessé d'être développé depuis [Lu et al., 2008]. Les dalles podotactiles sont toutefois limitées. D'une part, si elles aident à faciliter à la fois l'aspect *marche* en signalant les zones dangereuses, et l'aspect *orientation* de la mobilité en indiquant les directions principales, elles ne remplacent pas l'identification distante de points que permet la vision. D'autre part, et c'est probablement

---

3. Informations de <http://cannejaune.org/loi-belgique.php> et <http://questions.assemblee-nationale.fr/q13/13-18646QE.htm> (accès le 12 septembre 2019)

là leur principale limite, elles ne sont présentes que là où elles sont installées, et leur installation n'est pas sans contraintes :

1. Pour être efficaces, elles doivent être prises en compte lors de la conception des chemins [Lu et al., 2008];
2. Elles génèrent un surcoût [Bashin, 2001];
3. Elles modifient l'apparence de l'environnement urbain, et peuvent être jugées comme inesthétiques.

Pour ces raisons, les dalles podotactiles sont, en général, installées uniquement dans les lieux les plus fréquentés des grandes villes, et aux croisements. Leur pertinence aux croisements semble être discutée [Bashin, 2001], mais cela pourrait ne concerner que la difficulté de mise en œuvre correcte ou la non adaptation à certains environnements.

## 2.2 Aides électroniques à la mobilité

Depuis les années 1960, de nombreuses aides électroniques ont été développées. Les premiers travaux se sont concentrés sur des améliorations de l'aide de référence, c'est-à-dire la canne blanche. Nous appelons cette catégorie d'aides les cannes électroniques « simples ». La liste dressée ici est loin d'être exhaustive, elle vise à mettre en lumière la diversité des familles et principes existants d'aides électronique à la mobilité.

### 2.2.1 Cannes électroniques « simples »

Ces cannes sont conçues comme des extensions de la canne blanche traditionnelle (et du bâton), dans son rôle de détection d'obstacles – le rôle d'identification pouvant être conservé ou non. Bien que toutes ces aides ne prennent pas une forme allongée caractéristique ou ne viennent pas se greffer sur des cannes standards, nous les nommons *canne*s en raison de leur fonction similaire. Elles prennent la forme d'un capteur à tenir en main. De plus, dans ce cas, l'appellation n'est pas non plus complètement imméritée car leur usage s'apparente très approximativement à celui d'une canne « virtuelle ». Dans cette classe, nous regroupons deux familles d'aides électroniques identifiées par Brabyn [Brabyn, 1982] et Andò [Ando, 2003] (qui citent tous deux Foulke [?]). D'une part, il y a les *capteurs d'obstacles* indiquant la présence d'obstacles ou à l'inverse, d'un espace libre, de manière binaire ou avec une distance approximative. D'autre part, il y a les *capteurs environnementaux*, qui indiquent avec beaucoup plus de précision la distance des surfaces visées, et permettent de fournir une assistance à la perception.

#### Capteurs d'obstacles

L'appareil de Russel, le *Pathsounder* est une des premières aides à la mobilité basées sur un capteur ultrasonique [Russel, 1965]. Le *Pathsounder* est conçu pour être



FIGURE 2.2 – Le capteur d’obstacles *Pathsounder*, à gauche, se porte en collier sur la poitrine ; l’appareil visible à droite est son chargeur [Russel, 1965].

porté en collier sur la poitrine (figure 2.2) et utilisé conjointement avec une canne blanche, et il indique la présence d’obstacles par une variation sonore graduelle [Brabyn, 1982].

Le capteur *Mowat* est un autre capteur ultrasonique indiquant les obstacles. Il est prévu pour être utilisé conjointement avec une canne blanche et contrairement au *Pathsounder*, il communique ses informations par voie tactile [Pressey, 1977] [Brabyn, 1982].

La *Laser Cane* détecte les obstacles avec, comme son nom l’indique, un capteur laser [Jansson, 1975]. Elle possède des interfaces audio et tactiles [Brabyn, 1982].

Le *Miniguide* est un capteur ultrasonique commercial récent (figure 2.3) ; la dernière version date de 2005 [Research, 2005] et coûte actuellement 360 livres sterling<sup>4</sup>. Il semble posséder une large base d’utilisateurs. Avec un capteur ultrasonique, le *Miniguide* permet de signaler les distances des obstacles via une interface soit audio, soit tactile (vibrations). La distance de détection est réglable. Il est conçu pour aller de pair avec une canne ou chien, car il ne peut pas détecter les renforcements et trous au sol. Il a été étudié conjointement avec le système *Breeze*, décrit plus bas, dans une étude publiée en 2014 [Vincent et al., 2014]. Les participants de cette étude ont noté sa fiabilité et l’aide qu’il apporte dans la recherche de portes ; ils en étaient globalement satisfaits. Certains défauts ont été rapportés : l’appareil est inutilisable dans les environnements encombrés, car il vibre tout le temps ; il est nécessaire de le maintenir à l’horizontale, ce qui n’est pas évident ; il est également difficile de se souvenir dans quel mode il se trouve (distance de détection).

Le *Tom Pouce* est un détecteur d’obstacles fonctionnant grâce à un ensemble de capteurs infrarouges qui permettent de détecter les obstacles dans plusieurs directions [Farcy et al., 2006]. Fixé sur une canne blanche, il permet à la fois d’étendre

4. On peut le trouver en vente à l’adresse <http://pamtrad.co.uk/product/miniguide/> (accès le 26 août 2019)



FIGURE 2.3 – Le miniguide [Research, 2005].

la portée de celle-ci et de détecter les obstacles qui lui échappent, c'est-à-dire les obstacles en hauteur sans trace détectable au sol (par exemple, les volets d'une fenêtre). La présence d'un obstacle est signalée par des vibrations appliquées sur un doigt de l'utilisateur.

### Capteurs environnementaux

Le premier dispositif qui nous intéresse ici est le premier appareil électronique d'aide à la mobilité largement référencé dans la littérature : le capteur de Kay, nommé *Sonic Torch* [Kay, 1964]. Destiné à remplacer la canne blanche, il comporte un capteur ultrasonique et un haut-parleur. Quand l'utilisateur scanne son environnement avec la *Sonic Torch*, le signal ultrasonique est converti en signal audio. Le codage en signal auditif de la distance est réalisé par une variation de la fréquence audible, et le timbre permet d'identifier la surface qui a réfléchi le signal [Brabyn, 1982].

Le *Teletact*, qui est conçu pour être monté sur une canne classique (figure 2.4) utilise un capteur laser pour détecter les obstacles [Farcy et al., 2006]. Grâce au faible angle du cône de dispersion du laser, les utilisateurs scannant leur environnement avec un *Teletact* peuvent « sentir » le profil des obstacles présents dans l'environnement. Deux interfaces sont disponibles : une interface tactile qui code les obstacles détectés en vibrations sous des doigts différents selon les distances, ou bien une interface audio. Cette dernière permet plus de précision et de finesse dans la détection de l'environnement ; les distances jusqu'à 15 m sont codées par 28 notes, ce qui transforme le profil balayé de l'environnement en « mélodie ». C'est elle qui, nous semble-t-il, donne sa place au *Teletact* dans la catégorie des capteurs environnementaux plutôt que celle des simples détecteurs d'obstacles.

La *EyeCane* est une canne électronique avec des interfaces tactiles et audio [Maidenbaum et al., 2014]. Bien que classique dans son principe – elle détecte les obstacles avec des capteurs infrarouges et signale leurs distances par des signaux audios et vibratoires, elle est intéressante pour l'accent que ses concepteurs ont mis sur ses performances et caractéristiques non-fonctionnelles. Maidenbaum *et al.* ont observé



FIGURE 2.4 – Un Teletact (capteur laser) monté sur une canne blanche [Farcy et al., 2006].

que malgré l'accueil positif qu'elles reçoivent, les aides électroniques à la mobilité peinent à être acceptés et utilisés. Pour ces auteurs, cela tient à des problèmes faisant obstacle à l'utilisation de ces dispositifs dans la vie courante : taille et poids gênants, manque d'autonomie des batteries, coût prohibitif, manque de fiabilité, interférences avec les autres modalités sensorielles. Le défaut le plus particulièrement souligné est la difficulté d'apprentissage. L'*EyeCane* a donc été conçue pour être pratique et facile à utiliser après un apprentissage court ; dans les expérimentations, cinq minutes suffisent à la rendre utile aux utilisateurs. Il est à noter que l'*EyeCane* acquiert des données « fines » sur le relief de l'environnement car ses capteurs ont des cônes de détection étroits, ce qui permet à ses utilisateurs de ressentir leur environnement directement (attribution distale).

## Discussion

Les cannes électroniques simples sont très nombreuses et semblent, à première vue, similaires. Il existe en fait des différences importantes entre elles, et ce sur tous les points qui les caractérisent. Deux aspects des capteurs au moins ont une forte influence sur la fonction des appareils : le type ou principe du capteur et l'ouverture du cône de détection. Trois technologies dominent : les capteurs ultrasoniques, les capteurs infrarouges et les capteurs lasers.

Les capteurs ultrasoniques sont souvent utilisés. Ils ont un coût limité et une consommation basse pour une portée raisonnable : les capteurs MaxSonar EZ-0 consomment par exemple moins de 3 mA pour une portée de détection de 1,5 m pour des cylindres de 6,1 mm de diamètre et jusqu'à 6 m pour des obstacles à surfaces planes [Max, 2015]. Ces derniers souffrent de problèmes de réflexions spéculaires sur les surfaces se présentant avec un angle de plus de 40 degrés, et ne permettent pas de détecter des petites ouvertures [Farcy et al., 2006]. Les capteurs optiques – qui comprennent IR et laser – ne sont pas touchés par ces problèmes. En retour, ils sont sensibles à la lumière du soleil [Farcy et al., 2006] et ont une consommation plus élevée.

L'ouverture du cône de détection des capteurs est une caractéristique déterminante pour l'utilisation d'une aide électronique à la mobilité. En effet, un cône de détection large ne permettra pas à l'utilisateur de distinguer les détails dans les formes présentes dans l'environnement, les éléments les plus proches présents dans le cône étant les seuls dont la distance sera mesurée. Dans ce cas, on sait seulement si un obstacle – suffisamment gros pour être détecté – se trouve dans le cône de détection, ce qui donne une grosse granularité aux informations reçues. On doit alors scanner l'environnement jusqu'à trouver un endroit sans obstacle. À côté de cela, plus le cône de détection sera large, plus le balayage devra être d'une grande ampleur pour que le signal du capteur change, et plus la perception sera donc grossière.

La perception directe passe par des boucles sensorimotrices [Ahissar and Assa, 2016], qui sont bien plus directes avec des appareils ayant un cône de détection étroit ou un laser ponctuel, et qui permettent donc une perception fine. C'est d'ailleurs leur objectif, comme dans le cas du Teletact : « *The important thing is to be able to imagine the obstacle which generated the sounds as if the eyes had temporarily been opened. The person has to be able to go back to the obstacle at any moment without the help of the device.* » [Farcy et al., 2006]. On voit ici que la personne peut construire, grâce à l'objet, une représentation spatiale assez précise. Ce niveau de maîtrise, que tous ne parviennent pas à atteindre, requiert un apprentissage. Les sessions de formations sont données par des instructeurs de mobilité, qui doivent eux-même être formés pendant un mois et pratiquer pendant un an [Farcy et al., 2006]. Si le Teletact et le Tom Pouce jouissent d'un certain succès en France, c'est non seulement grâce à leurs qualités intrinsèques mais aussi grâce à la structure de formation qui les accompagne.

Les appareils appartenant à la deuxième catégorie, ont un fonctionnement qui se rapproche en fait de celui des systèmes de substitution sensorielle que nous allons présenter dans la section suivante.

### 2.2.2 Supplémentation, substitution sensorielle

La substitution ou supplémentation sensorielle a été développée pendant les années 1960 par Bach-y-Rita. Son principe est de transmettre les informations qui transitent normalement par une voie sensorielle (très souvent, la vision) par une autre voie sensorielle : tactile, électro-tactile ou bien auditive.

#### Tactile ou électro-tactile

Le premier système de substitution sensorielle a été développé par Bach-y-Rita [Bach-y Rita et al., 1969]. Il a eu un impact important sur de nombreuses équipes de recherche qui ont à sa suite développé de nombreux systèmes de substitution sensorielle. Ce système permet de faire transiter des informations « visuelles » par le biais de vibrations sur la peau. Les images sont capturées par une caméra de télévision manipulée par les participants et les vibrations produites par une matrice de 400 solénoïdes en contact avec leur dos (figure 2.5). Les expérimentations ont mon-

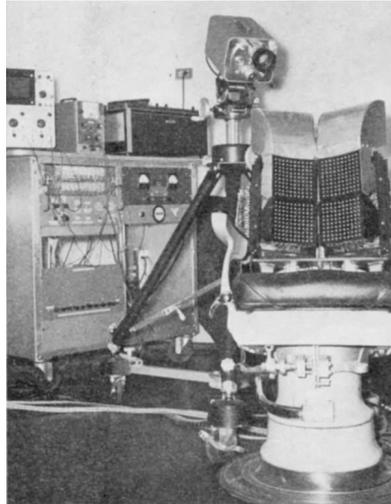


FIGURE 2.5 – Caméra et fauteuil instrumentés utilisés pour le système de substitution sensorielle de Bach-y-Rita *et al.* [Bach-y Rita et al., 1969].

tré que ce système permet à des personnes aveugles de naissance de reconnaître des lignes, des objets simples et même des objets présents mais partiellement occultés dans un environnement. La capacité de reconnaissance et le temps passé à observer avant reconnaissance sont très dépendants du temps d'entraînement préalable avec l'appareil.

Bach-y-Rita et son équipe n'ont ensuite pas cessé de concevoir de nouveaux appareils, de les tester et d'explorer de nouvelles pistes de substitution sensorielle. La langue offre une surface sensitive d'un encombrement moindre et d'une précision plus importante que le dos, car elle présente une très importante densité de récepteurs. Un système caméra-langue a été développé [Bach-y Rita et al., 1998]. L'utilisateur porte une paire de lunettes spéciale comportant une caméra, dont les images sont transposées sur une matrice de 7 x 7 stimulateurs sur la langue. L'objectif affiché était que ce système puisse à terme s'affranchir de câbles et fonctionner sans fil, permettant un port discret et une utilisation au quotidien.

Hsu *et al.* ont créé un dispositif de substitution visuo-tactile en tant qu'outil d'étude des aspects actifs de la perception [Hsu et al., 2013]. Bien que le système de substitution constitue ici un outil et pas une fin en soi, il est néanmoins pour nous un objet intéressant à part entière. L'interface est constituée par trois matrices de 32 x 32 stimulateurs, placés sur trois doigts, et la partie sensorielle par une caméra manipulée par la personne. Le point le plus important pour nous se trouve entre les deux aspects, à savoir le traitement appliqué aux images pour les rendre intelligibles. Ce traitement comporte plusieurs étapes : diminution de la résolution, filtrage passe-bas, seuillage, puis filtrages morphologiques (érosion et dilation de l'image) et enfin une nouvelle diminution de la résolution de l'image. Cette succession de filtres permet de faire ressortir les éléments importants des scènes (figure 2.6).

Kajimoto, Suzuki et Kanno ont développé un système de substitution visuo-tactile basé sur un smartphone [Kajimoto et al., 2014]. La partie matérielle est compo-

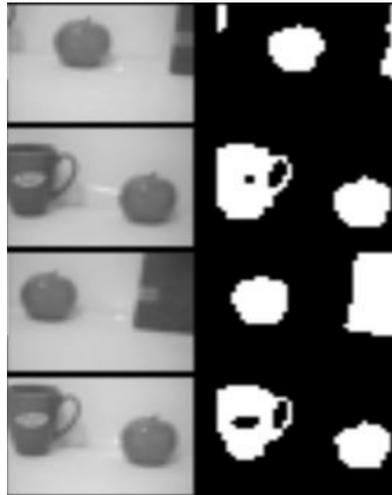


FIGURE 2.6 – Hsu *et al.* utilisent un algorithme de filtrage d’images pour faire ressortir les objets (à gauche, l’image en résolution diminuée et niveaux de gris ; à droite, la sortie de l’algorithme) [Hsu et al., 2013].

sée du smartphone et d’une matrice de 32 x 16 stimulateurs électriques plaquée sur celui-ci. Le smartphone est utilisé pour acquérir les images et les traiter. La transmission à la matrice de stimulateurs se fait d’une manière originale : l’écran du téléphone affiche les images traitées, qui sont « captées » (filmées de manière rudimentaire, en fait) par une matrice de phototransistors homogène et connectée avec les stimulateurs (figure 2.7). Le choix du smartphone permet de disposer d’un matériel performant, en particulier pour ce qui concerne la caméra, et de pouvoir mettre facilement le dispositif à jour : la chaîne logicielle est une simple application smartphone, et le matériel peut être changé facilement par les utilisateurs, pour suivre par exemple les progrès qui ne manqueront pas d’arriver dans le domaine des caméras de profondeur sur smartphone.

### Sonification

La sonification est la transmission sonore d’informations à l’exclusion du langage [Strumillo et al., 2018]. Dans le cas qui nous intéresse, on transmet les informations que la vision devrait acquérir, chez une personne avec une vision correcte, et c’est une forme de substitution sensorielle.

Le premier système développé est probablement le successeur et extension de la *Sonic Torch* de Kay : les *Sonic Glasses* [Kay, 1974]. Elles sont en effet considérées à la fois comme une aide à la mobilité et comme un système de supplémentation sensorielle par Lenay *et al.* [Lenay et al., 2003]. Elles permettent en effet de reconnaître des textures. Le capteur Mowat est aussi considéré par Lenay *et al.* comme un hybride entre une aide à la mobilité et un système de supplémentation sensorielle [Lenay et al., 2003]. Les objets que nous avons catégorisés comme « capteurs environnementaux » sont en fait de tels hybrides : ils sont avant tout des aides à la mobilité, mais fonctionnent ou peuvent fonctionner comme des systèmes de supplémentation sensorielle.

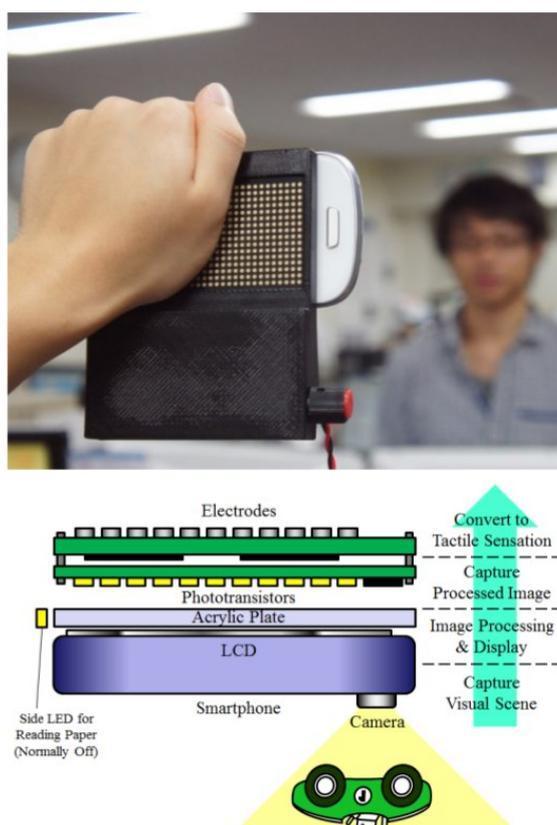


FIGURE 2.7 – Le système de Kajimoto, Suzuki et Kanno. En haut, une photo du dispositif en cours d'utilisation ; en bas, une représentation schématique de son fonctionnement matériel et logiciel [Kajimoto et al., 2014].

À partir du début des années 1990, Meijer a développé le système *The vOICe* pour transmettre des images sous forme auditive [Meijer, 1992]. Ce travail visait particulièrement à démontrer qu'un tel système pouvait être portable, économe en énergie et abordable, car basé sur des composants standards. Un morphisme de conversion simple est utilisé pour transformer en sons les images de 64 x 64 pixels en 16 niveaux de gris : les colonnes de pixels sont présentées séquentiellement ; chaque pixel (ou ligne) d'une colonne correspond à une fréquence ; un pixel, qui se trouve donc repéré par le moment où il est présenté (colonne) et une fréquence (ligne), voit son niveau de luminosité codé par l'amplitude de son signal. Toutes les fréquences des lignes d'une colonne sont additionnées et le signal sonore résultant est diffusé. La succession temporelle des colonnes forme une sorte de mélodie. Deux images successives sont séparées par un clic audible. Les formes simples sont faciles à identifier : par exemple, une ligne montante donne une unique fréquence qui augmente, un rectangle uniforme donne un bruit blanc filtré par un passe-bande. Les images complexes sont évidemment moins simples à interpréter, mais l'auteur parie sur la possibilité d'une analyse consciente qu'offre la simplicité de la conversion pour aider à l'apprentissage. Le système a été amélioré et est disponible – plutôt que le système complet du début, il s'agit d'une application (qui semble gratuite) disponible sous Android et Windows ; le matériel conseillé est une paire de lunettes intelligentes fonctionnant sous Android, les VISION-800 [Meijer, 2019]. Les apports de l'appareil ont été étudiés par Auvray, Hanne-ton et O'Regan [Auvray et al., 2007]. La version utilisée fonctionnait avec une webcam tenue à la main, un ordinateur portable, et un casque audio fermé. Les expérimentations évaluaient la capacité des participants dans des tâches de localisation et d'identification d'objet. Dans les deux cas, les participants ont réussi à accomplir les tâches requises.

Le projet *Sound of Vision* a été réalisé dans le cadre de l'appel à projet *H2020* de l'Union Européenne. Son objectif était de créer un système d'assistance portable (*wearable* en anglais) ayant suffisamment de capacités pour améliorer la situation des personnes malvoyantes dans les domaines de la perception environnementale et de la mobilité, en intérieur comme en extérieur, et ce, sans dépendance à une infrastructure [Sound of Vision Consortium, 2018]. Le système développé fournit plusieurs fonctionnalités et est à ce titre décrit plus en détails plus bas dans le texte. Dans un premier temps, nous allons nous attacher à décrire le fonctionnement de la partie centrale du système, qui transcrit les images sous forme sonore. L'utilisateur porte sur la tête un appareil comprenant une caméra stéréo 3D (figure 2.8), une caméra de profondeur, et une centrale inertielle [Caraiman et al., 2017]. La caméra stéréo est privilégiée en extérieur, et la caméra de profondeur en intérieur, mais les deux peuvent être utilisées en parallèle et leurs données fusionnées. La centrale inertielle sert à connaître l'orientation de la tête par rapport à l'environnement. Le sol et le mur sont détectés, ainsi que les objets les plus pertinents, dont le nombre peut-être paramétré pour éviter la surcharge cognitive. Les éléments spéciaux détectés sont les suivants : les trous et renforcements dans le sol (escaliers descendants par exemple), les portes, les escaliers, les panneaux indicateurs et les textes. Quatre modalités de présentation sonore ont été expérimentées. Elles sont des reprises ou améliorations des mécanismes utilisés dans le projet Naviton, qui précédait Sound of Vision [Strumillo et al., 2018]. Un son particulier est assigné à



FIGURE 2.8 – Les différents éléments développés dans le projet Sound of Vision [Sound of Vision Consortium, 2018].

chacun des éléments détectés : murs, escaliers, etc. La première modalité fonctionne de la manière suivante : la scène est balayée virtuellement par l'interface, du proche vers le lointain ; la progression du balayage est marquée, mètre après mètre, par des impulsions sonores de plus en plus discrètes. Quand le « faisceau » rencontre un objet, le son lui correspondant est diffusé de manière localisée, donnant l'impression que le son provient du lieu où se trouve l'objet. La hauteur et l'amplitude du son sont inversement proportionnelles à la distance et la durée du son liée à largeur de l'objet. La seconde modalité est identique, à la différence du balayage, qui y est effectué latéralement. La troisième modalité présente aussi les objets dans un ordre latéral, mais leur distance est indiquée par le nombre de répétitions du son qui leur correspond, et leur largeur par la fréquence du son. Avec la quatrième modalité, la scène est segmentée en trois régions gauche, centre et droite. Ces régions sont audibles chacune leur tour, de la manière suivante : si un ou plusieurs objets y sont présents, le son attaché à l'objet le plus proche est diffusé, avec la fréquence indiquant la largeur et le niveau sonore dénotant la proximité ; si, au contraire, aucun objet n'est présent, un son spécifique (discret) est diffusé.

## Discussion

La dénomination de ce principe est sujette à discussions. Initialement nommée *substitution sensorielle* par Bach-y-Rita (parfois traduit par *suppléance perceptivo* [Auvray, 2006]), la technique a aussi été appelée *supplémentation* par d'autres chercheurs par la suite [Lenay et al., 2003]. Ce changement d'appellation dénote que, plutôt que le remplacement de la vision, le phénomène qui se produit est l'émergence d'un nouveau sens, différent des autres. Dans cette thèse, nous utiliserons les deux termes indifféremment.

Une autre difficulté avec la supplémentation sensorielle est de définir le périmètre de ce qu'elle englobe. Il est plus large que ce qu'on pourrait penser dans un premier temps : par exemple, Bach-y-Rita y inclut la lecture du Braille [Bach-y-Rita and Kercel, 2003]. À partir de là, peut-on dire que toute information transitant par un sens et communiquée par le biais d'un autre sens constitue de la supplémentation sensorielle ? En considérant que oui, alors la description orale d'une scène représente une forme de substitution sensorielle. Mais cette définition de la supplémentation sensorielle est trop large à notre convenance, elle dilue l'intérêt que l'on y trouve en faisant de presque toute interaction possible d'une personne aveugle avec son environnement un acte de substitution sensorielle. Peut-on la restreindre en limitant le principe de substitution aux interfaces présentant une structure en grille matrice, analogue à une image filmée ? Cela exclurait l'expérimentation de Declerck et Lenay, dans laquelle ces auteurs ont développé un appareil « minimal » permettant de percevoir des images [Lenay and Declerck, 2018]. Le système est qualifié de minimal car son capteur se réduit à une unique diode photosensible et son interface avec l'utilisateur est un stimulateur vibratoire unique placé au bout d'un doigt. Cet appareil suffit à permettre à ses utilisateurs de percevoir des images. C'est-à-dire qu'ils ont l'impression de percevoir l'image directement et implicitement ; autrement dit, un phénomène *d'attribution distale* est présent. L'attribution distale est en fait ce qui permet peut-être de caractériser le mieux la supplémentation sensorielle. Elle correspond aux situations où, bien que la personne perçoive par le biais d'un outil (le système de supplémentation sensorielle), elle ressent directement ce qu'elle perçoit, comme si elle le percevait naturellement ; les sensations liées à l'outil lui-même disparaissent. Auvray *et al.* discernent trois éléments formant l'attribution distale : la conscience d'un lien entre mouvement et perception, la perception d'un objet et la perception spatiale. La perception de l'objet et la perception spatiale sont fortement liées à la possibilité de manipuler et d'agir sur ce qui est perçu [Auvray *et al.*, 2005]. L'importance du mouvement dans l'émergence de l'attribution distale a également été mise en évidence par le système minimal de Lenay et Declerck [Lenay and Declerck, 2018], et plus généralement, par la plupart des expériences de supplémentation sensorielle. Cette dépendance aux mouvements et aux actions n'est pas l'apanage de la supplémentation sensorielle, elle est aussi une composante importante des autres formes de perception, et en particulier de la vision [Schroeder *et al.*, 2010].

Les projets de substitution sensorielle, qu'elle soit visuo-tactile ou visuo-auditive, n'ont jamais débouché sur des systèmes suffisamment robustes pour être utilisés dans la vie courante [Loomis *et al.*, 2012]. Le rapport entre les bénéfices des appareils de supplémentation sensorielle et les contraintes qu'ils posent n'est donc pas suffisamment avantageux. Les progrès technologiques ont permis une miniaturisation importante des caméras et des processeurs, et les contraintes sont donc avant tout dues à l'interface homme-machine. Dans le cas des systèmes visuo-tactiles, se pose l'opposition entre la taille de l'interface et son débit. Le problème est induit par l'utilisation qu'on souhaite faire des systèmes de supplémentation sensorielle. Dans notre cas, c'est l'assistance à la mobilité. La vision, qui est le sens qu'on cherche à remplacer (au sens des fonctions qu'il fournit pour la mobilité), est le sens possédant le débit informationnel le plus élevé [Loomis *et al.*, 2012]. Pour le

remplacer, même dans un mode très dégradé, il est alors nécessaire que le système ait un débit élevé. C'est en particulier le cas de l'interface homme-machine. Ce débit élevé, dans le cadre de systèmes électro-tactiles, nécessite mécaniquement plus de surface puisque la peau a une densité limitée de récepteurs sensoriels (et que le codage par fréquence est également limité). Or, plus la surface est grande, plus les contraintes pesant sur l'utilisation sont lourdes : poids, confort, fragilité, autonomie, critères esthétiques, coût... Chacune étant individuellement de nature à limiter l'adoption de ces systèmes. Pour les systèmes audio, la bande passante représente également un obstacle à la performance. Les systèmes visuo-tactiles posent également des problèmes d'irritation de la peau voire de douleurs [Auvray et al., 2007] citant [Bach-y Rita, 1972].

Le développement de systèmes extrêmement peu contraignants, très limités en résolution mais capables de fournir des informations utiles aux personnes représente une piste intéressante. Un tel système aurait probablement une fonction limitée, spécialisée, plutôt que le remplacement intégral de la vision [Loomis et al., 2012, Elli et al., 2014]. Malgré les nombreux succès obtenus dans les expérimentations en laboratoire, une supplémentation sensorielle intégrale de la vision semble difficile à cause de la complexité de l'environnement, qui augmente la charge cognitive [Elli et al., 2014].

Pour ces raisons, malgré les débuts prometteurs qui laissaient imaginer un développement rapide et une adoption massive, les appareils de substitution sensorielle ne sont toujours pas utilisés dans la vie quotidienne des personnes déficientes visuelles.

### 2.2.3 Implants rétiniens

Les prothèses visuelles invasives, ou rétines artificielles, constituent une forme d'aide à la perception qui peut être rapprochée de la substitution sensorielle. On pourrait peut-être même les appeler elles-mêmes des appareils de substitution sensorielles, mais le terme de substitution sensorielle est encore discutable ici – cela dépend de ce qu'on entend par « sens ». Les implants sont constitués d'une partie pour l'acquisition d'images et d'une interface pour la transmission au nerf optique. L'interface de transmission est nécessairement implantée à l'intérieur de l'œil, où elle peut être posée sur la rétine, en dessous, dans la choroïde, ou la sclère : on parlera alors, respectivement d'implant épi-rétinien, sous-rétinien, supra-choroïdal ou épi-scléral (figure 2.9). Cette interface est constituée par une matrice d'électrodes capables de stimuler les cellules nerveuses avec lesquelles elles sont en contact, répliquant ainsi le fonctionnement de la rétine. L'acquisition d'image est généralement réalisée par une caméra externe, mais certains implants embarquent une matrice de photodiodes qui acquiert les images directement à l'intérieur de l'œil et transmet les données aux électrodes [?]. Il est à noter qu'il existe des travaux sur des implants s'interfaçant avec le cortex visuel plutôt que la rétine ou le nerf optique [?].

Comparées aux solutions non-invasives, les méthodes invasives ont leurs avantages et leurs défauts. L'avantage évident et principal est que, si une prothèse est bien exécutée et fonctionnelle, alors la personne retrouve l'usage de l'organe perdu, ici la vision. Cela peut donc paraître être la solution presque universelle à tous les

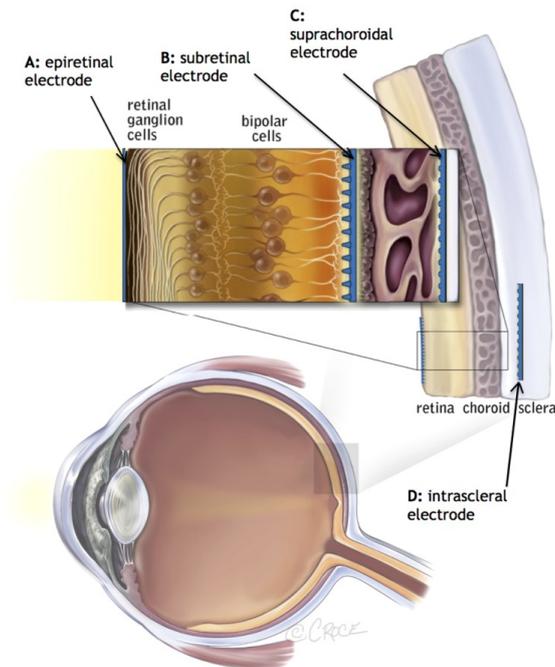


FIGURE 2.9 – Emplacements potentiels des implants rétiniens. A : implant épi-rétinal ; B : implant sous-rétinien ; C : implant supra-choroïdal ; D : implant épi-scléral. Source [?].

problèmes de cécité. Ce n'est pas le cas, en tout cas ni aujourd'hui, ni dans un futur proche. En effet, les implants rétiniens souffrent de nombreux problèmes qui les empêchent de représenter la solution idéale qu'ils peuvent paraître être. L'opération chirurgicale constitue un premier obstacle potentiel. Comme toute opération, elle comporte des risques, et ajoute son coût à celui de l'implant lui-même, et les infrastructures hospitalières doivent être adaptées. La question du retrait est également cruciale, car, s'il est impossible à effectuer, nombreux sont les problèmes induits : que faire en cas de défaillance ? Comment profiter de développements ultérieurs de la technologie si on est bloqué avec une ancienne prothèse ? La pose et le retrait de prothèses représentent un pan délicat de la médecine, les rejets étant possibles et le vieillissement pas toujours bien anticipé. Ainsi, des problèmes de vieillissement sont observés sur les dispositifs existants [Chuang et al., 2014]. De plus, les implants rétiniens sont *a priori* inadaptés aux personnes déficientes visuelles non totalement aveugles. La question du risque se pose pour eux de manière encore accentuée, puisqu'il est possible qu'ils perdent le peu de vision qu'ils possèdent.

Les implants rétiniens sont potentiellement inefficaces ou en partie inefficaces dans le cas de personnes aveugles de naissance, qui n'ont jamais développé de faculté de vision [Elli et al., 2014]. Même dans des conditions favorables, avec des risques et contraintes limités, il est probable que de nombreuses personnes choisiraient de ne pas subir d'implantation, optant plutôt pour des solutions non-invasives. Toutes ces raisons sont valables dans l'hypothèse où les implants sont parfaitement opérationnels et très performants. Or, c'est loin d'être aujourd'hui le cas. De nombreux problèmes techniques demeurent, au premier rang desquels on trouve l'interface avec le nerf optique, qui est assez « rustique » : elle est composée d'un faible nombre

d'électrodes, moins de 100 x 100, en général moins de 10 x 10 [Chuang et al., 2014]. Se pose donc le même problème fondamental de débit d'informations que dans le cas des systèmes de supplémentation sensorielle. À ce problème il faut ajouter les complications liées à l'implantation du système : source d'énergie et son positionnement, rechargement, etc. Au final, les implants rétiniens sont encore loin d'être une solution universelle.

## 2.2.4 Aides à la navigation

À partir des années 1990, les progrès dans le domaine des systèmes d'information géographiques (SIG) et la mise en place et l'ouverture de systèmes de positionnement par satellites (en anglais *Global Navigation Satellite Systems* ou GNSS), en commençant par le système GPS, puis d'autres (par exemple GLONASS, Beidou, maintenant Galileo) ont permis le développement d'aides à la mobilité orientés vers la navigation et la localisation. Ces systèmes permettent à un récepteur de connaître sa position absolue dans la zone de couverture des satellites avec un certain niveau de précision.

Un peu plus tard, un autre progrès est venu faciliter la conception de systèmes d'aide à la navigation : la miniaturisation des capteurs inertiels grâce aux technologies *microelectromechanical systems* (MEMS) ou microsystèmes électromécaniques. La miniaturisation a permis de passer de capteurs très coûteux, encombrants, lourds et grands consommateurs d'énergie à des puces extrêmement bon marché et consommant très peu. Il est très avantageux de combiner capteurs inertiels et systèmes de navigation par satellite car les défauts des premiers sont à peu de choses près les qualités des seconds, et vice versa. Ainsi, il est possible, en fusionnant les données qu'ils produisent, d'obtenir des données bien plus qualitatives (précision dans le temps) qu'avec une ou l'autre des technologies prises séparément.

### Appareils indépendants

Le premier système basé sur le GPS semble être celui de Loomis *et al.* [Loomis et al., 1994] (d'après la synthèse de Giudice et Legge [Giudice and Legge, 2008]). Il possédait plusieurs interfaces, dont une qui deviendra par la suite très classique, une synthèse vocale, et une autre plus originale, utilisant des sons spatialisés pour indiquer directement à la personne où se trouvent les objets (direction, distance). Le système complet était assez encombrant (figure 2.10).

Le *Geotact* est un outil d'aide à la navigation destiné à guider l'utilisateur avec des instructions qui indiquent la direction (en heures) et la distance (en mètres) du prochain point à atteindre [Farcy et al., 2006]. La position est estimée avec un récepteur GPS et des capteurs inertiels.

Le *Trekker Breeze* est un système commercial de navigation de l'entreprise HumanWare<sup>5</sup>. Il a été étudié par une équipe de recherche indépendante [Vincent et al.,

---

5. Il n'est plus disponible à la vente, mais du support est encore fourni <http://support>.



FIGURE 2.10 – Le *Personal Guidance System* de Loomis *et al.* (photo prise sur le site du système : <http://www.geog.ucsb.edu/pgs/main.htm> (accès le 13 septembre 2019)).

2014]. C'est un GPS parlant qui annonce le nom des rues, des intersections et des points d'intérêt sur le trajet. Il coûtait au moment de l'étude – publiée en 2014 – environ 1000 dollars canadiens. Le participant qui l'avait utilisé dans l'étude en était satisfait, appréciant particulièrement la possibilité d'enregistrer des lieux – utile en phase de familiarisation – la possibilité de savoir où il se trouvait à tout moment et la réduction du niveau de concentration nécessaire lors du parcours de trajets connus. Le participant notait cependant des erreurs sur les noms des rues, des problèmes liés à la précision limitée du GPS et à l'absence de couverture satellite dans certains lieux. Au final, il ne considérait pas le Breeze comme suffisant pour se déplacer seul dans un lieu inconnu ; il était plutôt considéré comme une assistance à la familiarisation, limitant la charge cognitive dans les lieux connus. Cet avis était partagé par les instructeurs de mobilité impliqués dans l'étude.

Des systèmes purement inertiels ont également été utilisés pour localiser les personnes. Le problème qui se pose alors est celui de la dérive de la trajectoire, et un mécanisme spécifique doit être adopté pour mitiger celle-ci. Zheng *et al.* ont utilisé à cette fin des capteurs inertiels fixés sur les chaussures des personnes [Zheng *et al.*, 2014]. Cette position particulière apporte la certitude que la vitesse et l'accélération appliqués aux capteurs sont toutes deux nulles à intervalle régulier, quand le pied touche par terre. Le système peut alors réinitialiser la double intégration de l'accélération vers la vitesse et de la vitesse vers le déplacement : cette technique est nommée *Zero-velocity Updates* ou ZUPT. C'est cette double intégration qui engendre la dérive de position, et la réinitialiser ainsi permet de limiter la dérive. Zheng *et al.* ont également utilisé un filtre de Kalman étendu pour estimer et corriger les paramètres d'erreur des capteurs. Le filtre utilise la phase statique du pied pour analyser les données du capteur et en déduire les paramètres d'erreur. Ces deux techniques permettent d'utiliser la centrale inertielle seule avec une erreur



FIGURE 2.11 – *Seeing Eye GPS* est une application mobile de navigation « virage par virage ». Image tirée de <https://www.senderogroup.com/products/shopseeingeyegps.html> (accès le 13 septembre 2019).

limitée à 1 % de la distance parcourue. Il est à noter qu’une erreur de 1 % de la distance parcourue constitue assez rapidement une dérive gênante.

### Applications pour smartphones

Les smartphones sont des appareils extrêmement intéressants pour les aides à la mobilité, particulièrement en ce qui concerne l’aide à la navigation. Plusieurs applications fournissent de l’assistance en utilisant les fonctions de localisation des smartphones, leur connexion à Internet et leur capacité à gérer des systèmes d’information géographiques.

L’application *Seeing Eye GPS*, disponible uniquement sous iOS, en fait partie (figure 2.11). Elle est centrée sur la navigation « virage par virage » (*turn by turn* en anglais), où la personne reçoit des instructions lorsqu’un changement de direction est nécessaire [Sendero Group, 2019]. Elle offre une interface tactile simplifiée, avec trois boutons accessibles en permanence : trajet, points d’intérêt (dont la source est sélectionnable entre les données d’Apple, de Foursquare et d’OpenStreetMap), et localisation actuelle. Les intersections où se trouve la personne ainsi que les points d’intérêt environnants sont décrits automatiquement par le système. Cette application n’est pas disponible en France ; elle est gratuite à l’achat mais nécessite à l’usage de payer un abonnement de 5,99 \$ par mois.

Une autre de ces applications est *Blindsquare*. Coûtant une quarantaine d’euros (précisément 43,99 euros au moment où ces lignes sont écrites), fonctionnant sous iOS, elle permet de décrire l’environnement, annonce les points d’intérêt et prévient les utilisateurs à l’approche des intersections via une interface fonctionnant avec de la synthèse vocale [MIPSoft, 2019]. Les données géographiques proviennent d’OpenStreetMap et de FourSquare, qui est elle-même une application référençant les points d’intérêt (restaurants, boutiques, lieux publics, etc.) avec des données annexes (horaires d’ouverture, par exemple). Un algorithme filtre les données sur

les lieux environnants pour ne présenter que les informations les plus pertinentes à l'utilisateur. L'interface d'entrée de BlindSquare, généralement basée sur une interface tactile accessible, utilise à son avantage les possibilités des smartphones : les utilisateurs peuvent par exemple secouer le smartphone pour demander à entendre leur localisation courante. Les utilisateurs peuvent enregistrer des lieux particuliers, définir des préférences de communication (pour, par exemple, n'entendre que les informations relatives aux rues et points d'intérêts personnels). Pour la navigation, elle se connecte à une application de navigation généraliste comme Google Maps.

## Discussion

Les smartphones disposent de plusieurs sources de données leur permettant d'estimer leur position : ils intègrent un récepteur GNSS ; ils communiquent avec des antennes radio de réseau mobile dont la position est connue ; ils peuvent détecter les points d'accès IEEE 802.11 WiFi, dont la position est souvent connue ; enfin, ils intègrent presque toujours des capteurs inertiels – accéléromètres, gyromètres, boussole [Goo, 2019]. Les performances du GPS sont améliorées grâce au système AGPS, qui transmet les données d'orbite des satellites à partir du réseau mobile plutôt qu'à partir des satellites eux-mêmes, processus beaucoup plus lent. Il est plus ou moins facile d'utiliser ces sources, selon que le système assure lui-même la fusion ou que les développeurs doivent s'en occuper.

Les applications smartphones présentent plusieurs avantages par rapport aux appareils spécifiques. Elles coûtent peu, surtout si on considère le smartphone comme déjà acquis – et même s'il est inclus, il est possible de se procurer un smartphone performant pour un prix limité, conséquence de leur production en très grandes quantités. Un avantage extrêmement significatif est la discrétion qu'une application smartphone apporte à la personne qui l'utilise. La vue de quelqu'un tenant un smartphone est commune et ne focalisera pas l'attention des passants, contrairement à celle d'un dispositif spécifique. Or, une bonne partie de la population des personnes déficientes visuelles cherche à passer inaperçu en public [Hersh, 2018]. Dans la même lignée, l'esthétique des dispositifs est considérée comme un critère très important pour beaucoup d'utilisateurs [Golledge et al., 2004].

### 2.2.5 Aides basées sur la vision

Les systèmes de vision par ordinateur ont connu depuis les années 1980 un progrès continu à toutes les étapes de la chaîne de traitements. Les capteurs d'acquisition d'images deviennent de plus en plus performants en sensibilité, résolution, et dynamique, dans des tailles qui diminuent. De nouveaux types de capteurs d'images – au sens large – émergent, comme les caméras temps de vol (*Time of Flight* en anglais) ou les LIDAR. Le rapport puissance sur taille et consommation des puces qui effectuent les calculs augmente, permettant une augmentation des capacités de calcul pouvant être embarquées à niveau égal d'autonomie. Enfin, les algorithmes traitant les images, en particulier ceux permettant d'extraire des informations de haut niveau des images se sont améliorés de manière spectaculaire au cours des

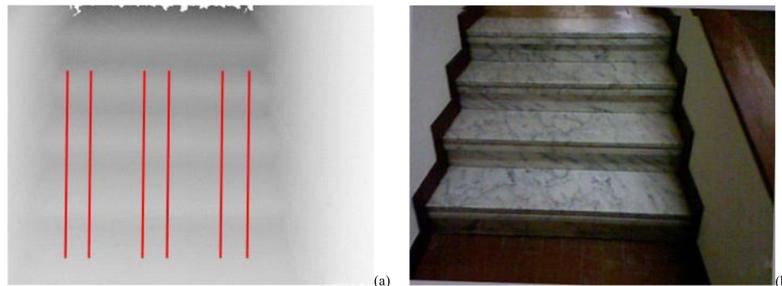


FIGURE 2.12 – Le dispositif de Filipe *et al.* analyse le profil de profondeur sur des lignes pour identifier les obstacles et escaliers.

années, comme en témoignent les performances des méthodes de *Deep Learning* appliquées à la reconnaissance d’objets et de scènes. Ces progrès ont permis le développement d’aides électroniques portables utilisant la vision par ordinateur.

Filipe *et al.* ont développé un système basé sur un capteur Kinect<sup>6</sup> qui détecte les obstacles et les escaliers [Filipe *et al.*, 2012]. La classification est effectuée sur des zones prédéterminées de l’image, où le profil de la profondeur est analysé (figure 2.12). Ces profils sont ensuite classifiés par un réseau de neurones ; la précision annoncée est de 99%.

Tapu *et al.* ont utilisé un smartphone et sa caméra pour développer une application qui détecte en temps réel les objets en mouvement [Tapu *et al.*, 2013]. Par l’analyse des images de la caméra du smartphone, l’application estime les trajectoires des objets présents dans l’image, détermine si elles sont menaçantes ou non. Cette analyse nécessite plusieurs phases : détermination des mouvements du smartphone lui-même, identification des obstacles, etc. qui sont représentés sur la figure 2.13. Les obstacles sont également classifiés (piéton, véhicule, vélo ou objet inanimé) par le système. Des tests de précision ont été effectués, et le nombre de faux négatifs, c’est-à-dire les obstacles qui n’ont pas été détectés et qui sont ceux qui nous intéressent le plus, est de 3,6 % pour les voitures, 31,3 % pour les vélos, 9,3 % pour les piétons et 21,2 % pour les objets inanimés. L’application nécessitait en moyenne 102 ms pour traiter une image sur un smartphone Samsung Galaxy S3 et 18 ms sur un ordinateur équipé d’un processeur Xeon à 3,6 GHz, 16 Go de mémoire vive et une carte graphique NVIDIA Quadro 4000.

La vision par ordinateur a été utilisée pour remplir des tâches plus spécifiques et moins directement liées à la mobilité, qui sont parfois inaccessibles par un autre sens que la vision. Dans cet esprit, une équipe de recherche du City College de New York s’est attelée à développer des applications basées sur la vision par ordinateur afin d’assister les personnes déficientes visuelles dans leur vie de tous les jours.

Tian *et al.* ont développé une application qui permet de reconnaître les portes à partir de connaissances déterminées *a priori* : les portes ont quatre coins et quatre lignes dont deux parallèles, et sont encastrées à l’intérieur des murs, contrairement aux meubles qui, eux, dépassent des murs [Tian *et al.*, 2010]. Comme il est pro-

6. Une combinaison de caméra RGB et de capteur 3D largement diffusé commercialement et disponible pour un coût raisonnable.

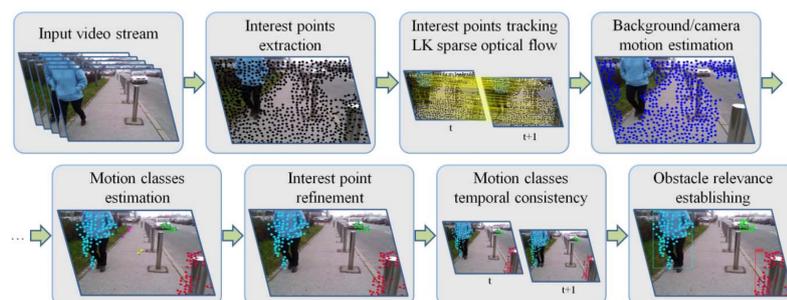


FIGURE 2.13 – Étapes d’analyse des images pour les obstacles de Tapu *et al.* [Tapu *et al.*, 2013].

bable que la personne tenant la caméra ne parvienne pas à cadrer la porte dans l’image – puisqu’elle ne la voit pas, en principe – l’équipe de conception a optimisé l’application pour la reconnaissance de portes partiellement visibles.

Yi *et al.* ont proposé d’utiliser une solution basée sur la vision par ordinateur pour la recherche d’objets [Yi *et al.*, 2013]. Le système utilise à la fois une caméra mobile tenue par la personne et un réseau de caméras fixes distribuées dans la maison. Quand la personne déclare être à la recherche d’un objet spécifique, les images des caméras fixes sont analysées pour y chercher la présence de l’objet, dont l’apparence a été précédemment apprise. Une fois trouvé, sa localisation est annoncée (par exemple « dans la cuisine »). Une fois dans la cuisine, la personne utilise la caméra qu’elle tient afin de trouver la position exacte de l’objet recherché.

La lecture de texte à distance est inaccessible à la plupart des personnes déficientes visuelles sans assistance. Plusieurs appareils ont donc été développés pour les aider dans cette tâche. Yi et Tian ont développé une méthode de reconnaissance de texte [Yi and Tian, 2015]. Elle fonctionne en deux phases : d’abord, le système se base sur les couleurs constantes et l’alignement horizontal que les textes présentent souvent pour rechercher les zones de textes dans les images. La seconde phase est la reconnaissance des caractères. Elle est implémentée sous la forme d’une application fonctionnant sur smartphone et présente des performances supérieures aux autres solutions comparables de l’époque.

Dans le cadre de la mobilité, la reconnaissance de panneaux indicateurs représente une problématique proche de la lecture de textes distants : elle n’est accessible que par la vision, en tout cas pour les panneaux classiques, et est très utile pour la navigation, particulièrement dans les lieux complexes comme les gares, aéroports, etc. Wang et Tian, de l’équipe du City College de New York, ont développé une application pour résoudre ce problème [Wang and Tian, 2011]. Les panneaux connus (toilettes, sorties de secours, flèches, etc.) sont recherchés dans les images ; la méthode utilisée permet de trouver plusieurs occurrences d’un même signe dans une scène. Les panneaux seuls sont correctement identifiés dans 85 % à 90 % des cas, tandis que les panneaux multiples le sont dans 70 % à 90 % des cas. Quatre panneaux différents ont été testés dans le premier cas, et six dans un second ; chacun de ces dix tests a été réalisé sur trente à quarante échantillons.

D’autres applications de reconnaissance spécifiques ont été développées dans la

même équipe : on peut citer, par exemple, un système de reconnaissance de billets de banque [Hasanuzzaman et al., 2012], ou un autre pour les vêtements [Tian and Yuan, 2010].

En raison des dangers et de la gêne que représentent les passages piétons, plusieurs systèmes ont été développés pour aider les personnes déficientes visuelles à les franchir. Une équipe du Smith-Kettlewell Eye Research Institute de San Francisco est particulièrement active sur ce thème [Ivanchenko et al., 2008, 2009, Shen and Coughlan, 2012, Coughlan and Shen, 2012, 2013].

Un de leurs systèmes vise à assister à l'orientation pour la traversée de passages cloutés – au sens premier du mot orientation, c'est-à-dire à aider la personne à s'aligner avec le passage [Ivanchenko et al., 2008]. L'algorithme fonctionne sur un Nokia N95, qui est un des premiers smartphones, dont la puissance était limitée par rapport aux appareils récents (332 MHz, 81 Mo de mémoire vive disponible). Les images sont analysées à la fréquence de 3 Hz par le smartphone pour y rechercher la présence d'un passage piéton et une tonalité est émise quand l'application en a détecté un. Les utilisateurs peuvent ainsi balayer leur environnement pour, d'une part, rechercher un passage piéton, et, d'autre part, s'aligner avec lui une fois qu'il est trouvé.

Autre système ou évolution plus tardive du précédent, le système *Crosswatch* vise à offrir une assistance bien plus complète que le simple maintien de l'orientation lors d'une traversée [Coughlan and Shen, 2013]. Cette dernière est incluse, pour deux types de marquage au sol qu'on rencontre aux USA. Une description des intersections est également fournie à l'utilisateur : type d'intersection, présence ou non de passages protégés, de feux tricolores ou de stops, de signaux commandés par un bouton ou non... *Crosswatch* assiste également dans la localisation de l'entrée des passages protégés, ce qui peut être non trivial surtout dans le cas des intersections complexes qui sont justement celles qui posent le plus de problèmes. Enfin, le système indique l'état des feux et les instants les plus propices à la traversée. Le système fonctionne sur un smartphone ; la caméra est employée, mais ne suffit pas à l'acquisition de toutes les informations. Le récepteur GPS et la boussole du smartphone sont utilisées à la fois pour le guidage jusqu'à l'entrée du passage et, en association avec des sources de données externes, pour la description des intersections. Les sources de données standards existantes, comme OpenStreetMap ou Google Maps, ne sont pas suffisantes à elles seules pour assurer les fonctionnalités du système : la présence de boutons de déclenchement de feux n'y est par exemple pas (encore) signalée, et elles ne suivent les changements qui se produisent qu'avec un délai aléatoire. Afin d'obtenir des informations à la fois plus poussées et plus fiables, les concepteurs de *Crosswatch* souhaitent mettre en place un système de données issues du *crowdsourcing* : chacun pourrait ainsi contribuer à décrire les intersections par le biais d'une plate-forme web.

D'autres systèmes de détection de passages piétons ont été développés, tel que celui de Ahmetovic *et al.* [Ahmetovic et al., 2014]. Le système de Ghilardi *et al.* est, lui, orienté vers un but voisin mais similaire, la détection de dalles et surfaces podotactiles [Ghilardi et al., 2016].

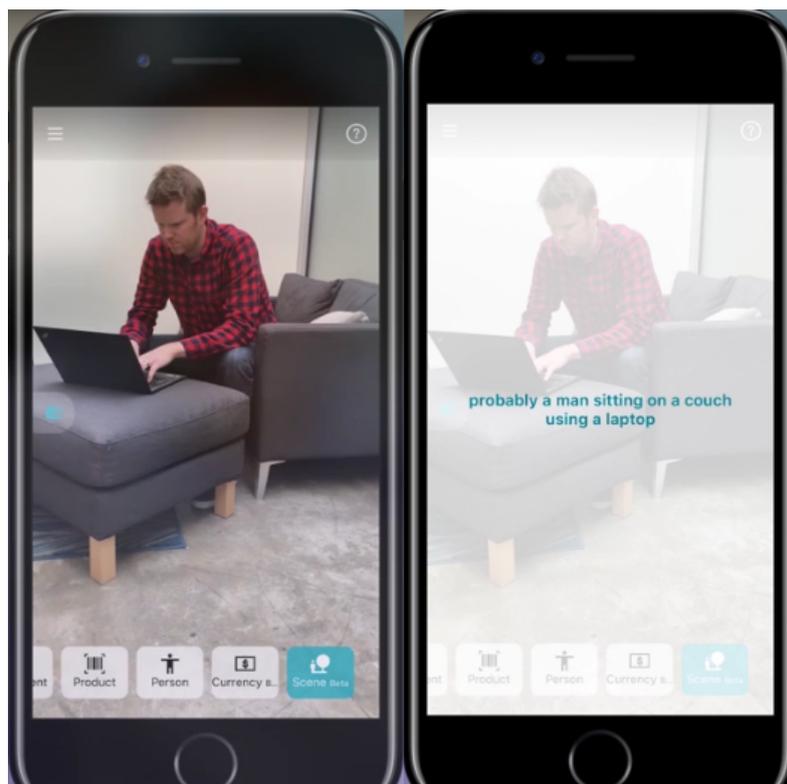


FIGURE 2.14 – L’application de reconnaissance Microsoft Seeing permet notamment de décrire des scènes (images prises sur le site de l’application <https://www.microsoft.com/en-us/ai/seeing-ai>).

Un autre champ s’est développé récemment : celui de la description de scène, dont un bon exemple est l’application *Seeing AI* de Microsoft. Cette application smartphone – disponible sur iOS et gratuite – est dotée de plusieurs fonctionnalités autour de la vision [Microsoft, 2017]. Elle est capable d’analyser et de décrire une scène (figure 2.14) ; elle est également capable de lire des textes courts ou longs, y compris manuscrits et elle peut reconnaître et décrire un grand nombre d’entités : les personnes et leurs émotions, les couleurs qu’on lui montre, les billets de banque, les produits grâce à leur code barre.

## Discussion

Les aides basées sur les technologies de vision par ordinateur sont séduisantes parce qu’elles résolvent le problème en utilisant la modalité sensorielle qui fait défaut, la vision. Elles sont en réalité difficiles à mettre en œuvre dans des appareils destinés à être employés dans la vie de tous les jours. L’extraction d’information d’images, bien qu’elle ait énormément progressé, reste une tâche très difficile, surtout dans le cas où l’environnement n’est pas contrôlé et connu *a priori*.

Les systèmes basés sur la vision par ordinateur ne semblent pas pour l’instant capables de fonctionner correctement dans toutes les conditions rencontrées dans un environnement non-structuré et non-contrôlé. Une luminosité très élevée, la nuit,

les réflexions, les objets partiellement occultés pénalisent les caméras et l'analyse des images ; dans notre cas, s'ajoute à ces problèmes la difficulté de correctement orienter la caméra vers une cible pour une personne déficiente visuelle. Même les systèmes conçus pour la détection d'un seul type d'objet dans les images ne sont pas parfaitement fiables. De plus, bien qu'il soit actuellement réalisable, le traitement d'images en temps réel nécessite beaucoup de ressources calculatoires (processeur, mémoire) et épuise rapidement les batteries qui alimentent les objets mobiles.

La différence fondamentale avec un système de supplémentation sensorielle visuo-tactile, qui repose aussi sur la vision, est l'usage qu'en fait la personne et la place de la décision dans le système {appareil+personne}. Dans le cas de la supplémentation sensorielle, même si le dispositif acquiert et effectue des traitements sur des images, il ne porte pas de réel jugement sur leur contenu, c'est la personne l'utilisant qui décide en fonction de ce qu'elle « sent » à travers l'outil. À l'inverse, dans le cas d'une aide basée sur la vision, c'est l'appareil qui juge la situation, prend une décision, et en informe l'utilisateur. La distinction n'est évidemment pas binaire ; les systèmes de supplémentation sensorielle tendent à prendre une voie où plus d'intelligence est embarquée dans le système car c'est un moyen efficace de réduire la charge cognitive. On peut voir un exemple de cette tendance dans le système Sound of Vision, qui identifie et classe les objets se trouvant dans la scène avant de les présenter à l'utilisateur.

Cette différence vient à jouer un rôle important pour au moins une raison pratique : elle détermine à la charge de qui se trouve la prise de risque. On a vu que le système de Filipe *et al.* est donné pour avoir une précision de 99% dans ses estimations. Un détail intéressant est que lors des tests, la présence d'escaliers descendants a été reportée comme étant une absence d'obstacle dans environ 1,5 % des cas (2 cas sur 130). Cela peut paraître peu, mais si on prend cela comme indicateur d'un risque de chute dans les escaliers tous les cents escaliers, la perspective est différente. Le problème est que la personne risque de faire confiance au système dans des cas où celui-ci se trompe.

### 2.2.6 Aides fonctionnellement basées sur des objets dispersés dans l'environnement

Ces systèmes sont ceux dans lesquels l'information provient d'objets dispersés dans l'environnement. Le périmètre de ce que nous définirons ici comme « objet » est assez large : un objet peut être dit *intelligent*, c'est-à-dire doté de capacités de traitement de l'information ou de décision propres, mais il peut également être un simple marqueur ou un conteneur d'informations – les tags RFID sont donc inclus. Les objets peuvent même être dépourvus d'électroniques, comme les QR-codes. Les contraintes sont au nombre de trois :

1. Les objets sont dispersés dans l'environnement ;
2. Ils sont en eux-mêmes sources d'informations ou d'interactions pour la personne ;

3. Présence d'au moins un appareil électronique dans les interactions – les bandes podotactiles, par exemple, en sont donc exclues.

Traverser des routes ou des carrefours représente un danger potentiel pour tout piéton. Il est bien plus difficile pour une personne malvoyante ou aveugle de limiter ce danger. Les feux de signalisation peuvent être équipés de signaux accessibles, indiquant quand les piétons peuvent passer. Selon les pays, la modalité de déclenchement de ce signal varie. En France, une partie des feux de signalisation est équipée de boîtiers permettant à une personne possédant une télécommande spécifique de déclencher le signal sonore. L'installation de ces équipements sur les indicateurs pour piétons est obligatoire dans les nouvelles installations de voirie et lors de travaux sur voirie existante<sup>7</sup> ; à terme, tous les feux de signalisation en seront équipés (pour l'instant, environ 300 000 le sont [Alexandre et al., 2018]). Ce système est également déployable dans d'autres lieux recevant du public, comme les gares ou les musées. Environ 40 000 établissements recevant du public en sont équipés [Alexandre et al., 2018]. Le fonctionnement de ce système est décrit dans la norme NF S32-002 : le signal émis autour de la fréquence de 868,3 MHz (bande ISM) par la télécommande déclenche la lecture d'un message pré-enregistré par le haut-parleur des balises, avec une portée de 5 à 20 m. Le système est bien reçu par les utilisateurs, qui sont – d'après une enquête de l'Institut de la Vision commandée par l'entreprise qui fabrique les balises – 80 % à déclarer que c'est l'aide la plus utile disponible et 75 % à fréquenter préférentiellement les lieux qui en sont équipés [Alexandre et al., 2018]. Le prix d'achat d'une balise complète est d'environ 500 euros<sup>8</sup>.

Les *Talking Signs* sont des objets permettant de transmettre un message à distance inventés par Loughborough [Loughborough, 1979] (d'après Brabyn, [Brabyn and Brabyn, 1983]). Destinés à être accolés à un panneau indicateur classique, ils transmettent le texte de celui-ci via un émetteur infrarouge. Quand un récepteur portable adapté se trouve pointé en direction de l'émetteur, il décode le message et le présente à l'utilisateur sous forme vocale [Brabyn and Brabyn, 1983]. Brabyn a comparé l'utilisation de ces objets à des étiquettes en braille pour la navigation. Les panneaux étaient situés au dessus des portes et les étiquettes sur les portes ; les deux types supportaient les mêmes informations. Les 21 participants aux tests étaient globalement plus rapides pour accomplir la tâche demandée, qui était d'atteindre un lieu précis, quand ils utilisaient le système Talking Signs.

Un grand nombre de projets basés sur la dispersion d'objets porteurs d'informations dans l'environnement utilise des tags RFID pour l'implémentation des objets. La technologie RFID, bien qu'ancienne (années 1960 [Landt, 2005]), présente des avantages certains : un très faible coût d'acquisition (quelques centimes) et un entretien potentiellement nul avec les tags passifs.

Ross et Blasch ont expérimenté tôt l'aide à la navigation par le biais de tags RFID [Ross and Blasch, 2002]. Dans leur système, les tags sont disposés aux intersections

---

7. Décret n°2006-1657 du 21 décembre 2006, disponible à l'adresse <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006054989> (accès le 31 août 2019).

8. On peut en trouver à la vente à l'adresse <https://okeenea-produit.com/collections/balise-sonore/products/balise-sonore-naviguel-plus-hifi> (accès le 31 août 2019)

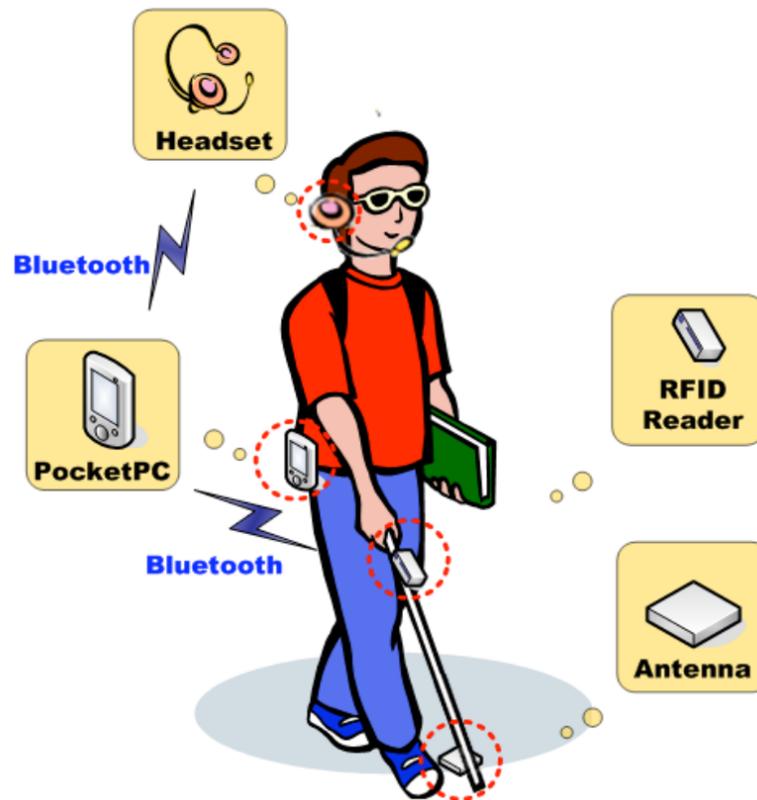


FIGURE 2.15 – La canne *iCane* de Chang *et al.* (image tirée de [Chang *et al.*, 2005])

des couloirs et un ordinateur embarqué les utilise pour connaître la position de la personne. Suivant un trajet calculé, le système est ainsi capable de donner des instructions à chaque intersection : « tournez à gauche », « allez tout droit », ou bien « tournez à droite ». La portée des tags passifs étant faible (environ 25 cm à 40 cm) il n'était pas évident pour les utilisateurs d'atteindre aux intersections une position à laquelle la canne était capable de lire les informations des tags. Pour résoudre cette difficulté, les concepteurs ont intégré un signal sonore indiquant la distance du tag.

Chang *et al.* ont proposé d'utiliser des tags passifs RFID pour enrichir la navigation des personnes malvoyantes. Leur appareil, nommé *iCane*, permet de proposer des informations géographiquement contextuelles à l'utilisateur par l'intermédiaire d'une interface vocale [Chang *et al.*, 2005]. Les tags RFID possèdent chacun un identifiant et sont posés au sol à l'emplacement des points d'intérêt. La personne transporte un ordinateur embarqué (un PDA) sur lequel est installée une base de données qui stocke toutes les informations sur les tags : à partir d'un identifiant, le système est ainsi capable de fournir à la personne les données qui sont liées au lieu (figure 2.15). Cette solution de stockage indirecte a été choisie en raison des difficultés de lecture des tags quand ils contiennent plus d'informations. En effet, la lecture doit être robuste malgré le déplacement de la personne tenant la canne et malgré la portée limitée des tags.



FIGURE 2.16 – Le robot de Kulyukin *et al.* (image tirée de [Kulyukin et al., 2004, 2006]).

Le système de Kassim *et al.* est semblable aux deux précédents : une canne est équipée d'un lecteur RFID et les tags sont déployés sur les dalles podotactiles [Kassim et al., 2013]. Le but est de les utiliser en complément du GPS pour augmenter la précision de celui-ci. De nombreux projets similaires existent. On peut citer par exemple celui de Shiizu *et al.* [Shiizu et al., 2007], celui de Chumkamon *et al.*, qui est particulièrement destinés aux environnements intérieurs où le signal des satellites des GNSS parvient difficilement [Chumkamon et al., 2008], ou encore celui de Willis et Helal [Willis and Helal, 2005]. Celui d'Alghamdi, van Schyndel et Khalil utilise des tags RFID actifs afin d'augmenter la distance de détection, ce qui permet de placer les tags en hauteur, au dessus des portes par exemple [Alghamdi et al., 2014]. D'autres approches que la canne équipée d'un lecteur de tags RFID ont été explorées. Kulyukin *et al.* ont travaillé sur un robot d'assistance guidant les personnes déficientes visuelles en intérieur, le robot lui-même se repérant grâce aux tags RFID présents dans les couloirs [Kulyukin et al., 2004, 2006]. Le robot, tenu en laisse par l'utilisateur, adopte ici un rôle similaire à un chien guide (figure 2.16).

De nombreux travaux ont été menés sur les méthodes de localisation basées sur une infrastructure de points d'accès WiFi. Lymberopoulos *et al.* ont comparé différents systèmes dont une partie importante utilise les points d'accès WiFi; ils sont en général associés à un autre moyen de localisation (infrarouge, capteurs inertiels) [Lymberopoulos et al., 2015].

### 2.2.7 Aides multifonctionnelles

La *GuideCane* est une canne dont l'extrémité en contact avec le sol est dotée de roues et motorisée, ainsi que de capteurs et d'un ordinateur embarqué [Borenstein and Ulrich, 1997]. Les deux roues sont pilotées par l'ordinateur, qui détermine la direction à prendre en tenant compte de la direction souhaitée et des obstacles détectés. Cette détection est réalisée par une rangée de capteurs ultrasoniques montés en arc de cercle sur la canne. La trajectoire et la vitesse instantanée de déplacement de la canne sont estimées par l'ordinateur grâce à une boussole (capteur) et aux données d'odométrie des roues<sup>9</sup>. Les utilisateurs signalent à l'ordinateur la direction qu'ils souhaitent prendre avec un joystick miniature. La *GuideCane* peut également être équipée d'un récepteur GPS pour connaître sa position absolue – avec 20 mètres de précision – et éviter les problèmes de dérive de la trajectoire estimée au cours du temps. L'utilisation de la canne est ainsi en quelque sorte comparable à celle d'un chien guide : la personne suit le dispositif qui évite les obstacles de manière autonome, tout en gardant à moyen terme le cap choisi. La *GuideCane* permet ainsi d'assister à la fois la tâche de la marche et, en partie, celle de l'orientation. Il est rare qu'un dispositif combine l'assistance à ces deux tâches, ou plutôt, cette combinaison est rare dans les dispositifs ne visant pas à fournir une assistance perceptive « totale », comme le font souvent les systèmes de supplémentation sensorielle visuo-tactiles.

La société Handisco propose à la vente la canne *Sherpa* au prix de 750 euros plus un abonnement de 99 euros par an [Handisco, 2019]. La canne propose quatre fonctionnalités :

1. La fonction « Où suis-je » donne la localisation – probablement sous la forme d'une adresse – et les points d'intérêt situés à proximité ;
2. La fonction « Navigation » donne des instructions pour suivre un trajet vers une destination énoncée ;
3. La fonction « Transports en commun » permet de repérer les stations et arrêts, d'obtenir les horaires de passage des prochains véhicules et, une fois à bord, de connaître les prochains arrêts.
4. Enfin, la fonction « Balises sonores » permet d'activer les balises sonores présentes dans l'environnement : celles présentes sur certains feux tricolores pour signaler leur état, celles présentes dans les gares par exemple<sup>10</sup>.

L'interface entre la canne et l'utilisateur est principalement vocale et auditive ; un clavier est également disponible. L'appareil offre 10 heures d'autonomie, et peut être fixé sur une canne traditionnelle ou sur le harnais d'un chien guide. Il est à noter que la canne ne détecte pas les obstacles. La motivation donnée pour justifier ce choix, outre l'augmentation de prix induite, est que cette fonctionnalité

---

9. Mesure de la distance qu'elles parcourent en se basant sur le nombre de tours qu'elles ont effectué et le diamètre des roues, qui est une constante déterminée lors de la conception.

10. Les gares SNCF sont équipées, par exemple, de balises indiquant la position dans la gare et de balises indiquant les directions, cf. <https://www.accessibilite.sncf.com/la-demarche-d-accessibilite/equipements/amenagements-en-gare/se-reperer/article/les-balises-sonores>.

compliquerait grandement l'utilisation de la canne, qui est, en l'état, donnée pour être très facile à utiliser : Handisco annonce une prise en main en 10 minutes.

Le projet européen *Sound of Vision*, décrit plus haut, a permis la réalisation d'un système complexe, capable de d'accomplir plusieurs tâches [Sound of Vision Consortium, 2018]. Cet appareil comprend plusieurs éléments matériels :

1. Une structure portée sur la tête, embarquant des caméras 3D, une centrale inertielle et des haut-parleurs ;
2. Une ceinture haptique comportant 60 moteurs vibrants, portée sur l'abdomen ;
3. Une unité de calcul, embarquée dans un sac à dos et reliée à la tête et à la ceinture par des câbles ;
4. Une télécommande sans fil.

Le système capture des images en 3D, les analyse afin de détecter les objets les plus importants et les signale à l'utilisateur par voie sonore et/ou haptique. La fonctionnalité principale est la détection de scène, pour laquelle deux variations sont disponibles : l'utilisateur peut choisir de ne « voir » que les objets les plus proches situés devant, ou, au contraire, la scène complète. D'autres outils sont également disponibles : détection et signalement de la distance du premier objet situé exactement en face de la caméra (ce mode est nommé « torche ») ; détection et signalement des espaces libres ; détection et lecture de textes ; détection et description de panneaux ; description de scène par une synthèse vocale. Un processus parallèle informe des dangers, obstacles ou trous.

L'entreprise GoSense a développé un « bouclier virtuel », nommé Rango, qui se fixe sur une canne blanche et détecte les obstacles [GoSense SAS] (figure 2.17). Le GoSense Rango utilise (apparemment) trois capteurs ultrasoniques pour détecter les obstacles situés devant la personne, au sol comme suspendus. Il se connecte au smartphone de l'utilisateur et propose également une fonctionnalité de localisation, ainsi que les horaires des transports à proximité des arrêts. L'appareil coûte 900 euros, pèse 107 g, est étanche aux jets d'eau venant de toutes directions (IP64), et possède une autonomie de 3h30.

## 2.3 Discussion

À cause de la complexité de la problématique et de la diversité des systèmes réalisés, la réflexion autour de l'état de l'art est loin d'être évidente. Il est par exemple difficile d'en trouver une classification unique qui soit entièrement satisfaisante. Un ensemble de classifications semble plus à même de montrer les aspects intéressants des différentes approches par rapport à ce qui nous intéresse, c'est-à-dire la problématique de l'amélioration de la mobilité.



FIGURE 2.17 – Le dispositif GoSense Rango (image tirée de [GoSense SAS]).

### 2.3.1 Analyse par rapport aux fonctionnalités

On peut analyser les dispositifs du point de vue de leurs fonctionnalités. Cette manière d'analyser est probablement la plus directe – d'ailleurs, c'est plus ou moins celle qui structure la section précédente – puisque la fonctionnalité fournie par un appareil est sa raison d'être et son aspect le plus important. La progression historique a, naturellement, structuré en partie les développements en familles : dans de nombreux cas, les concepteurs d'appareils nouveaux se sont inspirés d'autres déjà existants, soit parce qu'ils étaient leur continuation directe, soit par la reprise de principes.

Les grandes familles de fonctionnalités sont les suivantes : détection d'obstacles, supplémentation sensorielle, navigation et points d'intérêt, et reconnaissance d'objets. Elles peuvent être subdivisées avec plus de précision. Les appareils de détection d'obstacles peuvent détecter les obstacles eux-mêmes ou les espaces libres ; ils peuvent détecter tous les obstacles ou bien seulement ceux inaccessibles aux cannes blanches, typiquement, en hauteur. Les systèmes de supplémentation sensorielle peuvent être visuo-tactiles (ou visuo-électrotactiles) ou visuo-auditifs ; ils peuvent transmettre une image complète ou une information plus limitée, comme la luminosité. Quant aux systèmes de navigation, ils peuvent donner des informations sur les points d'intérêt, ou seulement fournir un guidage lors de trajets ; la nature des instructions de guidage est également importante, comme l'est leur capacité ou non à fonctionner en intérieur.

### 2.3.2 Technologie des capteurs

Une autre classification intéressante est celle qui se base sur les capteurs employés par les différents appareils. Une autre manière de définir la fonctionnalité d'un appareil est de décrire les informations qu'il transmet à son utilisateur. Dans la mesure où on se limite à des systèmes manipulant des informations, celles qui sont transmises *représentent* la fonctionnalité fournie. Les informations fournies à l'utilisateur sont dépendantes des informations récoltées par le système. La dépendance est plus ou moins directe selon le degré de manipulation et de traitement effectué, mais elle est toujours présente. Les informations récoltées par le système sont fortement dépendantes des capteurs embarqués. On a ainsi un lien plus ou moins direct qui existe entre le ou les capteurs et la fonctionnalité des appareils. Cette approche peut paraître au premier abord purement artificielle, si on considère un capteur comme n'étant qu'un composant technique *simple* fournissant les données qui forment la base d'une fonctionnalité. Un tel raisonnement conduit souvent à considérer les capteurs comme des systèmes abstraits : capteur de distance, capteur d'image, microphone, etc. Cette approche est pertinente mais pour comprendre le problème plus en détail, il est nécessaire de s'intéresser aux particularités des types de capteurs. Ce ne sont en effet pas des composants abstraits mais des instruments avec des défauts et des limites qui contraignent et définissent les capacités d'un appareil. Si chaque modèle de capteur a ses caractéristiques propres, certaines propriétés importantes sont déterminées par le mécanisme du capteur, c'est-à-dire son type. Les caractéristiques les plus intéressantes dans notre contexte sont les défauts des capteurs, car ils déterminent en bonne partie les conditions dans lesquelles les appareils qui les embarquent peuvent être utilisés, et donc dans quelle mesure ces derniers sont susceptibles d'être adoptés. Il est donc nécessaire d'avoir, à côté de la classification, une idée des limites de chacune des familles de capteurs les plus employés.

1. Les limitations des capteurs ultrasoniques ont été décrites par Borenstein et Koren [Borenstein and Koren, 1988] et Farcy *et al.* [Farcy et al., 2006]. Les capteurs fonctionnent en envoyant un signal ultrasonique qui est réfléchi par l'obstacle ; la durée écoulée entre l'émission du signal et la détection du signal réfléchi donne la distance. Si l'angle que fait la surface de l'obstacle avec la direction du capteur est supérieur à un certain seuil (environ 40 °) seule une infime partie de l'énergie du signal émise est renvoyée dans la direction du capteur. Ce problème est limité dans le cas de surfaces présentant des rugosités d'échelle supérieure ou égale à la longueur d'onde du signal ultrasonique ; mais étant donné que cette longueur d'onde est de l'ordre de 5 à 10 mm (la fréquence d'un capteur ultrason classique étant autour de 40 à 50 kHz), de nombreux obstacles apparaissent « lisses » et donc potentiellement invisibles pour le capteur (c'est typiquement le cas des murs).
2. Les capteurs optiques « scalaires » ont généralement un défaut principal : leur sensibilité aux perturbations environnementales. Les autres types de capteurs y sont également sensibles, mais pour les capteurs optiques, une source d'interférence puissante et ubiquitaire existe : la lumière du soleil. De manière générale, hormis dans les environnements strictement contrôlés, la puissance du signal émis par les capteurs optiques est limitée par rapport

à celle de la luminosité ambiante et ne permet qu'un faible rapport signal sur bruit [Kube, 1996]. Plus le capteur est éloigné de l'objet, plus cela est vrai. Ce problème d'intensité lumineuse pure concerne surtout les capteurs de distance infrarouges. Les capteurs lasers sont, en général, moins sensibles à ce problème en raison de leur focalisation importante, permettant une puissance de signal par unité de surface supérieure à la celle de la lumière du soleil [Konolige et al., 2008]. Un autre problème affectant les capteurs infrarouges est l'écart des mesures selon la réflectance de l'objet, variable selon que l'objet a une couleur foncée ou claire, et une surface polie ou non.

3. Les caméras standards sont également affectées par la lumière. Contrairement aux capteurs optiques scalaires, elles sont des récepteurs passifs n'émettant pas de signal. Elles ne sont donc pas affectées par les interférences de la lumière ambiante. En revanche, quand le niveau de luminosité descend sous un certain seuil, qui dépend des caractéristiques (sensibilité/gain du capteur, transmission de l'optique), elles sont inopérantes. Ce problème, bien qu'évident, représente tout de même une limite. Un autre problème lié aux caméras, qui concerne en fait plutôt la chaîne complète d'analyse des images, est la difficulté de la gestion des changements de luminosité :
  - (a) Un passage de l'ombre au soleil ou inversement change énormément la luminosité ;
  - (b) Si le soleil se trouve dans le champ visuel de la caméra, celle-ci a d'importantes chances d'être au moins partiellement « aveuglée » (saturée) ;
  - (c) La dynamique des caméras est limitée, rendant difficile l'extraction simultanée d'informations à la fois dans des zones fortement exposées et faiblement exposées d'une scène ;
  - (d) Les changements de lumière, modifient fortement l'apparence (couleurs, directions des ombres) des objets, par exemple entre le matin, le midi, le soir, et la nuit (éclairage artificiel) ;
  - (e) Les zones d'ombre et de lumière d'un objet ont souvent un contraste plus fort que celui entre une partie de l'objet et l'environnement, ou entre des facettes distinctes de l'objet, rendant la segmentation et la reconnaissance très difficiles en conditions extérieures.

La gestion de la lumière naturelle est globalement complexe. Les deux premiers points concernent la caméra elle-même et sont plus ou moins correctement gérés aujourd'hui, grâce, par exemple, à l'adaptation automatique de la sensibilité ou de l'ouverture du diaphragme – ce dont est naturellement capable l'œil. Les deux derniers, par contre, relèvent de l'analyse qui est faite en fin de chaîne sur les images, et non du capteur d'image à proprement parler. Ils semblent néanmoins pertinents car l'utilisation d'une caméra implique au minimum un traitement des images, et souvent des algorithmes évolués pour reconnaître certains objets présents dans les scènes. Ces opérations sont fortement impactées par les deux derniers points précédents, et un système utilisant une caméra y sera donc exposé. Les caméras sont aussi affectées par les occlusions partielles d'objets.

4. Les caméras de profondeur offrent une information sur la distance de chaque pixel en plus des caractéristiques visuelles classiques (luminosité, couleur).

Trois grandes catégories existent : les caméras stéréographiques, les caméras à lumière structurée et les caméras temps de vol. Les premières sont constituées par une paire de caméras classiques décalées. À la manière d'une paire d'yeux, il est alors possible d'extraire l'information de profondeur des images. Ces caméras souffrent des mêmes problèmes que les caméras classiques puisqu'elles ont les mêmes caractéristiques. Les caméras à lumière structurée sont un peu différentes : un émetteur projette un motif précis ayant une forme bien déterminée (des points par exemple, dans le cas du Kinect [El-laithy et al., 2012]) sur les objets dont la forme doit être analysée. Ce motif épouse les surfaces et est déformé par les volumes de l'objet, et une caméra 2D capture ces déformations. Comme le motif original est connu, un traitement des images permet alors de retrouver la géométrie de l'objet. Puisqu'il est dépendant de la projection d'une image lumineuse, ce type de caméra possède un défaut similaire à celui des capteurs optiques scalaires : la portée diminue en plein soleil. Dans le cas du Kinect, un des capteurs les plus populaires, le phénomène est particulièrement marqué puisque les données de profondeur sont absentes [El-laithy et al., 2012]. La troisième famille de capteurs de profondeur est formée par les capteurs temps de vol. Ces capteurs fonctionnent en émettant un signal 2D et en mesurant, pour chaque pixel, la durée de retour du signal ou son déphasage. Puisqu'ils sont basés sur une émission lumineuse, ces capteurs souffrent eux aussi d'une diminution de portée induite par la lumière du soleil [Butkiewicz, 2014] .

5. Les défauts des récepteurs des systèmes de positionnement par satellites sont bien connus. Le démarrage est long, le temps que le récepteur obtienne les données de trajectoire des satellites (éphémérides). Le fait de disposer d'une mauvaise couverture satellite (en ville par exemple) augmente les risques de d'interruption du téléchargement et la durée du démarrage. Ensuite, la précision du GPS simple est limitée. Dans une étude en conditions réelles, Schipperijn *et al.* rapportent une précision médiane de 3,9 m pour un piéton en ville, voire 5,2 m dans les canyons urbains [Schipperijn et al., 2014]. Cette baisse de précision dans les canyons urbains est liée – mais pas uniquement – à l'impossibilité pour le signal des satellites de traverser la plupart des obstacles ; il faut que les satellites soient en ligne de vue. Ainsi, l'intérieur des bâtiments est généralement coupé des signaux satellitaires. Des mécanismes d'amélioration des systèmes de positionnement ont été mis au point : en particulier, le *GPS assisté* permet d'écourter le temps d'initialisation des récepteurs et le *GPS différentiel* (DGPS) permet d'améliorer de manière spectaculaire la précision du système, grâce à une connexion avec un récepteur terrestre fixe dont la position est connue. Cette contrainte d'un récepteur fixe de position connue est difficilement conciliable avec un système d'aide aux personnes malvoyantes qui se veut, idéalement, ubiquitaire.
6. Les capteurs inertiels sont de deux sortes : accéléromètres et gyromètres. Les accéléromètres servent à mesurer l'accélération linéaire de leur boîtier. Les puces MEMS regroupent trois accéléromètres orthogonaux pour mesurer le vecteur accélération dans l'espace. Les gyromètres, eux, mesurent la vitesse angulaire ou vitesse de rotation de leur support. Comme les accéléromètres, ils sont souvent regroupés par trois pour mesurer les trois

composantes de la vitesse angulaire dans l'espace. Souvent, accéléromètres et gyromètres triaxiaux sont regroupés dans une même puce, formant ainsi une centrale inertielle qui suffit, en théorie, à calculer les déplacements de la puce. Quand les centrales inertielles sont utilisées à des fins d'estimation de la position, la méthode de calcul est basée sur l'intégration de déplacements « infinitésimaux » (en pratiques de petits déplacements) : c'est une forme de navigation à l'estime (*dead-reckoning* en anglais). Ce calcul est remarquable car il n'utilise pas de référence externe, mais il est difficile à réaliser en pratique. En effet, il repose sur l'utilisation conjointe de l'intégration du signal du gyromètre, qui donne les rotations, avec la *double* intégration du signal de l'accéléromètre, qui donne les déplacements linéaires (une simple intégration donnant la vitesse). La double intégration est un véritable problème car elle s'applique également aux erreurs. Ces erreurs sont toujours présentes ; elles peuvent être dues à l'accéléromètre (on y reviendra), mais sont de toute façon inévitables, car les intégrations ne sont que des approximations numériques<sup>11</sup> d'un signal qui n'est lui-même qu'un échantillonnage du mouvement. Une source d'erreur importante est due au fait que les accéléromètres mesurent non seulement leur accélération, due au mouvement, mais également l'accélération de la gravitation. Ces deux accélérations sont confondues dans le signal mesuré par le capteur ; or, il est nécessaire, pour calculer les déplacements, d'éliminer la gravité des mesures. En effet, cette accélération parasite est une constante d'une amplitude souvent supérieure aux accélérations d'un piéton, et elle est intégrée deux fois. Pour la soustraire, il faut connaître à chaque instant la direction du capteur par rapport à la verticale, ce qui est aisé en phase statique puisqu'on sait qu'il n'y a pas de mouvement, mais plus difficile au cours du mouvement : il faut alors utiliser un filtre d'estimation qui exploite les signaux des gyromètres. On risque de voir des résidus du signal dus à la gravité se retrouver dans la double intégration. Les *zero-velocity updates* sont une solution élégante pour limiter les problèmes de dérive des calculs de navigation à l'estime piétonne : on monte les centrales inertielles sur les pieds, pour pouvoir ainsi limiter la durée de l'intégration dans le temps (un pas), et donc la dérive du calcul de la position [Zheng et al., 2014]. Cette dernière n'est cependant jamais complètement éliminée.

7. Les magnétomètres sont des capteurs indiquant la direction du champ magnétique, qui pointe, en l'absence de perturbation, vers le nord magnétique. L'utilisation la plus commune des magnétomètres est celle de la boussole pour connaître le cap de la personne. Des perturbations magnétiques sont cependant présentes dans l'environnement, rendant l'orientation grâce au nord magnétique moins robuste. Ces perturbations sont particulièrement importantes dans les bâtiments, à cause par exemple de leurs équipements électriques et des structures métalliques ; l'effet des perturbations est suffisant pour affecter la navigation, et faire dériver l'estimation d'une trajectoire jusqu'à la rendre inutilisable [Afzal et al., 2010].

---

11. Par exemple, la méthode de Simpson génère moins d'erreur que la méthode des trapèzes, mais elle en génère quand même.

En plus de ces défauts, qui concernent uniquement la validité des informations produites par les capteurs, il est nécessaire de prendre en compte leur consommation d'énergie. Les capteurs optiques projetant un signal sont généralement assez gourmands en énergie, alors que les capteurs ultrasoniques le sont moins. Les capteurs inertiels et les magnétomètres sont également très économes. Les capteurs d'images sont consommateurs avant tout à cause des calculs qu'ils impliquent pour exploiter les données, qui nécessitent des processeurs puissants et gourmands en énergie.

Un des intérêts de l'approche des aides électroniques à la mobilité par l'angle des capteurs qu'elles embarquent est qu'elle permet de voir comment certaines d'entre elles exploitent plusieurs capteurs afin de pallier les faiblesses des uns avec les forces des autres. Cela est particulièrement visible dans le domaine des aides à la navigation basées sur les systèmes de positionnement par satellites, en raison des limitations nettes de ces derniers. L'intégration de plusieurs capteurs permet de servir deux objectifs : d'une part, l'augmentation de la robustesse de l'acquisition des données, et d'autre part, l'intégration de plusieurs fonctions dans un même appareil.

### 2.3.3 Forme et port

Les aides électroniques à la mobilité prennent de nombreuses formes.

1. La plus attendue et probablement la plus répandue est la canne électronique, ou sa variante sous la forme d'un boîtier à monter sur une canne classique. Cette forme tire probablement une partie de sa popularité de la filiation avec les cannes blanches et des représentations qui s'y rattachent, mais elle offre également des avantages certains aux utilisateurs. Tout d'abord, elle est moins dépayssante pour ceux qui sont habitués à l'usage d'une canne blanche. Ensuite, elle est à même de détecter les obstacles au sol, soit classiquement, par contact, soit par des capteurs qui peuvent se trouver très bas, près du sol. Cette seconde configuration reste néanmoins rare, peut-être à cause de la répartition peu avantageuse du poids qu'elle entraîne.
2. Un autre format commun est le boîtier à tenir dans la main dont la personne se sert à la manière d'une torche pour inspecter son environnement. Ce format est compact, facile à ranger, et discret. Le succès du Miniguide est très intéressant, car cet appareil ne fournit qu'une fonctionnalité très limitée, la détection d'obstacle [Research, 2005]. Ses avantages se situent sur les autres terrains : il est petit et discret, possède une autonomie importante et est visiblement fiable, et laisse à l'utilisateur le choix d'une interface tactile ou vibratoire.
3. Les smartphones, même s'ils ne sont pas destinés et conçus *a priori* à l'assistance aux personnes déficientes visuelles, sont néanmoins très intéressants car ils contiennent des capteurs inertiels, des caméras de qualité correcte et offrent la localisation par GPS, antennes de réseau mobile et points d'accès WiFi. Ils sont compacts et puissants, ont une autonomie acceptable, une connexion à Internet et une pile logicielle évoluée. Ces caractéristiques proviennent de la concurrence que se livrent leurs fabricants et de leur immense

diffusion dans la population mondiale, qui fait baisser les coûts de production. Cette popularité leur apporte deux avantages supplémentaires pour les applications qui nous intéressent : d'une part, il est fort possible qu'une personne déficiente visuelle possède un smartphone, indépendamment de toute aide à la mobilité ; ensuite, le smartphone est discret et n'engendre pas d'attention non souhaitée de la part du public, même quand il est utilisé ; enfin, un smartphone est flexible et il peut être utilisé avec de nombreuses applications selon les besoins du moment, évitant dans une certaine mesure le besoin de cumuler les appareils. L'intérêt des smartphones a été remarqué par Ahmetovic *et al.* [Ahmetovic et al., 2014] et plus généralement par tous les développeurs qui y publient des applications d'assistance.

4. On trouve également de nombreuses aides prenant place sur la tête de la personne, lunettes instrumentées voire, parfois, casques. Ces appareils exploitent les mouvements de la tête et le lien qui existe entre l'orientation de la tête et la représentation spatiale en plaçant des capteurs d'obstacles ou des caméras sur celle-ci. Le format casque semble peu à même de convaincre de nombreuses personnes du fait de son incongruité pour un piéton.
5. Certains appareils prennent la forme de ceintures ou, parfois, de vestes. L'intérêt de ce format est la possibilité d'ajouter une interface vibratoire non intrusive pour les gestes quotidiens : elle laisse les mains libres et offre une grande surface tactile. Les vestes sont utilisées quand une surface encore plus grande est nécessaire, pour la supplémentation sensorielle par exemple. La ceinture, et plus encore, la veste sont cependant gênantes et représentent une contrainte importante.
6. L'instrumentation des chaussures présente un intérêt non négligeable pour la possibilité qu'elles offrent aux systèmes de navigation d'utiliser le mécanisme du *Zero Velocity Update*, qui offre une précision correcte de localisation avec des capteurs inertiels. Les chaussures instrumentées représentent toutefois une contrainte importante pour la vie de tous les jours.
7. Certaines équipes de conception ont fait le choix d'aides qui prennent la forme d'un robot mobile que les utilisateurs suivent. Si cela peut paraître surprenant au premier abord, c'est en fait relativement naturel, les robots guides étant aux chiens guides ce que les cannes électroniques sont aux cannes blanches : un « remplacement ». Bien entendu, la comparaison est déséquilibrée : les capacités sensorielles et perceptives d'un chien sont largement plus difficiles à surpasser que celles d'une canne blanche. Mais ces robots, bien qu'encore loin d'une implémentation commerciale pratique, sont intéressants par les questions qu'ils soulèvent. D'abord pour la similarité évidente entre robotique mobile et aides électroniques à la mobilité : les outils sont en partie les mêmes, en particulier dans le domaine sensoriel des systèmes. Ensuite, la question de l'indépendance est intéressante. Selon les appareils, elle est plus ou moins marquée : le robot de Kulyukin *et al.* ou le *HARUNOBU-6* de Mori et Kotani ([Mori and Kotani, 1998]) sont capables, au moins matériellement, de déplacements sans utilisateur, ce n'est pas le cas de la *Guidecane*. Cette absence ou présence d'indépendance laisse imaginer des modes de fonctionnement différents : dans le cas où le sys-

tème est dépendant d'un utilisateur, on aurait tendance à supposer à une propriété individuelle, alors que dans l'autre cas, une mise à disposition publique pourrait être envisagée. Un ou des robots seraient attachés à un lieu et disponibles à la demande pour amener les personnes malvoyantes là où elles le désirent, puis capables de retourner à leur point d'attache – à la manière des assistants dans les aéroports ou les gares. Une étude des solutions robotiques existantes a été faite par Cuturi *et al.* [Cuturi et al., 2016]. Ils ont analysés dix robots d'assistance de la littérature et les ont comparés à quatorze appareils d'assistance classiques. Les robots offraient généralement de meilleures performances pour l'évitement des obstacles et la navigation (il suffit de les suivre), mais posaient évidemment des problèmes de portabilité et avaient une mobilité limitée, en particulier dans les escaliers.

### 2.3.4 Interface utilisateur

Outre les caractéristiques techniques, il est important de prendre en compte les interfaces utilisateur dans l'analyse des aides à la mobilité. Les différentes interfaces employées par les appareils présentent toutes des avantages et des inconvénients. Une interface audio est à même de communiquer plus d'informations qu'une interface tactile vibratoire ponctuelle, mais elle risque de couper l'utilisateur du monde extérieur, ce qui est inconfortable [Brabyn, 1982, Giudice and Legge, 2008]. À l'intérieur même d'un type d'interface, on trouve des variations importantes. Ainsi, dans les interfaces audio, on trouve par ordre de gêne décroissante : casque fermé, casque ouvert, haut parleur à proximité des oreilles, et conduction des sons par les os.

D'un autre côté, une interface vibratoire peut être inintelligible ou inutile quand elle vibre trop (ainsi que l'ont rapporté les participants de l'étude de Vincent *et al.* [Vincent et al., 2014]).

En conclusion, l'interface a autant, sinon plus, d'importance que la technologie derrière. Ross et Blasch rapportent que les utilisateurs sont généralement plus intéressés par l'interface des appareils d'assistance que par la technologie qui les fait fonctionner [Ross and Blasch, 2002]. Ce jugement est légitime, car, dans un appareil, c'est l'interface qu'ils seront amenés à manipuler. La technologie en elle-même a moins d'importance que les fonctionnalités qu'elle offre et la manière dont elle le fait. Dans le cas de systèmes d'assistance à la perception tels que les systèmes de supplémentation sensorielle, cette interface revêt une criticité d'autant plus grande que c'est elle qui détermine si l'appareil fonctionne effectivement ou pas : elle doit être transparente pour permettre aux percepts d'émerger.

### 2.3.5 Fiabilité

La fiabilité est un élément important des systèmes. On a vu qu'elle était mentionnée par les participants de l'étude de Vincent *et al.* sur le *Miniguide* et le *Breeze* [Vincent et al., 2014]. En effet : pour une personne dénuée de vision, les erreurs ont

des conséquences très lourdes. D’abord, parce que le domaine de ce qui constitue une erreur est considérablement élargi. Même des actions qui pourraient sembler triviales, comme tourner une rue trop tôt sur un trajet, peuvent mener très rapidement à l’égarement. Ensuite, parce que les conséquences des erreurs sont amplifiées : une personne qui heurte un objet a plus de chances de tomber, une personne qui tombe ne sait pas où elle tombe, une personne égarée n’a pas ou presque d’information sur le lieu où elle se trouve et ne peut probablement pas retourner sur ses pas. Il est facile pour une personne voyante de revenir sur ses pas et de retrouver un point connu, mais pas pour une personne aveugle. Il est nécessaire que les aides à la mobilité ne soient pas un vecteur de risque pour leurs utilisateurs : leur fiabilité doit donc être exemplaire. Les personnes n’utiliseront un système d’assistance que si elles sont convaincues qu’elles ne risquent pas de se trouver dans une mauvaise situation. L’étude de Vincent *et al.* est d’ailleurs révélatrice de ce phénomène : les utilisateurs et les instructeurs de mobilité mentionnent l’usage qui peut être fait de ces aides dans des contextes *déjà connus* mais pas lors de la découverte de nouveaux lieux, qui doit se faire avec une assistance externe.

### 2.3.6 Infrastructures

Bien que les infrastructures soient souvent peu visibles dans la sélection d’aides électroniques à la mobilité étudiées ici, elles représentent un point central des problématiques. Nous définissons ici les infrastructures comme tout équipement électronique déployé dans l’environnement pour permettre une fonctionnalité. Les antennes de téléphonie mobile et Internet en sont un exemple. Certains équipements peuvent être considérés ou non, selon le point de vue, comme formant une infrastructure : c’est le cas par exemple des points d’accès WiFi. Pour la personne qui déploie un seul point d’accès, celui-ci n’est pas une infrastructure, alors que c’en est une pour le système qui utilise les points d’accès pour se localiser. Ici, nous nous plaçons du côté des appareils utilisateurs, et le périmètre de ce que nous considérons comme infrastructure est donc large.

Par définition, les infrastructures nécessitent que des équipements soient déployés sur les zones où l’on souhaite voir fonctionner le système. Ce déploiement représente presque toujours un coût important, pour plusieurs raisons. Tout d’abord, à l’évidence, il y a le coût incompressible d’achat des équipements. Ensuite, on trouve leur coût d’entretien ou de maintien, dont un exemple typique dans notre secteur est le changement de batterie des objets intelligents. Des pannes éventuelles des équipements doivent également être prises en compte. Ensuite, il y a le coût de gestion de l’espace où se trouvent ces équipements. Ils ne peuvent pas être installés n’importe où ; même si les lieux réceptacles ne sont pas forcément à proprement dit loués, il est extrêmement probable que le déploiement requiert d’importantes réflexions, discussions, autorisations et autres remplissages de dossiers. Tout ceci représente un coût soit monétaire, parce qu’il faut payer des personnes pour s’en occuper<sup>12</sup>, soit, au minimum, un coût en temps. Si une fonctionnalité peut être fournie de plusieurs manières, des problématiques de manque d’interopérabilité

---

12. Cela devient inévitable passé un certain niveau de déploiement.

entre les infrastructures administrées par les différents acteurs deviennent inévitables. Les problèmes liés aux infrastructures sont amplifiés par l'échelle à laquelle il est souhaitable qu'elles soient déployées. Malheureusement, notre cadre profite d'une couverture la plus étendue possible : une aide à la mobilité qui fonctionne dans la moitié d'une ville est moyennement utile ; c'est probablement encore pire si la couverture du service est morcelée. Une couverture nationale voire continentale ou mondiale est aussi souhaitable que difficile, pour ne pas dire impossible à réaliser à moyen terme.

Lors de la conception d'un système d'assistance, il est nécessaire de faire des choix relatifs aux infrastructures : utilisation ou non, quelle infrastructure, etc. Si toutes ces questions doivent être étudiées dans leur contexte, il est néanmoins possible d'extraire des points clés qui ont une influence considérable sur les probabilités de succès pratique du système. Nous avons identifié trois de ces points clés :

- Qui supporte la charge de l'infrastructure ? En particulier, est-elle gérée par un seul, un petit ensemble, ou par un grand nombre d'acteurs ?
- Quelle est la destination de l'infrastructure ? Est-elle spécifique à l'application d'assistance aux personnes malvoyantes, ou bien vise-t-elle à fournir d'autres services, sans souci particulier pour les problématiques liées à la malvoyance ?
- L'infrastructure est-elle déjà pré-existante ou bien faut-il la déployer ?

Le tableau 2.1 recense toutes les combinaisons possibles de points clés et des exemples réels pour certaines de ces combinaisons.

TABLE 2.1 – Combinaisons possibles des caractéristiques d'une infrastructure.

Charge	Destination	Pré-existence	Exemple	Référence
Peu d'acteurs	PMV	Oui	Balises sonores	[Alexandre et al., 2018]
Peu d'acteurs	PMV	Non	Talking Signs	[Loughborough, 1979]
Peu d'acteurs	Autre	Oui	Sherpa / balises	[Handisco, 2019]
Peu d'acteurs	Autre	Non	Reconnaissance points d'intérêt	[Serrão et al., 2012]
Nombreux acteurs	PMV	Oui		
Nombreux acteurs	PMV	Non	Talking Signs	[Loughborough, 1979]
Nombreux acteurs	Autre	Oui	Localisation WiFi	[Lymberopoulos et al., 2015]
Nombreux acteurs	Autre	Non	Chapitre IoT	

Les contraintes sur la réussite d'un système seront différentes selon le type d'infrastructure. L'utilisation d'une infrastructure existante est beaucoup plus favorable que la dépendance au déploiement d'une nouvelle infrastructure. Cette infrastructure peut être utilisée dans son optique originelle, comme dans le cas de la possibilité d'utiliser les balises sonores avec une canne Sherpa. Elle peut également être

utilisée d'une manière différente de celle pour laquelle elle a été conçue : c'est le cas de la localisation avec des points d'accès WiFi, qui sont destinés avant tout à fournir un accès au réseau. La question de la charge ou responsabilité de l'infrastructure est également importante. D'un côté, plus la responsabilité est concentrée sur un faible nombre d'acteurs, plus l'infrastructure sera homogène et les informations la concernant centralisées, mais d'un autre côté, plus la charge sera lourde pour cet acteur. À l'inverse, une responsabilité plus distribuée abaisse considérablement les aspects d'investissement, le rendant plus probable, mais augmente les risques d'avoir des problèmes d'interopérabilité. Les réseaux sans fil locaux sont, encore une fois, un exemple de bonne situation distribuée. Les points d'accès sont à la charge des utilisateurs, ce qui facilite leur déploiement et leur gestion, et les problèmes d'interopérabilité sont absents grâce à la domination de la famille de standards IEEE 802.11.

Les cartes sont soumises à des problématiques relativement similaires à celles des infrastructures. En effet, en plus de l'importance primordiale de l'aspect géographique qui leur est caractéristique, la structuration du coût est similaire. Toutes deux nécessitent des dépenses à la fois à la création et lors de « l'entretien », c'est-à-dire la mise à jour des cartes pour prendre en compte les changements survenus dans la réalité. Il est à noter que des mises à jour fréquentes sont particulièrement importantes dans les applications d'assistance aux personnes déficientes visuelles en raison des dangers qu'une carte désynchronisée peut induire ; dans l'idéal, les cartes devraient au moins être actualisées avec les données concernant les travaux de voirie. L'administration et la responsabilité des données cartographiques est également une source de difficultés. Si, en extérieur, il est possible de se fournir en cartes auprès d'acteurs offrant une très bonne couverture, les données OpenStreetMap<sup>13</sup> étant en accès libre, ou les données Google Maps<sup>14</sup> avec lesquelles il est également facile de s'interfacer, la situation concernant les espaces intérieurs est tout autre. En effet, il n'existe pas de standard (y compris *de facto*) largement adopté et diffusé pour la représentation et l'échange de données cartographiques sur les intérieurs. Plusieurs obstacles s'y opposent. La question de ce qui doit être cartographié (la couleur des portes doit-elle être prise en compte ?) est un véritable problème ; cela dépend des applications. Il est donc relié à un autre obstacle : l'utilité de la cartographie intérieure, ou plutôt, de la diffusion des plans, est limitée. En fait, des plans existent dans presque tous les cas : plans d'architectes, plans d'évacuation, etc. Seulement, il ne sont pertinents que pour des contextes très spécifiques pour lesquels la diffusion publique (autrement que locale et visuelle) n'est pas considérée comme nécessaire. En fait, cette diffusion publique est même potentiellement considérée comme néfaste, pour des raisons de confidentialité et de sécurité. Par exemple, si quelqu'un utilise les plans d'un bâtiment pour y mener une attaque terroriste, qui en prendra la responsabilité ? En réalité, les considérations de cartographie et de confidentialité ne se posent presque jamais : le besoin de cartographie n'est tout simplement pas exprimé. Certains lieux vastes et complexes, comme les gares et les aéroports, sont de plus en plus intégrés dans les fonds

---

13. <https://www.openstreetmap.org>

14. <https://www.google.com/maps/>

des fournisseurs de cartes en ligne. Le projet OpenStationMap<sup>15</sup> vise par exemple à cartographier les gares. Mais la cartographie en intérieur demeure un domaine encore peu stable, en témoignent les discussions et différents standards proposés au sein même d'OpenStreetMap<sup>16</sup>.

### 2.3.7 Analyse par rapport au positionnement perceptif

Malgré la grande diversité qui caractérise les aides électroniques à la mobilité, on peut observer deux doctrines qui structurent la conception d'une grande partie d'entre elles. Certains appareils visent à proposer une assistance à la mobilité la plus générale possible, tandis que d'autres cherchent à proposer *une* fonctionnalité la plus efficace possible. Ces deux tendances sont bien représentées par les appareils de substitution visuo-tactile d'un côté (tel que [Bach-y Rita et al., 1998]), et les systèmes de reconnaissance d'un type particulier d'objet de l'autre côté (à l'instar de [Yi and Tian, 2015]). Ce clivage a été relevé dans la littérature : Giudice et Legge mentionnent l'opposition entre l'approche consistant à fournir aux utilisateurs autant d'informations que possible sur l'environnement, et celle présupposant que l'utilité repose sur la simplicité de l'interface [Giudice and Legge, 2008]. Malgré les nombreux succès obtenus dans les expérimentations en laboratoire, une supplémentation sensorielle intégrale de la vision semble difficile à cause de la complexité de l'environnement, qui augmente la charge cognitive [Elli et al., 2014].

Cette opposition est également visible au sein de la sous-catégorie des appareils destinés à l'assistance à la *marche* : le *Miniguide* est ce qu'on pourrait appeler un détecteur d'obstacles minimaliste, alors que les *Sonic Glasses* ou la partie substitution sensorielle du système *Sound of Vision* présentent une interface beaucoup plus complexe à l'utilisateur. À un niveau plus haut de différenciation, on trouve la forme qui est probablement la plus franche de la première tendance : l'objectif pour un appareil d'assister plusieurs tâches de la mobilité. Il est naturel de vouloir comparer les deux approches dans l'absolu, de chercher laquelle est la plus pertinente. Le succès du *Miniguide* et les échecs des appareils de substitution sensorielle pourraient inciter à prendre une position radicalement orientée vers les systèmes à fonctionnalité restreinte, mais la problématique n'est pas si simple. Le potentiel des systèmes de supplémentation sensorielle est immense ; mais difficile à concrétiser en un seul et même appareil réellement utilisable. Cette difficulté est particulièrement marquée dans le domaine des systèmes de *substitution* visuo-tactiles, qui sont historiquement les premiers et les plus marquants médiatiquement. Ils ont donc une influence forte sur l'association entre « mauvaise conception » et systèmes de *supplémentation* sensorielle. Des systèmes plus limités sont également possibles avec la supplémentation sensorielle : Loomis *et al.* mentionnent leur préférence pour de tels systèmes plutôt que ceux à visée générale [Loomis et al., 2012] et donnent des conseils pour éviter les pièges à la conception, qui s'articulent généralement autour de la prise en compte de la sensibilité des utilisateurs potentiels et pas seulement des considérations techniques et informationnelles. Il serait dommage d'écarter les

---

15. <https://www.openstationmap.org>

16. [https://wiki.openstreetmap.org/wiki/Indoor\\_Mapping](https://wiki.openstreetmap.org/wiki/Indoor_Mapping)

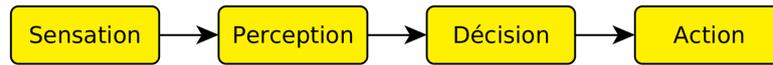


FIGURE 2.18 – Modèle de la chaîne perceptive

aides intégrant plusieurs fonctionnalités pour des raisons théoriques, car leurs avantages sont importants : ils permettent en fait, s'ils sont réussis, une simplification de l'interface présentée aux usagers car ceux-ci n'ont à gérer qu'un seul point d'entrée. L'application Microsoft *Seeing AI* en est un bon exemple : elle propose de nombreuses fonctionnalités dans un format bien intégré [Microsoft, 2017]. Cet exemple n'est en fait qu'à moitié représentatif : le cas est plutôt simple car les fonctionnalités incluses sont similaires. Dans les situations plus classiques (navigation, obstacles, etc.), les difficultés pratiques de réalisation de tels dispositifs provoquent souvent l'appartition de défauts.

L'opposition des approches entre appareils à visée simple et appareils à visée générale est également intéressante parce qu'elle met en lumière ce que nous appellerons le *positionnement perceptif*. On peut observer cela dans la séparation des deux types de cannes électroniques : les cannes détectrices d'obstacles dirigent beaucoup plus l'utilisateur que les capteurs environnementaux, qui enrichissent la représentation spatiale de l'utilisateur afin de lui permettre la prise de décision. Plus précisément, nous définissons ce positionnement perceptif comme étant l'étape de la chaîne perceptive à laquelle un système intervient. Cette « chaîne perceptive »<sup>17</sup> est la succession des étapes visibles sur la figure 2.18 :

- « Sensation » : les informations environnementales sont acquises ;
- « Perception » : les informations sont intégrées dans une représentation mentale ;
- « Décision » : les informations sont utilisées pour porter un jugement conscient ou non ;
- « Déplacement » : la réalisation du mouvement déterminé.

Ce découpage reste cependant grossier et approximatif : on l'a vu au chapitre précédent, les frontières entre sensation, perception et mouvement ne sont pas aussi hermétiques qu'on pourrait l'imaginer dans un premier temps. L'étape de la décision est elle aussi discutable : peut-on parler de décision dans le cas où celle-ci est inconsciente, comme peut par exemple l'être un évitement d'obstacle ? Malgré son caractère approximatif, nous pensons que ce modèle présente un intérêt parce qu'il peut s'appliquer non seulement aux personnes mais également aux dispositifs d'assistance – en fait, le modèle les décrit plus justement. En mettant en parallèle la chaîne perceptive d'une personne et celle d'un dispositif et en spécifiant le cheminement de l'information, on forme le système  $\{personne + dispositif\}$  d'assistance et il est alors possible de décrire le « positionnement perceptif » des dispositifs par rapport aux usagers (figure 2.19). Nous avons relevé sept cheminements distincts de l'information à travers le système qui semblent pertinents (figure 2.20) :

1. Le dispositif se contente de collecter des données sur l'environnement, de leur appliquer un traitement simple et de les présenter à l'utilisateur. Nous considérons ici comme « simple » un traitement dont la finalité n'est pas l'ex-

17. Que nous nommons ainsi en raison du rôle central que la perception y tient.

traction complète d'une information sémantique. La plupart des systèmes de supplémentation ou substitution sensorielle visuo-tactiles fonctionnent ainsi : les images sont floutées et converties en un nombre limité de niveaux de gris avant d'être présentées à l'utilisateur. Les systèmes de sonification effectuent aussi un filtrage pour simplifier les images, qui contiennent trop d'informations. Le fonctionnement de l'interface ne permet pas l'émergence du phénomène d'attribution distale : quand les informations sont présentées, la personne « sent » l'interface plus qu'elle ne perçoit directement les objets présentés. Elle peut se servir de ces informations, mais cela nécessite des efforts importants ;

2. Comme précédemment, le dispositif collecte des données environnementales, leur applique un traitement simple et les présente à la personne. Mais dans ce cas, l'interface est efficace et permet la perception directe de l'objet. C'est ainsi que fonctionnent ou devraient fonctionner la plupart des systèmes de substitution sensorielle ;
3. Dans ce cas, le dispositif analyse les données qu'il collecte afin d'en extraire des informations de haut-niveau, typiquement présentant des objets sémantiques identifiés dans l'environnement, et ce sont ces informations qui sont présentées à l'utilisateur. Néanmoins, comme dans le premier cas, le fonctionnement de l'interface n'est pas efficace et ne permet pas l'émergence de l'attribution distale et la perception directe de l'objet. C'est le cas d'un système de supplémentation ou substitution sensorielle évolué (extraction sémantique) mais inopérant ;
4. Comme le cas précédent mis à part que le système est opérant et les informations transmises permettent la perception directe des objets. L'étape de perception à l'intérieur du dispositif est intéressante pour réduire la quantité d'informations à transmettre à la personne, en raison de la bande passante limitée des sens autres que la vision [Loomis et al., 2012]. Cette approche est notamment employée dans le projet Sound of Vision [Strumillo et al., 2018] ;
5. Dans cette situation, le dispositif d'assistance prend une décision, ou émet un jugement sur l'action que devrait effectuer la personne par rapport à la situation environnementale. Ce jugement est présenté à l'utilisateur, souvent sous forme linguistique, soit dans notre cas par synthèse vocale, soit, dans les cas où le jugement est simple, par une interface scalaire – vibrations de plus en plus fortes, fréquence changeante d'une tonalité – voire binaire. La personne entend cette décision et s'en sert, avec d'autres critères, pour élaborer sa propre décision. C'est ainsi que le fonctionnement de la plupart des capteurs d'obstacles simples est décrit dans le modèle. Par exemple, si un dispositif indique un obstacle à un endroit où la personne vient de passer et qu'elle sait que ce dispositif a une fiabilité limitée, la personne pourra aller contre le jugement du dispositif et traverser l'endroit sans particulièrement se méfier. Les systèmes de navigation dits « *turn-by-turn* » suivent également ce mode de fonctionnement, puisqu'ils donnent à la personne des instructions directes : « tournez à gauche », « allez tout droit sur 50 pas », etc. ;
6. Le dispositif d'assistance fonctionne ici comme dans le cas précédent, la

différence se trouvant dans le comportement de la personne face aux instructions reçues. Dans cette situation, la personne n'élabore pas sa décision propre en utilisant d'autres sources, elle se contente de suivre la décision du dispositif. Bien que ce comportement soit en première apparence déterminé par des phénomènes subjectifs (comme la personnalité), il peut être motivé par des contraintes objectives, en particulier celle du manque d'information venant d'une autre source – connaissances, perception auditive, autre dispositif d'aide, etc.

7. Dans le dernier cas, le dispositif d'assistance prend en charge toute la chaîne perceptive, jusqu'à l'action : la personne est passive. Des informations de contrôle, destinées à informer la personne, peuvent lui être transmises, mais ce n'est pas obligatoire. Une personne qui prend un taxi et est transportée jusqu'au point désiré se trouve dans ce cas de figure. En fait, cette situation n'est pas limitée aux personnes déficientes visuelles : toute personne empruntant un transport en commun s'y trouve incluse. Bien entendu, les notions de décision et d'action sont subordonnées au niveau hiérarchique occupé par la sous-tâche de mobilité considérée : si on reprend l'exemple du taxi, le moyen (chemin) de parvenir à destination forme un ensemble de décisions et d'actions que prend en charge le chauffeur, mais la destination elle-même est décidée par la personne transportée.

Ces différents cheminements sont résumés dans le tableau 2.2. Certaines situations ne sont pas couvertes par les sept cheminements décrits ; de plus, de nombreux processus réels utilisent plusieurs cheminements à la fois. Par exemple, en fonction de son entraînement et de son niveau de fatigue, une personne utilisant un système de supplémentation sensorielle flottera entre le premier et le second cas. Un système de navigation pourra alterner entre des instructions de guidage (5<sup>e</sup> cas) et des indications des points d'intérêt, qui pourront venir enrichir la représentation spatiale de la personne (4<sup>e</sup> cas). Ce cas des aides offrant plusieurs fonctionnalités est intéressant, car l'interface doit transmettre des informations de natures différentes, et empruntera souvent plusieurs chemins pour ce faire. Les variations temporelles et contextuelles sont également probables entre les chemins 6 et 7, selon les informations dont peut disposer la personne. De manière générale, un dispositif ou un ensemble de dispositifs peut emprunter plusieurs cheminements : un en particulier dans un mode préférentiel et un autre dans un mode « dégradé ». En général, un mode dégradé est, du point de vue de la fonctionnalité du dispositif, identifié par une stagnation ou une régression dans la chaîne perceptive effective, car cela signifie que la personne est amenée à refaire une opération que le dispositif prend en charge.

La charge mentale est affectée par la longueur de la chaîne se situant du côté de l'utilisateur. Même si les importantes variations possibles entre des instances différentes du même cheminement peuvent conduire des cheminements théoriquement plus efficaces à engendrer une plus forte charge mentale en réalité, la caractéristique générale des cheminements est la diminution de la charge mentale quand les tâches à effectuer par la personne diminuent. En contrepartie de cet allègement, la personne est moins libre et plus isolée de son environnement. Pour maximiser le potentiel d'adoption d'un dispositif, un bon équilibre doit être trouvé entre ces

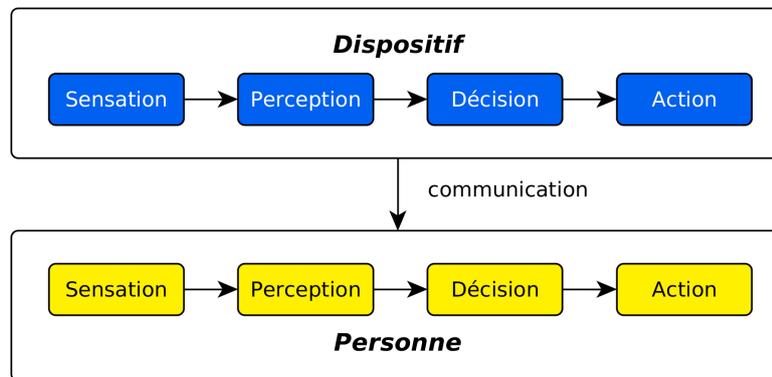
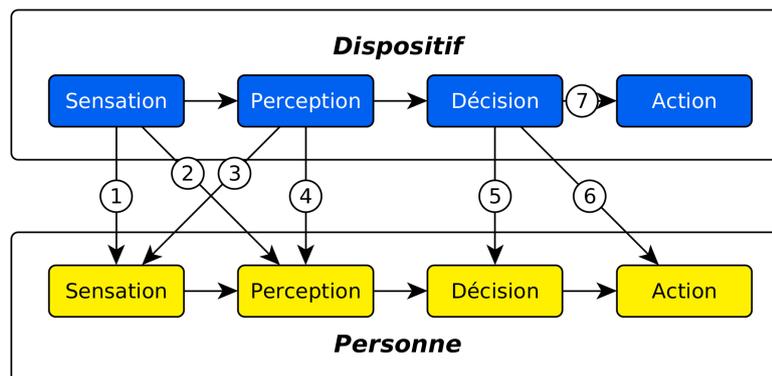
FIGURE 2.19 – Modèle du système  $\{dispositif\ d'assistance + personne\}$ .FIGURE 2.20 – Cheminements possibles de l'information dans le système  $\{dispositif\ d'assistance + personne\}$ .

TABLE 2.2 – Tableau décrivant les différents cheminements possibles suivant la numérotation de la figure 2.20.

n°	Description	Informations transmises	Exemple
1	Le dispositif collecte des données et les transmet à la personne, qui les « ressent » et les interprète ; peu efficace	Données brutes ou pré-traitées	Dispositif de supplémentation sensorielle inopérant, sans attribution distale
2	Le dispositif collecte des données et les transmet, faisant émerger des percepts ; efficace	Données brutes ou pré-traitées	Dispositif de supplémentation senselessness opérant, avec attribution distale, comme le TVSS [Bach-y Rita et al., 1969]
3	Le dispositif collecte des données, les interprète et les transmet à la personne, qui les « ressent » et les interprète ; peu efficace	Données interprétées	Dispositif de supplémentation sensorielle inopérant, sans attribution distale
4	Le dispositif collecte des données, les interprète et les transmet, faisant émerger des percepts	Données interprétées	Dispositif de supplémentation sensorielle opérant, avec attribution distale, comme le dispositif Sound of Vision [Sound of Vision Consortium, 2018]
5	Le dispositif collecte des données, les interprète et décide d'une action propice, qui est transmise et jugée par la personne	Décision	Dispositif de navigation par instructions de guidage (« tournez à gauche » par exemple) utilisé comme support dans un lieu connu
6	Le dispositif collecte des données, les interprète et décide d'une action propice, qui est transmise et suivie par la personne	Décision	Dispositif de navigation par instructions de guidage (« tournez à gauche » par exemple) utilisé dans un lieu inconnu
7	Le dispositif collecte des données, les interprète, décide d'une action propice, et l'implémente	Aucune ou contrôle	Dispositif prenant totalement en charge la personne (p. ex. : taxi).

deux pôles – cela en partant de l'hypothèse que les interfaces sont correctement implémentées et efficaces. Dans l'idéal, cet équilibre pourrait être ajustable par les usagers, qui auraient le choix de la quantité d'informations sur l'environnement. En pratique, il est difficile de permettre une telle flexibilité, car il existe une affinité forte entre fonctionnalités et cheminement de l'information. Les dispositifs à plusieurs fonctionnalités sont une exception dans ce domaine : le niveau d'assistance peut être modifié par la simple activation ou désactivation de fonctionnalités.

### 2.3.8 Résumé par rapport à la problématique de la thèse

Les principales caractéristiques des dispositifs étudiés ont été réunies dans le Tableau 2.3. C'est un résumé qui n'apporte sans doute pas toutes les informations nécessaires pour un jugement objectif des aides électroniques à la mobilité<sup>18</sup>, mais apporte rapidement des informations sur les différentes approches, capteurs et interfaces existantes.

La difficulté de la problématique est mise en valeur par la faible adoption des aides électroniques à la mobilité. Un très grand nombre a été développé, mais peu d'entre elles dépassent véritablement le stade du prototype et sont réellement utilisées [Roentgen et al., 2008]. Même quand elles le sont, ce n'est que partiellement, à la fois parce qu'elles sont peu répandues et parce qu'elles ne sont pas utiles dans un grand éventail de situations ; les aides peuvent par exemple être abandonnées après une période initiale de familiarisation avec un environnement [Roentgen et al., 2009].

Les tests d'aides à la mobilité sont peu nombreux et les méthodologies déployées peu qualitatives [Roentgen et al., 2009] : petits nombres de participants, populations hétérogènes ou biaisées et ne correspondant pas aux populations ciblées, sources internes des critères de jugement. Il n'est donc pas évident de comparer des dispositifs entre eux.

Ce constat peut paraître plutôt négatif pour les aides électroniques mais il doit être tempéré, au vu du déploiement également limité des aides traditionnelles que sont les chiens guides et les cannes blanches. L'exemple des chiens guides est frappant : en France, sur 1,7 million de personnes malvoyantes et aveugles, moins de 1 % utilisent un chien-guide [Zucker and Alexandre, 2018]. Cela tient probablement à plusieurs raisons, au centre desquelles on trouve l'investissement personnel et financier important [Wirth and Rein, 2008] qu'ils nécessitent. Le cas de la canne est plus intéressant. En effet, elle n'a pas les mêmes désavantages objectifs que les chiens-guides : elle est abordable financièrement, présente peu de contraintes à l'usage, sinon qu'il faut l'emporter, et permet d'améliorer la sécurité. Si la canne est plus répandue que le chien guide, elle n'est néanmoins pas non plus adoptée par la majorité de la population concernée : aux États-Unis, il était estimé que seuls 42 % de la population aveugle utilisait une canne en 1990 [Manduchi and Kurniawan, 2011].

---

18. Par exemple les deux applications smartphone Blindsquare et Seeing Eye GPS y sont représentées équivalentes alors qu'elles sont à l'usage différentes ; les dissocier nécessiterait l'introduction d'une caractérisation plus fine de l'interface.

Pour être adoptée par une population diverse, une aide électronique à la mobilité doit être suffisamment satisfaisante au regard de nombreuses contraintes. Ces contraintes ont été identifiées et analysées dans les sous-sections précédentes : une fonctionnalité suffisamment pertinente, la robustesse de l'acquisition des données et de l'appareil en général, une forme la plus discrète et confortable possible, une interface capable de transmettre les informations sans créer de gêne, une dépendance aux infrastructures bien gérée.

La suite de cette thèse présente les travaux sur le développement d'une aide dont la conception pourrait satisfaire ces contraintes et répondre à la problématique initiale.

TABLE 2.3: Tableau récapitulatif des dispositifs analysés, classés suivant leurs fonctionnalités et leurs caractéristiques (*SV* désigne la synthèse vocale, *SS* la substitution ou supplémentation sensorielle, *3D* les caméras de profondeur (double capteur, temps-de-voil ou à lumière structurée), *IMU* désigne une centrale inertielle (*Inertial Measurement Unit* en anglais) et *rec.* est l'abréviation de *reconnaissance*).

Système	Référence	Année	Fonctionnalités	Capteurs	Forme/port	Interface
Pathsounder	Russel [1965]	1965	obstacles	ultrasonique	poitrine	audio
Mowat	Pressey [1977]	1977	obstacles	ultrasonique	main	tactile
Laser Cane	Jansson [1975]	1975	obstacles	laser	canne	Audio + tactile
Miniguide	Research [2005]	2011	obstacles	ultrasonique	main	audio + tactile
Tom Pouce	Farcy et al. [2006]	2006	obstacles	infrarouge	main/canne	tactile
Sonic Torch	Kay [1964]	1964	relief	ultrasonique	main	audio
Teletact	Farcy et al. [2006]	2006	relief	laser	canne	audio + tactile
EyeCane	Maidenbaum et al. [2014]	2014	relief	infrarouge	canne	audio + tactile
TVSS	Bach-y Rita et al. [1969]	1969	SS visuo-tactile	caméra vidéo	dos	tactile
TVSS – langue	Bach-y Rita et al. [1998]	1998	SS visuo-tactile	caméra vidéo	langue + caméra	électro-tactile
Sonic Glasses	Kay [1974]	1974	obstacles/SS tactile	caméra vidéo	lunettes	audio
The vOICE	Meijer [1992]	1992	SS audio-tactile	caméra vidéo	lunettes	audio

Sound of Vision	Sound of Vision Consortium [2018]	2018	SS audio-tactile + rec. Textes, panneaux	Caméra + 3D + IMU	casque + ceinture	audio + tactile + SV
PGS	Loomis et al. [1994]	1994	navigation	GPS	sac	audio 3D, SV, tactile
Geotact	Farcy et al. [2006]	2006	navigation	GPS + IMU	épaule	SV
Breeze	Vincent et al. [2014]	2014	navigation	GPS	main	SV
Seeing Eye GPS	Sendero Group [2019]	2019	navigation	localisation smartphone	smartphone	SV, tactile
Blindsquare	MIPSoft [2019]	2019	navigation	localisation smartphone	smartphone	SV, tactile
Filipe et al.	Filipe et al. [2012]	2012	obstacles	3D	?	?
Tapu et al.	Tapu et al. [2013]	2013	obstacles	caméra vidéo	smartphone	?
Tian et al.	Tian et al. [2010]	2010	reconnaissance portes	caméra vidéo	smartphone	?
Yi et al.	Yi et al. [2013]	2013	recherche objets	caméra + réseau caméra	smartphone	SV
Yi et Tian	Yi and Tian [2015]	2015	lecture texte	caméra vidéo	smartphone	SV?
Wang et Tian	Wang and Tian [2011]	2011	reconnaissance signalisation	caméra vidéo	smartphone	?
Ivachenko et al.	Ivachenko et al. [2008]	2008	alignement passages piétons	caméra vidéo	smartphone	audio

Crosswatch	Coughlan and Shen [2013]	2013	multi-aides piétons	passages	caméra localisation smartphone	+	smartphone	SV + audio + tactile
Seeing AI	Microsoft [2017]	2019	description scène + objets	scène +	caméra vidéo		smartphone	SV
Balises feux	Alexandre et al. [2018]	2018	signalisation piéton	passage	radio 868 MHz		télécommande	audio
Talking Signs	Loughborough [1979]	1979	informations contextuelles	context-	infrarouge		télécommande	SV
Ross et Blasch	Ross and Blasch [2002]	2002	navigation		RFID		canne	SV
iCane	Chang et al. [2005]	2005	informations contextuelles	context-	RFID		canne	SV
Kassim et al.	Kassim et al. [2013]	2013	navigation		GPS + RFID		canne	?
Kulyukin et al.	Kulyukin et al. [2006]	2006	navigation en intérieur	inté-	RFID		robot-guide	suivi
GuideCane	Borenstein and Ulrich [1997]	1997	obstacles + navigation	navigation	ultrasoniques + odométrie + GPS		Robot-guide	suivi
Handisco Sherpa	Handisco [2019]	2019	navigation + trans-ports + balises	trans-	GPS + radio 868 MHz + ?		canne	SV + audio

# Chapitre 3

## Développement de la canne intelligente 2SEES

Afin de travailler sur les problématiques de mobilité des personnes déficientes visuelles introduites dans le chapitre 1, un prototype de canne intelligente a été conçu et fabriqué. Cette canne, que nous avons nommée *Local MIMO Smart Cane*<sup>1</sup> (LMSC), est au centre du dispositif 2SEES, qui comporte également un smartphone et une plate-forme *cloud computing*. Le 2SEES est basé sur le système *Smart Environment Explorer Stick* (SEES) développé précédemment dans l'équipe SMIR<sup>2</sup>. Ce chapitre présente le système SEES dans une première section, suivie d'une autre consacrée aux motivations pour le développement d'un nouveau système et son architecture. Vient ensuite une section décrivant la plate-forme matérielle (mécanique et électronique). Les deux dernières sections exposent les travaux sur les deux fonctionnalités de base du système : détection d'obstacles et navigation.

### 3.1 Développement du SEES

Le projet *Smart Environment Explorer Stick*, abrégé en SEES, a été initié par l'équipe SMIR en collaboration avec l'association GAIPAR et plus particulièrement avec de son président, M. Daniel Jacquet. Son principal acteur a été Muhammad Yusro, dont c'était le sujet de la thèse de doctorat [Yusro, 2017].

#### 3.1.1 Motivation et concept

##### Motivation

L'objectif du projet SEES était la création d'une aide à la mobilité réunissant plusieurs fonctions [Yusro, 2017] :

1. Ce qui signifie à peu près « canne locale intelligente à plusieurs entrées et plusieurs sorties ».
2. Équipe Système Multisensoriel Intégré Intelligent Réparti. Site web : <http://edss.isima.fr/sites/smir/> (accès le 14 septembre 2019).

1. Améliorer l'intégration sociale des personnes malvoyantes ou aveugles ;
2. Permettre la navigation dans les milieux connus et inconnus, en intérieur comme en extérieur ;
3. Aider à l'orientation et améliorer la conscience de l'espace ;
4. Aider à la détection et l'évitement des obstacles ;

Ces fonctions devaient être réalisés par un système intégrant un smartphone et une canne intelligente, laquelle comporterait plusieurs capteurs et des mécanismes d'adaptation au contexte.

## Concept

Le système SEES est construit autour de trois composants : une canne intelligente nommée *SEE-Stick*, un serveur local embarqué dans un smartphone nommé *SEE-phone*, et un serveur distant nommé *iSEE* (figure 3.1). Le serveur *iSEE* est conçu pour fournir des services web :

1. Il permet à des tiers de confiance de suivre les personnes malvoyantes à distance ;
2. Il permet un mode d'assistance directe, dans lequel la personne filme son environnement, transmet les images à un opérateur humain distant qui en retour guide vocalement la personne malvoyante. Ce mode est prévu pour pallier les défaillances – réelles ou perçues – des fonctionnalités du système lorsque celui-ci travaille en autonomie ;
3. Il gère les données cartographiques pour le système : stockage, distribution, analyses, etc.

Le *SEE-phone* est un smartphone du commerce utilisé comme serveur local pour fournir des services à la canne intelligente *SEE-stick*. Il est responsable du calcul des itinéraires, et de la gestion de l'accès à Internet, qui comprend en particulier l'accès au serveur *iSEE*, nécessaire pour obtenir les cartes locales.

Le smartphone est également destiné à accueillir d'autres fonctions : une application de détection de l'état des feux tricolores et une autre destinée à la détection des poteaux ont ainsi été développées par M. Yusro [Yusro et al., 2013b, Yusro, 2017]. Des fonctionnalités comme la reconnaissance vocale, natives sur smartphone, sont également destinées à être utilisées pour améliorer l'accessibilité du système. La canne *SEE-stick* est destinée à être connectée en permanence avec le *SEE-phone* par une liaison WiFi. Tandis que le smartphone est responsable de l'assistance à l'orientation, la canne collecte des données et gère l'assistance à la *marche* (sous-fonction de la mobilité). Le SEES permet ainsi à une personne de connaître sa position actuelle. Cette position est également accessible à d'autres personnes, des tiers de confiance (famille ou amis par exemple), qui peuvent ainsi vérifier que les déplacements de l'utilisateur du SEES se font en sécurité.

Un point clé du SEES est sa connexion à Internet, qui permet aux personnes malvoyantes ou aveugles d'obtenir aide et support depuis n'importe quel lieu du monde (disposant d'une couverture réseau). Cette caractéristique fait du SEES une canne intelligente ubiquitaire [Yusro et al., 2013a].

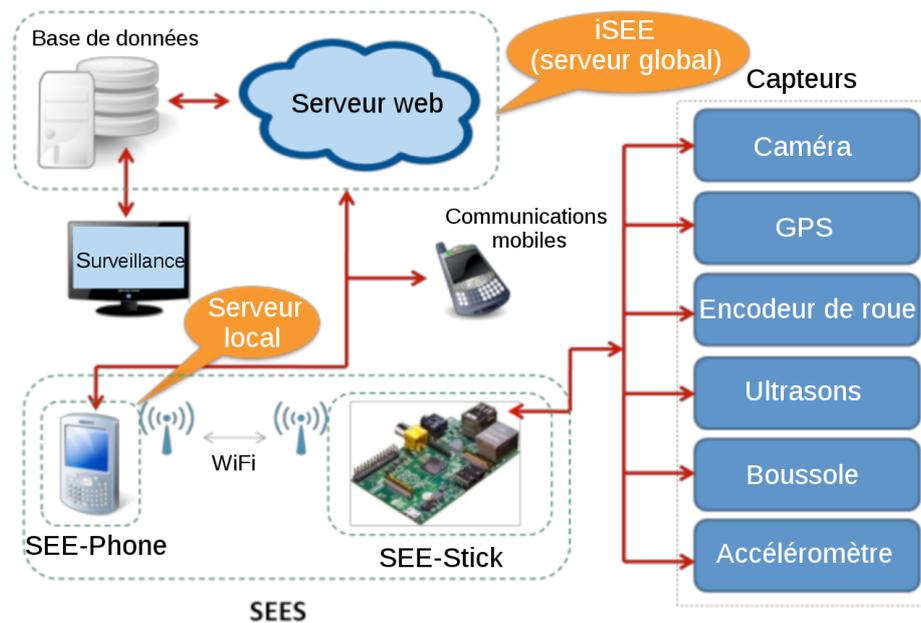


FIGURE 3.1 – L’architecture globale du projet SEES.

Une autre concept central du SEES est *l’active multi-sensor context awareness* [Yusro et al., 2013a]. Ce concept signifie la chose suivante : le système embarque plusieurs capteurs, dont certains sous-ensembles sont destinés à acquérir des informations cohérentes (p. ex. : capteurs de distance) et d’autres, des informations différentes (p. ex. : capteurs de localisation et capteurs d’obstacles). En fonction de l’environnement, certains capteurs sont plus nécessaires que d’autres, soit parce que la situation demande une grande précision pour une mesure particulière (par exemple si le système doute de la présence d’un obstacle), soit parce que certains capteurs sont inadaptés (comme des capteurs infrarouges en plein soleil), soit parce que certains capteurs sont à même, seuls, de fournir des informations de qualité suffisante (comme un récepteur GPS en terrain découvert). Le système doit être capable d’estimer cette pertinence des capteurs face au contexte : il va alors les allumer ou les éteindre dynamiquement selon la progression de la situation. Pourvu que l’ensemble des capteurs embarqués soit adapté, ce principe donne de la flexibilité au fonctionnement de la canne : on peut réduire la consommation d’énergie du système en éteignant des capteurs dans les cas propices, et, à l’inverse, dans les cas difficiles, on peut augmenter la précision en les allumant.

La maîtrise de la consommation d’énergie est une contrainte importante pour tout système embarqué, car l’autonomie en est directement dépendante. L’autre alternative pour augmenter cette dernière<sup>3</sup> est l’augmentation de la capacité de la batterie, qui implique des augmentations de l’encombrement et du poids du dispositif défavorables à l’utilisation du dispositif. Étant donnée l’importance de l’autonomie, il est donc nécessaire de faire baisser la consommation d’énergie, et le concept *d’ac-*

3. À l’exclusion des mécanismes de collecte d’énergie, qui sont difficiles à mettre en œuvre, et n’annulent d’ailleurs pas la contrainte de frugalité énergétique.

*tive multi-sensor context awareness* est particulièrement adapté à cet objectif. En effet, certains capteurs, en particuliers optiques, ont une consommation très élevée, soit directement comme les capteurs infrarouges qui émettent de la lumière, soit indirectement comme les caméras qui nécessitent une importante puissance de calcul (consommatrice) pour exploiter leurs images, voire même les deux, comme les caméras 3D à lumière structurée.

### 3.1.2 Plate-forme

#### Matériel

Certains des composants du système SEES sont des éléments standards fabriqués en grande série (p. ex. : smartphone), et d'autres sont gérés par des acteurs externes (p. ex. : plate-forme serveur *cloud*). *A contrario*, la canne intelligente SEE-stick est un équipement spécifique : un prototype de canne intelligente a dû être conçu et fabriqué afin de pouvoir conduire des expérimentations. Ce prototype était destiné à embarquer des capteurs de natures adéquates et en nombre suffisant pour tester les fonctionnalités souhaitées – principalement la navigation et la détection des obstacles, ainsi que les composants de collecte de données, de calcul et de connectivité nécessaires pour traiter et transmettre les données des capteurs dans le cadre de l'architecture choisie. Pour simplifier la conception mécanique, le choix a été fait de baser le prototype sur une plate-forme pré-existante : une poussette pour enfant en bas âge. Malgré son encombrement important et son incongruité pour les observateurs extérieurs, le choix d'une poussette comme support apporte des avantages. D'une part, spacieuse et robuste, une poussette peut facilement accueillir un grand nombre de capteurs, cartes électroniques, batteries, et est capable de supporter un ordinateur portable, ce qui est très confortable lors des tests. D'autre part, une poussette possède des roues, auxquelles on peut fixer un encodeur angulaire, capteur très utile car permettant de connaître la distance parcourue par la personne avec une très bonne précision et une excellente robustesse – si l'état du sol le permet, ce qui était le cas dans les expérimentations avec le SEES. Le prototype avait donc la forme d'une poussette instrumentée (figure 3.2), équipée des capteurs suivants :

- Huit capteurs de distance ultrasoniques MaxBotix EZ-0 [Max, 2015], disposés en arc de cercle à l'avant de la poussette pour détecter les obstacles. Ils sont capables de détecter et de mesurer la distance d'obstacles situés jusqu'à environ 6,45 m, avec une résolution de 1 pouce (environ 2,54 cm). Afin d'éviter les interférences entre les signaux des différents capteurs, ils sont câblés de manière à ordonnancer leurs temps d'émissions ;
- Une puce ST Microelectronics LSM303DLHC contenant un accéléromètre trois axes et un magnétomètre trois axes [ST, 2013]. L'accéléromètre a une plage de mesure sélectionnable entre  $\pm 2$  g,  $\pm 4$  g,  $\pm 8$  g, ou  $\pm 16$  g (la résolution diminuant avec chaque élargissement de l'étendue), et celle du magnétomètre est au choix de  $\pm 1,3$  gauss à  $\pm 8,1$  gauss avec 5 paliers intermédiaires ;

- Une puce ST Microelectronics LSM330DLC contenant un accéléromètre trois axes et un gyromètre trois axes [ST, 2012]. L'accéléromètre est similaire à celui embarqué dans la puce LSM303DLHC, et le gyromètre a une plage de mesure sélectionnable :  $\pm 250$  degrés/s,  $\pm 500$  degrés/s, ou  $\pm 1000$  degrés/s ;
- Un encodeur optique AEDR-8300 fixé sur la roue arrière droite de la poussette [Ava, 2007] ;
- Un GPS Fastrax UP501 [Fas, 2010] ;
- Une webcam HD Logitech C615<sup>4</sup>.

Les cartes électroniques suivantes sont présentes :

- Le cœur du SEE-stick est un Raspberry Pi modèle B révision 2<sup>5</sup>. Il embarque un *system-on-a-chip* (SoC) Broadcom BCM2835, puce réunissant un microprocesseur (un ARM11 à 700 MHz) et des périphériques nécessaires au fonctionnement du système : ports série, gestionnaire de mémoire de stockage, entrées-sorties programmables (en anglais *General Purpose I/O* ou *GPIO*), processeur graphique, processeur audio, etc. Ce Raspberry Pi est connecté au réseau WiFi par une carte WiFi IEEE 802.11n Alfa AWUS036NEH<sup>6</sup> connectée à un de ses ports USB ;
- Un autre Raspberry Pi modèle B révision 1.2 a été périodiquement utilisé pour les expérimentations de traitement d'image ;
- Une carte *iLive* accompagnée d'une carte d'extension dédiée est branchée sur le connecteur d'extension du Raspberry Pi principal ; elle fait l'interface entre les capteurs ultrasoniques et ce dernier, et embarque les deux puces ST Microelectronics formant la centrale inertielle. La carte *iLive* a été développée par l'équipe SMIR en tant que plate-forme pour la recherche sur les réseaux de capteurs sans fil. Elle est dotée de capteurs (inertiels, ainsi que luminosité et température qui ne sont pas utilisés par le système SEES) et de capacités de connectivité sans fil par le protocole IEEE 802.15.4<sup>7</sup>. La carte *iLive* est connectée au Raspberry Pi par une liaison série.

Deux périphériques de sortie destinés à la transmission d'informations aux utilisateurs sont également présents : un moteur vibrant Parallax 28821<sup>8</sup> pour signaler les obstacles et une prise jack 3.5 mm sur laquelle un haut parleur ou des écouteurs peuvent être branchés pour une sortie en synthèse vocale. L'ensemble des éléments

---

4. La version a changé (numéro de pièce différent), mais une page produit est disponible sur le site de Logitech au moment où ces lignes sont écrites : <https://www.logitech.fr/fr-fr/product/hd-webcam-c615> (accès le 17 septembre 2019).

5. Site de la fondation Raspberry Pi : <https://www.raspberrypi.org/>. De la documentation sur le modèle *1 B v1.2* en particulier est disponible sur le site *Electronic Designers in the South East of England* : <https://raspberrypi-projects.com/pi/pi-hardware/raspberry-pi-model-b/hardware-general-specifications> et sur Wikipédia : [https://fr.wikipedia.org/wiki/Raspberry\\_Pi#Mod%C3%A8le\\_1\\_B](https://fr.wikipedia.org/wiki/Raspberry_Pi#Mod%C3%A8le_1_B) (pour les trois URL, accès le 17 septembre 2019).

6. Informations disponibles à l'adresse [https://www.alfa.com.tw/products\\_detail/10.htm](https://www.alfa.com.tw/products_detail/10.htm) (accès le 17 septembre 2019).

7. Informations disponibles sur le site du groupe de travail de l'IEEE : <http://www.ieee802.org/15/pub/TG4.html> (accès le 17 septembre 2019).

8. Informations disponible à l'adresse <https://parallax.com/product/28821> (accès le 17 septembre 2019).

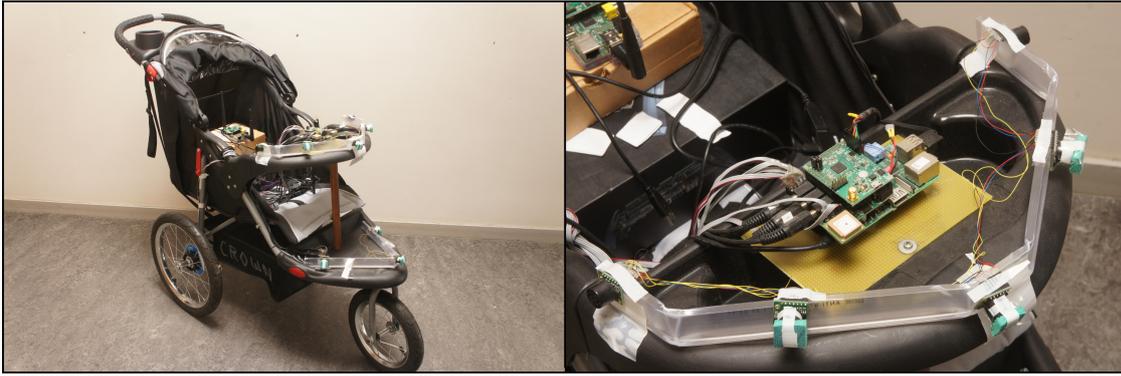


FIGURE 3.2 – Le prototype de SEE-stick. À gauche, une vue d’ensemble du prototype, à droite, une vue plus détaillée sur la carte à laquelle sont connectés la plupart des capteurs.

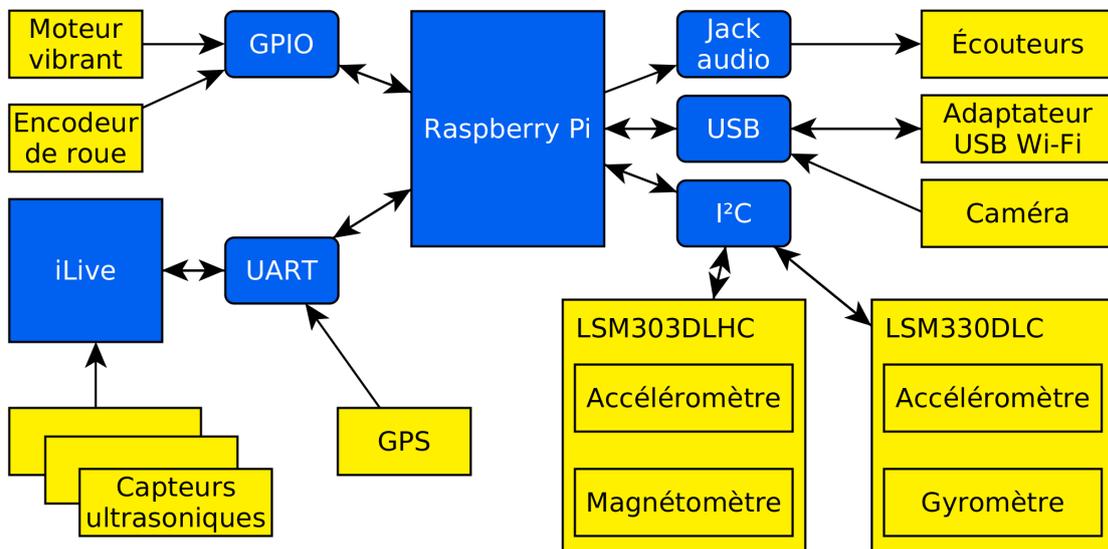


FIGURE 3.3 – Composants électroniques constitutifs du prototype de SEE-stick et leurs interconnexions.

et de leurs connexions est représenté sur la figure 3.3, et leur répartition sur la plate-forme du SEE-stick est visible sur la figure 3.4.

### Logiciel

Le système d’exploitation fonctionnant sur le Raspberry Pi principal est une distribution GNU/Linux Raspbian basée sur Debian 7.1 pour architecture ARM 32 bits, avec le noyau Linux en version 3.10.25. Le programme qui gère les données est écrit en C++ ; il détecte les obstacles tout en calculant les déplacements de la poussette et envoie les données de position au logiciel serveur. Ce dernier est installé sur un ordinateur standard ; il est composé d’un site web écrit en HTML, CSS, et d’un *back-end* écrit en PHP avec une base de données MySQL.

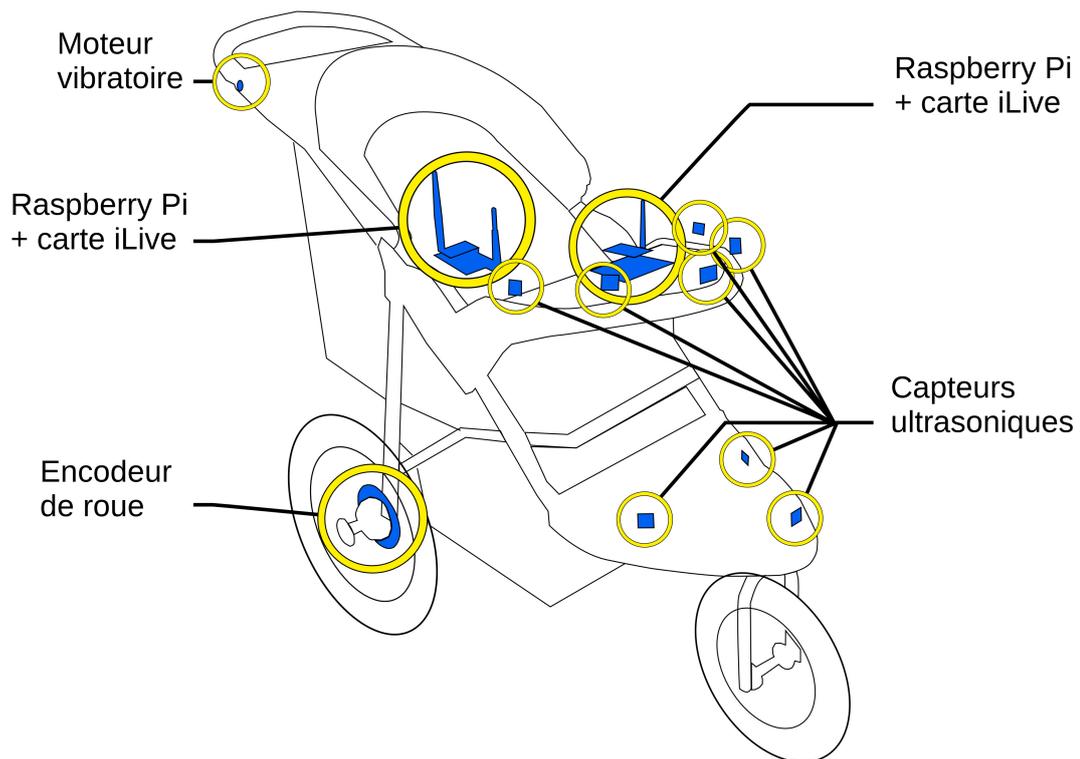


FIGURE 3.4 – Répartition des composants électroniques sur la plate-forme mécanique du prototype de SEE-stick.

### 3.1.3 Fonctionnalités et expérimentations

Plusieurs fonctionnalités ont été développées et testées sur le prototype de SEES.

#### Détection d'obstacles

Une fonction de détection des obstacles statiques a été implémentée sur le prototype [Yusro et al., 2013a]. Pour ce faire, six capteurs ultrasoniques ont été utilisés pour détecter les obstacles dans deux gammes de distances : de 0 m à 1 m et de 1 m à 2 m. Quand un obstacle est détecté, le moteur vibrant est utilisé pour signaler sa présence à l'utilisateur, avec deux fréquences différentes selon la distance approximative (proche ou lointaine). Cette fonctionnalité a été testée avec succès à des distances différentes (50 cm à 150 cm) et avec des obstacles de hauteurs différentes (40 cm, 100 cm et 180 cm).

#### Navigation en extérieur

Yusro a développé une fonction de navigation en extérieur reposant sur le GPS du smartphone (figure 3.5). Cette fonction a pour objectifs de localiser et suivre la personne en temps réel, et de fournir l'information à une tierce personne à partir du serveur [Yusro, 2017].



FIGURE 3.5 – Trajet en extérieur réalisé avec le SEES. L'image est tirée de l'interface web du système [Yusro, 2017].

### Navigation en intérieur

La fonctionnalité de navigation en intérieur est cruciale pour le système SEES. En effet, si les solutions de navigation en extérieur sont nombreuses, relativement matures et partiellement satisfaisantes (grâce aux systèmes basés sur les GNSS : Trekker Breeze, applications smartphone comme Blindsquare), la situation est toute autre en intérieur, où le problème est encore loin d'être résolu, y compris technologiquement [Lymberopoulos et al., 2015].

Une première méthode, simple, de localisation a été développée : elle repose sur le principe de navigation à l'estime (en anglais, *dead-reckoning*) avec les données de l'encodeur de roue et du gyromètre. Pour calculer la position d'un objet à un instant donné, il suffit de connaître sa position à l'instant précédent dans le temps et le déplacement qu'il a effectué pendant le laps de temps qui sépare les deux instants. Dans notre cas, on part de l'hypothèse qu'on connaît la position initiale de la canne ; reste à calculer son déplacement à tout instant ultérieur. Ce déplacement total est égal à l'intégrale des déplacements infinitésimaux – un déplacement étant le produit d'une vitesse et d'une durée (figure 3.6 et équation 3.1, où  $\vec{r}(t)$  désigne la position à un instant  $t$ ,  $d\vec{r}$  un déplacement infinitésimal et  $\Delta r$  une variation de déplacement petite mais non infinitésimale). Par définition, il est impossible en pratique de mesurer des quantités infinitésimales et la navigation à l'estime est donc une approximation de cette intégrale : c'est la somme des déplacements discrets (équation 3.2). On cherchera à minimiser la durée  $\Delta t$  de mesure des déplacements  $\Delta r$  pour approcher au mieux l'intégrale.

$$\vec{r}_{t_b} = \vec{r}_{t_a} + \int_{t=t_a}^{t_b} d\vec{r} \quad (3.1)$$

$$\vec{r}_{t_b} \approx \vec{r}_{t_a} + \sum_{t=t_a}^{t_b} \Delta \vec{r} \quad (3.2)$$

Il est donc nécessaire de mesurer les déplacements ; de nombreuses méthodes et capteurs existent. Dans le cas du SEES, le choix de M. Yusro *et al.* a été d'utiliser

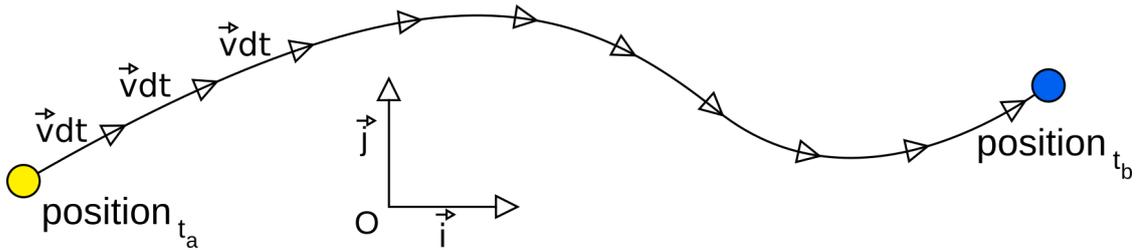


FIGURE 3.6 – Fonctionnement de la navigation à l'estime : une intégration de déplacements infinitésimaux permet de calculer un déplacement.

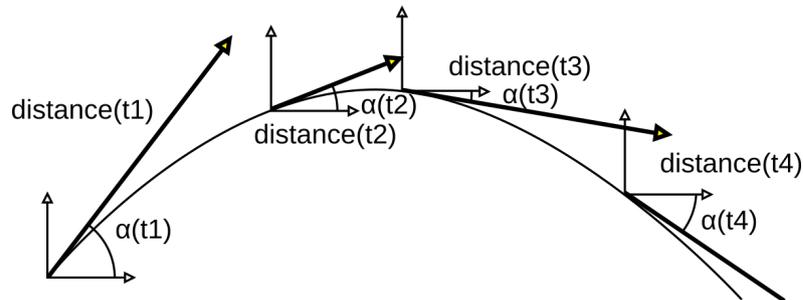


FIGURE 3.7 – La reconstruction de la trajectoire avec les données de distance parcourue et de d'orientation (notée  $\alpha$  sur la figure).

l'encodeur de roue pour mesurer les distances parcourues et le gyromètre pour les angles. On mesure donc régulièrement le mouvement rotatif de la roue instrumentée et en le multipliant par la circonférence de la roue (on néglige l'écrasement du pneu), ainsi que la vitesse angulaire, qui est intégrée pour obtenir l'angle de rotation. Ces deux données permettent de reconstruire la trajectoire (figure 3.7).

Au fur et à mesure de l'avancée du processus de navigation à l'estime, la trajectoire estimée *dérive* : l'erreur (c'est-à-dire l'écart entre la position réelle et la position estimée) augmente progressivement jusqu'à – si rien n'est fait – rendre le résultat de l'estimation inutile. Si l'erreur d'estimation augmente, c'est à cause du principe de fonctionnement même de la navigation à l'estime :

- La somme discrète n'est qu'une approximation imparfaite de l'intégration ; pour qu'elle soit correcte, il faudrait que l'intervalle de temps entre deux échantillonnages soit infinitésimal. À cause de cela, on a une perte d'information sur la trajectoire : on ne connaît la vitesse et l'orientation qu'à un nombre limité d'instants, alors que ces grandeurs varient continûment ;
- Dans les faits, ce n'est pas une sommation de *déplacements* qui est réalisée mais bien une sommation *d'estimations* de déplacements réalisées à partir des données capteurs. Ce processus va également additionner les erreurs sur chaque estimation de déplacement, qui vont alimenter la dérive.

La majeure partie des erreurs était due aux données des capteurs. Deux corrections ont donc été mises en place pour les limiter afin de conserver une estimation valable plus longtemps. Le premier problème corrigé était l'absence de prise en compte de l'orientation du gyromètre. Originellement, le gyromètre était – consciemment ou inconsciemment – supposé horizontal, et les données alors logiquement échantillon-

nées uniquement sur l'axe  $z$  (qui était donc supposé vertical). L'orientation n'était pas en réalité précisément contrôlée, la puce étant soudée sur une carte, qui était connectée à une autre carte, elle-même fixée à une partie de la poussette dont la position n'était pas connue et même, à cause de vibrations, variable. Le problème des vibrations a été atténué avec une rigidification du support de la poussette par l'ajout d'un tasseau en bois, tandis que celui de la position « stable » l'a été par une phase initiale de calibration au démarrage du système. Cette phase de calibration exploite l'accéléromètre présent dans la même puce que le gyromètre : au repos, celui-ci mesure uniquement l'effet de la gravité, dont le vecteur est, par définition, vertical. On peut alors mesurer la direction de la verticale du repère externe (celui de la pièce) dans le repère interne (celui de la puce), et ensuite utiliser cette information pour corriger les valeurs lues sur l'axe vertical du gyromètre. Si on suppose que le sol est horizontal, la correction est valable.

La seconde correction, qui vise à atténuer l'erreur du gyromètre, est également effectuée en phase de calibration. On peut modéliser la sortie réelle du gyromètre comme étant ce qu'il mesure, soit la vitesse angulaire sur ses trois axes, additionnée d'une erreur elle-même décomposable en sous-erreurs : une partie est constante, une partie dépend de la température, une autre de la vitesse angulaire, etc. Dans notre cas, nous avons cherché à corriger le biais statique, c'est-à-dire l'erreur constante qui s'additionne en permanence au signal du gyroscope. Pendant cette phase de repos, il est très probable qu'une majeure partie de l'erreur présente soit le biais statique ; elle est donc mesurée et enregistrée. Par la suite, elle est soustraite au signal du gyromètre. Cette correction est valable si l'hypothèse faite dans le modèle d'erreur est correcte.

Les trajectoires corrigées étaient acceptables sur une plus longue durée et une plus grande distance (figure 3.8), mais la dérive finissait par dépasser notre seuil de précision.

### Vision par ordinateur

M. Yusro a développé une application de vision sur le SEE-phone qui assure l'assistance dans deux tâches complémentaires à la navigation pure et la détection générale d'obstacles : la détection de l'état des feux tricolores et la détection des poteaux [Yusro, 2017]. La première fonctionnalité est évidemment importante pour la mobilité, puisque les traversées de routes représentent des étapes complexes et dangereuses pour les personnes déficientes visuelles ; on a vu dans l'état-de-l'art plusieurs dispositifs d'assistance à leur traversée (comme les balises sonores et le système Crosswatch [Alexandre et al., 2018, Coughlan and Shen, 2013]). Quand aux poteaux, ce sont des obstacles fréquents et difficiles à détecter par les capteurs de distance classiques (infrarouges, ultrasoniques) en raison de la finesse de leur profil. La figure 3.9 représente le fonctionnement global de l'application.

La détection de l'état des feux tricolores fonctionne par filtrage teinte/saturation/luminosité, avec un seuil bas et un seuil haut sur chaque composante définissant un sous-ensemble de l'espace pour chacune des trois couleurs que peut prendre un feu tricolore (figure 3.10). La détection des poteaux est plus sophistiquée. Plusieurs

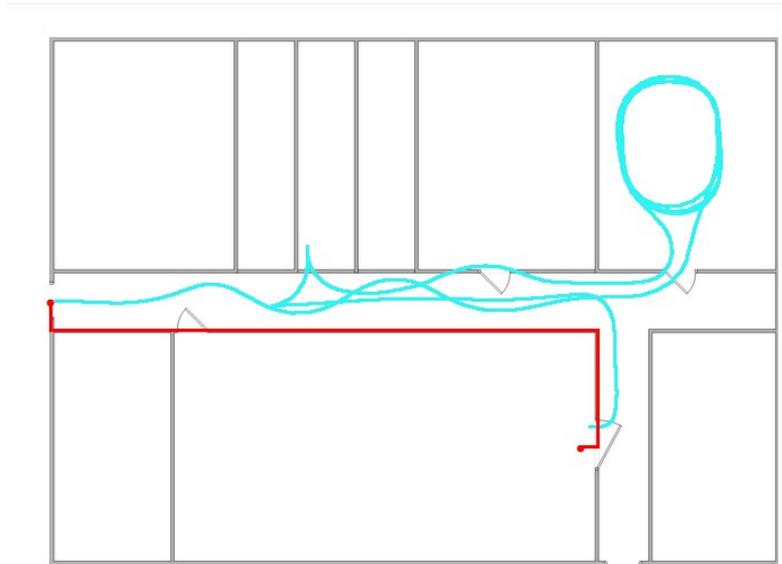


FIGURE 3.8 – Résultat du calcul d'un parcours en intérieur visant à mettre en évidence la dérive en effectuant plusieurs tours à l'intérieur d'une pièce : l'estimation est assez bonne.

étapes sont nécessaires (figure 3.11). Par rapport à ses concurrents, l'algorithme a l'avantage de limiter le nombre de faux positifs sans augmenter le nombre de faux négatifs.

Lin a développé une application de détection des portes et des poignées et plaques de portes (qui portent l'identification de la pièce et les noms de ses occupants) capable de fonctionner en temps réel sur Raspberry Pi [Lin et al., 2015]. La détection de porte est assurée par deux méthodes successives pour accélérer le processus de détection et limiter la consommation énergétique du SEES. Dans un premier temps, l'image est réduite en résolution (de 640 x 480 pixels à 40 x 30 pixels) puis filtrée par couleur, la couleur des portes étant, dans l'environnement de test, approximativement constante. Le filtrage est une isolation des pixels appartenant à un sous-ensemble de l'espace teinte/saturation/luminosité. La taille des zones de la bonne couleur est ensuite filtrée : au-dessus d'un seuil haut pré-défini, la zone est considérée comme étant une porte ; en-dessous d'un seuil bas pré-défini, elle est considérée comme n'en étant pas une ; entre les deux seuils, elle est analysée plus finement. Cette deuxième méthode d'analyse est géométrique : les lignes de la porte sont isolées avec un filtre de Canny et leur orientation avec une transformée de Hough. Si les lignes forment un rectangle, la zone est considérée comme étant une porte, sinon, comme n'en étant pas une. La détection des poignées et des plaques de portes n'est déclenchée, toujours pour des raisons d'économie, qu'une fois la personne arrivée à destination. De plus, elle est optimisée de la même manière que la détection de portes, avec une première phase de filtrage par seuillage de couleurs et de taille de zone dans l'image. Une fois que des zones candidates ont été repérées, elles sont analysées avec un détecteur de caractéristiques visuelles *Scale-Invariant Feature Transform* (SIFT)<sup>9</sup> et une base de données de caractéristiques visuelles

9. Traduit sur Wikipédia par « transformation de caractéristiques visuelles invariante à

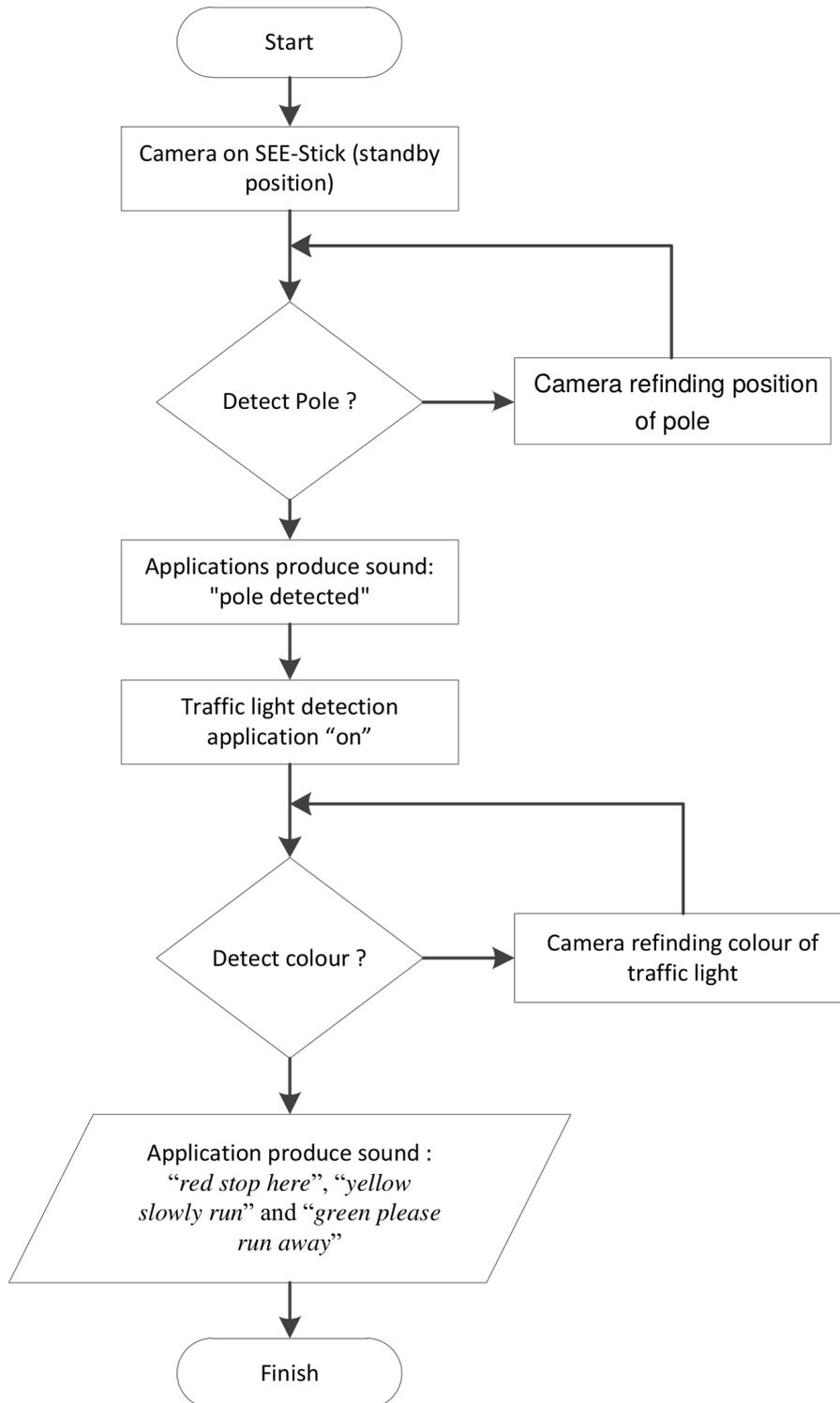


FIGURE 3.9 – Fonctionnement global de l’application de détection visuelle de l’état des feux et de poteaux ( [Yusro, 2017]).

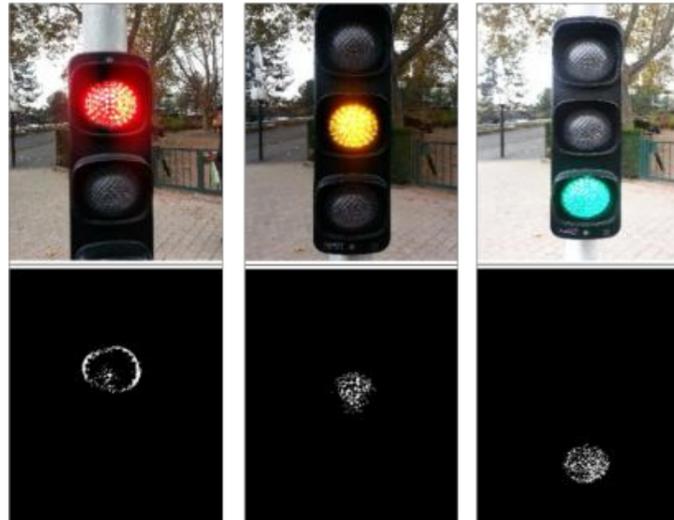


FIGURE 3.10 – Détection de l'état d'un feu tricolore ([Yusro, 2017]).

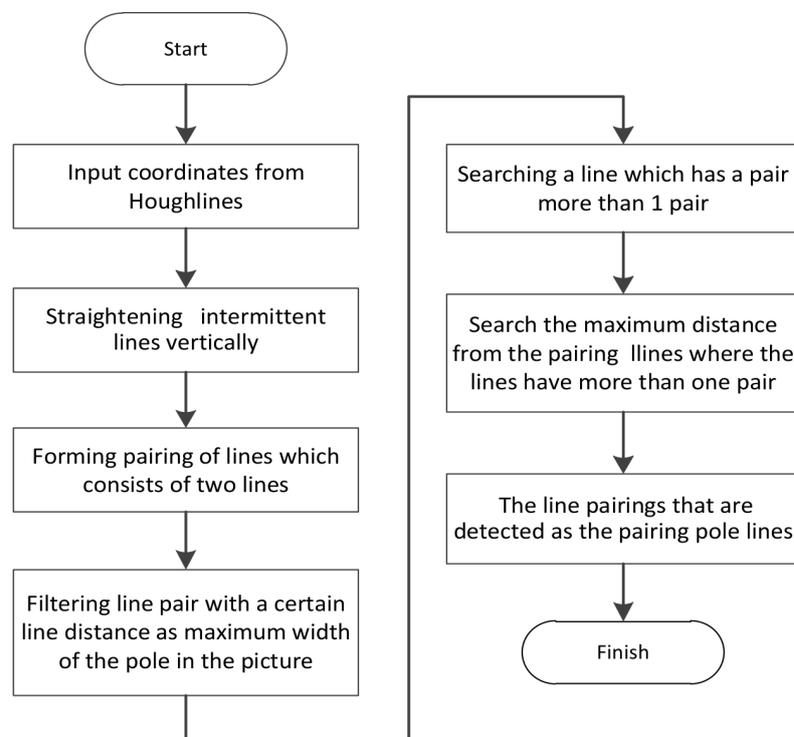


FIGURE 3.11 – Fonctionnement de l'algorithme de détection de poteaux (figure de [Yusro, 2017]).

pré-enregistrées.

## 3.2 Système 2SEES et architecture

Le système 2SEES est au cœur des travaux effectués pendant cette thèse. Il est le prolongement du système SEES, dont il conserve les orientations générales de son prédécesseur. Il repose sur la même architecture à trois tiers : canne intelligente, smartphone et *back-end* dans le *cloud*. Les trois rôles restent globalement inchangés :

1. La canne intelligente, qui se nomme désormais *Local MIMO Smart Cane* (LMSC)<sup>10</sup>, est chargée de collecter la plupart des données ;
2. Le smartphone sert de passerelle entre le LMSC et Internet, en particulier le *back-end*. Il sert également d'interface et est localement responsable de la gestion des informations de haut niveau : cartographie, informations sémantiques sur l'environnement, etc.
3. Le *back-end* stocke toutes les cartes nécessaires, et est l'interface avec les acteurs externes au système, comme, par exemple, les gestionnaires de *back-ends* d'objets intelligents.

Les fonctionnalités sont organisées en trois groupes : détection d'obstacles, navigation et assistance à la perception environnementale. Les deux premières sont le prolongement des travaux du SEES et, plus généralement, des autres dispositifs d'assistance à la mobilité. Le troisième groupe, l'assistance à la perception environnementale, comprend les fonctionnalités destinées à fournir des informations pertinentes sur l'environnement, non limitées à l'assistance à l'orientation ou à la marche.

### Sûreté de fonctionnement

Le principal point de différenciation et de progrès dans le passage du SEES au 2SEES est la prise en compte explicite des problématiques de sûreté de fonctionnement dans la conception du système.

La sûreté de fonctionnement (*dependability* en anglais) représente une des cinq propriétés fondamentales des systèmes, les quatre autres étant les suivantes : fonctionnalité, utilisabilité, performance, coût [Avižienis et al., 2001]. Elle est définie par Avižienis, Laprie et Randell : « *Dependability is the ability to deliver service that can justifiably be trusted.* » [Avižienis et al., 2001], ce qu'on peut traduire par : « La sûreté de fonctionnement est la capacité d'un système informatique à fournir un service dans lequel les utilisateurs peuvent avoir, à raison, confiance. » .

Penchons-nous sur cette notion dans notre contexte. Tout d'abord, le terme *système informatique* pourrait paraître légèrement déplacé : les dispositifs d'assistance

---

l'échelle » ([https://fr.wikipedia.org/wiki/Scale-invariant\\_feature\\_transform](https://fr.wikipedia.org/wiki/Scale-invariant_feature_transform) (accès le 19 septembre 2019)).

10. Ce qu'on pourrait traduire par « canne intelligente locale à plusieurs entrées et plusieurs sorties ».

sont plus que des systèmes informatiques classiques en raison de l'importance primordiale de leurs interactions avec l'environnement d'un côté et les utilisateurs de l'autre. Nous pensons cependant qu'elle est tout à fait pertinente car, techniquement, ce sont des systèmes de traitement de l'information. L'article d'Avizienis, Laprie et Randell donne d'ailleurs des exemples voisins, comme les systèmes de contrôle de vol embarqués dans les avions. Cette définition met l'accent sur le mot confiance : la finalité de la sûreté de fonctionnement est la confiance que doivent avoir les utilisateurs dans le service fourni. Or, c'est fortement le cas des dispositifs d'assistance, qui sont destinés à supporter une part importante de la mobilité de leurs utilisateurs. Pour une entité endossant un tel rôle de support, la confiance est essentielle. Le terme « à raison » (*justifiably*) impose que la cause de cette confiance trouve ou puisse trouver source dans les qualités du système lui-même, et non pas dans des externalités indépendantes de ces qualités (croyances, marketing, etc.). Le *service* que fournit l'application est défini comme étant le comportement du système *tel que perçu par ses utilisateurs*. Dans le cadre générique de l'article, les utilisateurs sont les autres systèmes, humains ou artificiels. Dans notre cas, les premiers utilisateurs sont les personnes déficientes visuelles.

Les concepts fondamentaux de la sûreté de fonctionnement sont de trois natures : les menaces (*threats*), les composantes (*attributes*) et les moyens (*means*).

Les menaces sont structurées de la manière suivante :

1. Les *défaillances* sont des disruptions du service fourni par le système. Elles sont provoquées par des erreurs affectant le service ;
2. Les *erreurs* sont les états imprévus du système, et sont potentiellement sources de défaillance ;
3. Les *fautes* sont des problèmes de conception ou d'implémentation, qui peuvent générer des erreurs. Elles sont dites actives si elles produisent une erreur, ou dormantes dans le cas contraire ;

Le programme 3.1 est un exemple de distinction entre fautes, erreur et défaillances. Il contient une *faute* causée par la personne qui l'a développé : la fonction `memcpy()` est utilisée avec `strlen()` comme paramètre de longueur sur la chaîne à copier. Or, en langage C, pour copier une chaîne de caractère, il ne faut pas uniquement copier ses caractères mais aussi le symbole de terminaison de chaîne `'\0'` (numériquement égal à 0). Sans ce symbole, une suite de caractères ne constitue pas une chaîne de caractères valide. Puisque `strlen()` renvoie le nombre de caractères que contient la variable `nom`, le symbole de terminaison n'est pas copié dans ce programme. La situation dépend ensuite de l'état du tableau `sortie` au moment de son allocation : s'il contient des zéros, ce qui est possible mais pas garanti, alors le premier zéro suivant les caractères sera interprété comme un symbole de fin de chaîne `'\0'` par les fonctions manipulant la variable, et tout fonctionnera bien. Si, au contraire, le huitième élément du tableau ne contient pas un zéro mais une autre valeur, alors `sortie` ne sera pas une chaîne de caractères valide ; ce sera une *erreur* dans le programme. Cette erreur pourra alors causer une *défaillance* du programme, soit un comportement en-dehors des spécifications. Si la chaîne n'est pas affichée, ou bien seulement traitée par des fonctions qui ne tiennent pas compte de la présence ou non du symbole de fin de chaîne, il n'y aura pas de défaillance. Si par contre

on utilise `sortie` avec une fonction qui utilise le symbole de terminaison, comme `printf("%s\n", sortie)` dans le programme 3.1, alors une défaillance se produira. Cette défaillance peut prendre la forme d'une sortie erronée : si le huitième élément de sortie était initialisé à 'o' et le neuvième à zéro, le programme affichera « Bernardo ». On peut également avoir un crash simple du programme : si la fonction `printf()` parcourt le tableau sans rencontrer de zéro jusqu'à sortir de la zone mémoire autorisée, on aura une erreur de segmentation. Selon la destination de l'application, l'une ou bien l'autre de ces défaillances peut avoir des conséquences graves. Il peut aussi se produire une défaillance plus générale : une personne mal intentionnée peut exploiter ce type de faille pour manipuler des données auxquelles elle n'est pas censée avoir accès.

```

1 #include <stdio.h>
2 #include <string.h>
3
4
5 int main(void)
6 {
7     char nom[16] = "Bernard";
8     char sortie[16] = "";
9
10    memcpy(sortie, nom, strlen(nom));
11
12    printf("%s\n", sortie);
13
14    return 0;
15 }

```

Listing 3.1 – Programme d'exemple des menaces sur la sûreté de fonctionnement.

Le programme 3.1 met donc en évidence la chaîne des menaces : une faute peut provoquer une erreur, qui elle-même peut provoquer une défaillance du système. Pour les éviter, il est nécessaire de connaître leurs raisons, c'est-à-dire d'étudier les erreurs, fautes, et défaillances possibles ainsi que les liens causaux qui les relient. Des caractérisations générales ont été développées et facilitent l'analyse des menaces qui doit se faire dans le cadre particulier de chaque système.

Les *composantes* (*attributes* en anglais) du système sont ses caractéristiques du point de vue de la sûreté de fonctionnement. Elles sont au nombre de six [Avižienis et al., 2001] :

- La *disponibilité* est la propriété de pouvoir fournir le service aux utilisateurs au moment où ils le désirent ;
- La *fiabilité* est la continuité du service correct, c'est-à-dire sans défaillance ;
- La *maintenabilité* est la propriété de pouvoir être réparé et modifié.
- La *confidentialité* est l'absence de divulgation d'information hors des spécifications ;
- La *sécurité* est la propriété liée à l'absence de conséquences néfastes pour l'environnement et les utilisateurs ;
- L'*intégrité* est l'absence de modifications du système hors des spécifications ;

Les trois dernières propriétés peuvent être regroupées dans une méta-propriété *sécurité*.

Une autre composante dérivée nous intéresse particulièrement : la robustesse, c'est-à-dire la sûreté de fonctionnement en présence d'entrées erronées [Avižienis et al., 2001]. Dans notre cas, les entrées erronées sont principalement constituées par les données des capteurs qui peuvent, dans certaines les situations, ne pas refléter fidèlement la réalité.

Afin d'atteindre la sûreté de fonctionnement d'un système, ou tout du moins l'améliorer, quatre *moyens* sont disponibles [Avižienis et al., 2001] : prévention des fautes, tolérance aux fautes, élimination des fautes et prévision des fautes. La prévention des fautes vise à éviter l'apparition de fautes, autrement dit, à augmenter la « qualité » du système avec, par exemple, des procédures de construction rigoureuses, des marges de tolérance importantes pour le matériel, etc. Les mesures passives de défense contre les fautes mal intentionnées font partie de cette catégorie. La tolérance aux fautes vise à permettre au système de continuer à fonctionner correctement en présence de fautes, autrement dit à éviter les défaillances malgré la présence de fautes. Pour cela, il faut que les erreurs soient détectées et identifiées, puis neutralisées. L'élimination des fautes consiste à analyser le système, depuis les spécifications et la conception jusqu'en phase d'exploitation, pour y trouver les fautes inactives et les éliminer. La validation des spécifications, les tests et l'injection volontaire de fautes, et la maintenance sont des techniques d'élimination des fautes. La prévision des fautes, enfin, consiste à analyser le système pour y identifier les fautes potentielles, leurs probabilités d'apparition, et leurs conséquences. Les analyses de risques et les statistiques de défaillances sont des techniques de prévision des fautes.

La redondance est un outil privilégié pour mettre en place la tolérance aux pannes [Arlat et al., 2006]. La méthode la plus directe est la redondance spatiale, c'est-à-dire la réplique des unités de calcul. Si deux unités de calcul effectuant le même traitement sur les mêmes données fournissent des résultats différents, une erreur est présente [Arlat et al., 2006]. Selon la configuration des unités, la gamme des erreurs détectables varie :

- Si les deux unités sont identiques, on peut détecter les erreurs dues à des fautes physiques internes ;
- Si le matériel et le logiciel des unités sont différents, on peut également détecter les erreurs dues aux fautes dans la conception de ces éléments.

Si trois répliques des unités fonctionnent en parallèle, on peut leur adjoindre une unité qui donne la valeur majoritaire fournie par les unités (vote). Ainsi, cela permet de directement avoir la valeur correcte en sortie sous l'hypothèse d'une seule défaillance parmi les trois unités. Cette technique n'est pas applicable pour beaucoup de systèmes embarqués, y compris le nôtre : en effet, la consommation d'énergie entraînée par la duplication ou le triplement des composants fonctionnant en parallèle de manière permanente est trop élevée.

La redondance temporelle est un autre type de redondance : lorsqu'une erreur est détectée, le système revient dans un état valide précédent (reprise). Cela nécessite un moyen de détection d'erreur et une sauvegarde des états. Cette technique a un

surcoût en consommation énergétique bien moins élevé.

Nous avons analysé globalement le 2SEES pour trouver les points les plus sensibles. Cette analyse n'était pas dirigée par les fonctionnalités du 2SEES, elle était moins précise : elle observe les points de rupture possibles dans le système. Ces points de rupture sont classifiés sur deux aspects. D'un côté, leur place dans le système : détection et acquisition des informations de l'environnement local, traitement de ces informations, et communications entre les composants du système et les systèmes externes. De l'autre côté, des critères de niveaux correspondant grossièrement à la classification de Marr des niveaux d'analyse de la perception [Marr, 1982] : niveau matériel, niveau implémentation et niveau informationnel ou computationnel. Le tableau 3.1 présente dans chacune de ses cellules l'élément du système qui est susceptible de provoquer une défaillance du système complet.

Le niveau matériel est le plus évident : si un composant a une défaillance et qu'il est essentiel au service, alors le service est interrompu. Cela est vrai pour tout type de composant : capteur, processeur, et canal de communication – émetteur-récepteur ou bien ligne de communication si le canal n'est pas sans fil.

Le second niveau est un peu moins évident. Il correspond aux implémentations logicielles qui fonctionnent sur les composants cités plus haut : programmes d'acquisition et de traitement des informations échantillonnées par les capteurs, pile de protocoles communication. Au même titre que les éléments matériels, ce sont des composants du 2SEES : si l'un d'entre eux subit une défaillance critique, le système dans sa globalité ne fonctionnera plus. Comme plus haut, l'atténuation du risque de panne complète passe par leur duplication : il faudrait en théorie deux implémentations de chaque algorithme de traitement de l'information et deux piles de communication. Cette méthode de renforcement peut paraître exagérée mais constitue une option réelle, voire un passage obligé, pour améliorer la sûreté de fonctionnement des systèmes critiques ; elle est notamment mentionnée par Avižienis, Laprie et Randell [Avižienis et al., 2001].

Le troisième niveau est le niveau computationnel ou informationnel. S'agissant des moyens de communication, il est un peu flou, mais on peut imaginer des modes de communication radicalement différents : la dépendance ou indépendance du système au réseau Internet ou pas représente par exemple une grande divergence dans les possibilités fonctionnelles et la robustesse du système. Le niveau computationnel–informationnel est également un peu flou pour le traitement des données, mais, encore une fois, on peut chercher des algorithmes s'appuyant sur des principes très différents : localisation absolue et navigation à l'estime, par exemple. Là où ce niveau d'analyse est le plus intéressant, c'est dans le domaine de la détection. En effet, dans chaque catégorie d'information, certains types de capteurs sont bien adaptés à certaines configurations de l'environnement, et moins à d'autres. Cela est lié au mode de fonctionnement même des capteurs (cf. section 2.2.1). Pour augmenter la robustesse du système, il est alors bénéfique de panacher les types de capteurs et les modes de détection afin de couvrir le plus de contextes possibles. La construction d'un système 2SEES tolérant aux fautes passe par une phase de sécurisation de tous ces aspects.

TABLE 3.1 – Matrice des sources potentielles de défaillance.

Niveau	Rôle dans le système		
	Détection	Traitement	Communication
Matériel	Un seul capteur	Un seul processeur	Une seule interface
Implémentation logicielle	Une seule implémentation	Un seul programme/fonction	Une seule pile de protocoles
Computationnel ou informationnel	Un seul type de capteur	Un seul algorithme	Un seul mode de communication

### 3.3 Conception et réalisation de la plate-forme matérielle

Cette section présente les travaux de conception, de fabrication et d'amélioration de la plate-forme qui sert de support matériel au LMSC. La première étape de la conception mécanique du LMSC a été définir son type. Ce choix était en effet structurant pour la suite de la conception. La conception à proprement parler et la fabrication de la plate-forme ont ensuite suivi. En parallèle, la plate-forme électronique a été conçue en suivant le concept du 2SEES, avec pour objectif d'augmenter la tolérance aux fautes.

#### 3.3.1 Forme et conséquences

Le LMSC fait suite au SEE-stick, et nos travaux sur et avec ce dernier nous ont fourni des informations précieuses sur la forme que peuvent prendre les dispositifs d'assistance. Le SEE-stick était basé sur une poussette, ce qui lui donnait une forme inhabituelle : l'utilisateur *déplaçait* un dispositif roulant devant lui. Les dispositifs vus au chapitre 2 qui se rapprochent le plus de ce format sont les robots d'assistance et la *GuideCane*. Les robots sont en fait assez différents, puisqu'ils se meuvent seuls et n'impliquent d'effort physique de poussée de la part de la personne. La *GuideCane* est plus semblable à la poussette, mais elle aussi présente une différence importante, puisqu'elle est active : ses roues sont motorisées et évitent les obstacles d'elles-mêmes. La fonction essentielle de la poussée est de signaler qu'on veut se déplacer, et dans quelle direction ; la *GuideCane* pourrait très bien assister la personne en traction de manière à faire disparaître les efforts de poussée. La poussette a des défauts : un encombrement important causant des difficultés dans les lieux exigus (qui sont gérables si on est voyant mais probablement pas pour des personnes aveugles), l'impossibilité de monter des escaliers, et, bien sûr, son incongruité. Avec ces défauts, ce format devient finalement un handicap supplémentaire pour la personne, et n'est pas envisageable en utilisation réelle. C'est acceptable dans un contexte d'expérimentations bien que cela éloigne les possibilités d'avoir un ressenti général la forme et la manipulation, qui sont d'une grande importance.

Pour le LMSC, nous avons cherché un intermédiaire entre la poussette et la canne blanche classique ou les boîtiers portables (« télécommandes ») qui permettent une utilisation légère et intuitive. Le format « canne à roue(s) » représente un compromis entre ces deux extrêmes (tableau 3.2) : comme pour la poussette, le poids n'est pas critique car supporté par la (ou les) roue(s), mais avec un encombrement plus limité et une transportabilité plus proche de ce que peut offrir une canne classique. Elle permet d'embarquer un encodeur de roue, et, de manière générale, plus de capteurs qu'une canne simple, sur toute sa longueur car il n'y a pas de porte-à-faux ; on peut donc placer des capteurs au plus près du sol. En revanche, on ne peut pas s'en servir en balayage, mais sa moindre mobilité apporte en contrepartie une plus grande stabilité aux capteurs, dont les positions et orientations par rapport à la personne sont ainsi connues. Enfin, elle permet d'emporter un encodeur de roue. Nous avons donc choisi ce format pour le LMSC, qui possède la flexibilité nécessaire à un pré-prototype destiné à l'expérimentation sans être aussi incongru que le format original du SEE-stick.

TABLE 3.2 – Comparaison des qualités respectives des dispositifs d'assistance à la mobilité des personnes malvoyantes.

	Canne instrumentée	Boîtier portable	Poussette	Canne à roues
Poids	Critique	Important	Peu important	Moyennement important
Perception directe	Possible	Possible	Très difficile	Difficile
Stabilité	Faible	Faible	Excellente	Bonne
Nombre de capteurs	Limité	Limité	Illimité	Moyen
Escaliers	Oui	Oui	Non	Moyens
Manœuvrabilité	Bonne	Bonne	Faible	Moyenne

### 3.3.2 Conception de la plate-forme mécanique

La plate-forme mécanique est composée d'un bloc support de roue – une seule roue dans un premier temps, pour des raisons de simplicité – que nous nommons *tête de canne*, et d'un manche. Ces deux éléments peuvent être instrumentés, mais la plupart des capteurs sont destinés à être portés par la tête de canne. La fonction de support de la roue est la plus critique, car soumise à des contraintes de dimensionnement mécanique avec des tolérances relativement faibles, contrairement à la fixation des capteurs qui est simple à mettre en œuvre. Ces contraintes proviennent du capteur servant à mesurer les rotations de la roue : nous avons fait le choix d'un encodeur rotatif magnétique, car ce type de capteur est insensible à la plupart des perturbations environnementales, contrairement au capteur optique qui équipait le SEE-stick et qui est sensible aux poussières, salissures, etc. En contrepartie, l'encodeur magnétique nécessite un positionnement précis : à la fois pour l'alignement entre le capteur et l'aimant qui doit être fixé sur l'axe, et pour l'écartement entre

### 3.3. CONCEPTION ET RÉALISATION DE LA PLATE-FORME MATÉRIELLE99

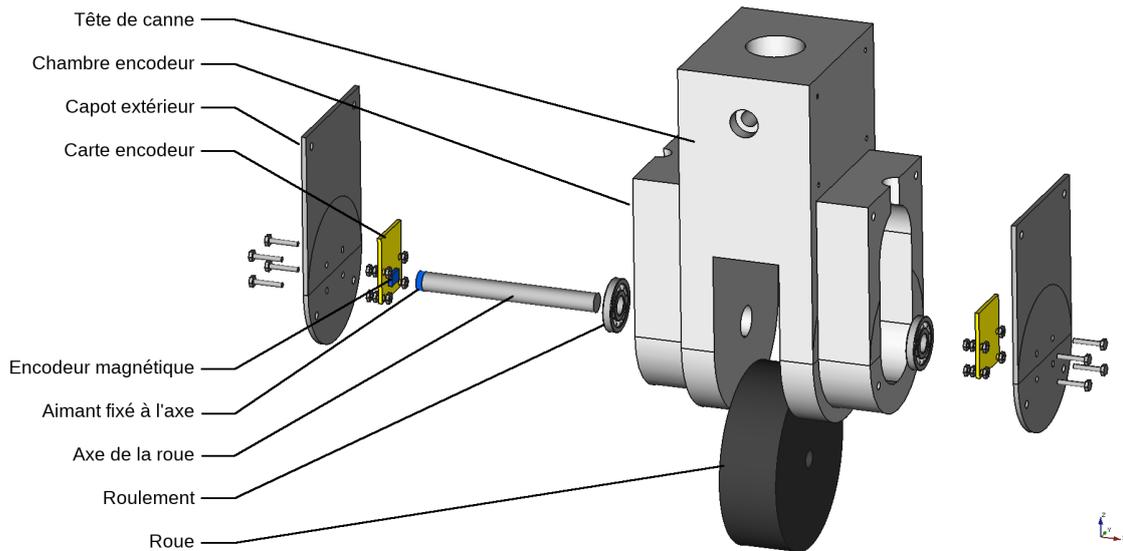


FIGURE 3.12 – Vue 3D éclatée de la tête de canne du LMSC.

ces deux éléments, qui doit être faible. Ces critères imposent des contraintes à la tête de canne, dont les dimensions dépendent également du diamètre de la roue, qui doit être assez grande pour pouvoir rouler sur un terrain irrégulier tout en limitant le poids et l'encombrement du dispositif.

La tête de canne a été réalisé d'une seule pièce, en ABS imprimé en 3D (figures 3.12 et 3.13).

La fixation des cartes électroniques sur la tête de canne n'a pas posé de problèmes particuliers. Des plaques en aluminium ont été fixés par vis auto-taraudeuses dans des alésages laissés dans le plastique de la tête de canne. C'est le cas des capots qui ferment les logements d'encodeurs magnétiques et des plaques supports des cartes électroniques. Ces dernières ont été fixées avec des entretoises pour laisser de l'espace pour les têtes de vis M2,5 qui servent à fixer les cartes (figures 3.12 et 3.13). L'aluminium a été choisi pour la facilité avec laquelle il peut être travaillé (pliage, perçage), et pour sa légèreté par rapport à l'acier.

La tête de la canne étant destinée à accueillir trois groupes de capteurs d'obstacles : un pour détecter les obstacles avec emprise au sol et de taille moyenne, un pour détecter les obstacles au ras du sol (p. ex. : bordures de trottoirs), et un troisième pour détecter les obstacles en hauteur<sup>11</sup>. Il était nécessaire que tous ces capteurs gardent une orientation constante afin de toujours scruter l'espace là où les obstacles sont susceptibles de se trouver (figure 3.14). Ce n'est pas l'utilisateur qui scanne l'environnement, mais la canne qui contrôle en permanence la présence d'obstacles. Or, l'orientation des capteurs n'est pas garantie car la canne possède une liberté de rotation autour de l'axe de sa roue : la personne qui tient la canne peut l'incliner plus ou moins vers le haut ou vers elle. La première solution pour compenser ce degré de liberté pour les groupes de capteurs d'obstacles a été l'uti-

11. Pour simplifier le câblage, ce groupe de capteurs destiné aux obstacles en hauteur n'a pas été installé sur le prototype.

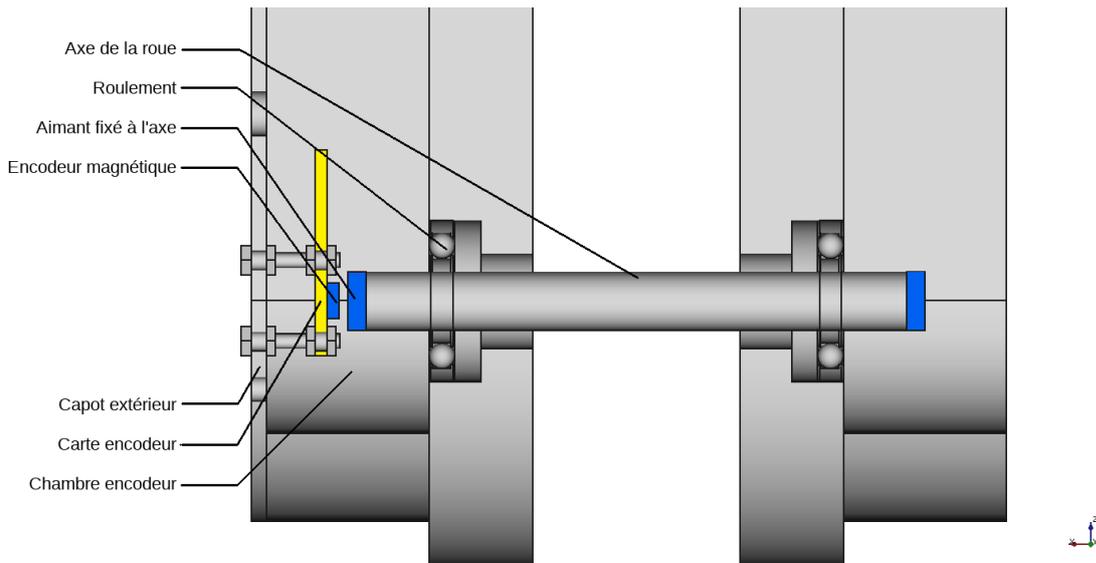


FIGURE 3.13 – Détail du positionnement de l’encodeur magnétique et de l’axe de roue.

lisation d’un balancier. Au lieu d’être directement fixés sur la canne, les capteurs ont été fixés à un anneau entourant la canne et pivotant librement autour d’un axe horizontal transversal. Un contrepoids fixé à l’arrière assurait l’équilibrage par rapport aux capteurs fixés à l’avant (figure 3.15). Cette solution a été abandonnée car la fréquence de résonance du balancier était proche de la fréquence à laquelle le manche était naturellement balancé, générant des oscillations importantes et annulant le bénéfice escompté.

Il aurait été possible de remplacer ce balancier passif par un système actif : un inclinomètre ou une centrale inertielle mesurant l’angle du balancier, et un contrôleur estimant la réponse à appliquer pour maintenir le balancier horizontal et commandant un actionneur (moteur, piston, etc.) en conséquence. Un système de ce type est néanmoins beaucoup plus complexe à mettre en place qu’un simple balancier.

La solution choisie a été, au contraire, tournée vers la simplification : plutôt que de supprimer le degré de liberté seulement pour les capteurs en les découplant de la tête de canne, nous avons choisi de maintenir le couplage et de supprimer le degré de liberté sur l’ensemble de la canne. Pour ce faire, des roues supplémentaires ont été installées à l’arrière de la roue principale. Elles sont fixées sur une structure en aluminium découpé et plié (figure 3.16). Quand les trois roues touchent le sol, les capteurs sont orientés correctement. Cette configuration a deux avantages supplémentaires : le contact au sol sur trois points permet à la canne de se tenir seule « debout », et la structure en aluminium offre un espace supplémentaire, non prévu initialement, pour la fixation de nouvelles cartes électroniques, elles-mêmes non prévues à l’origine.

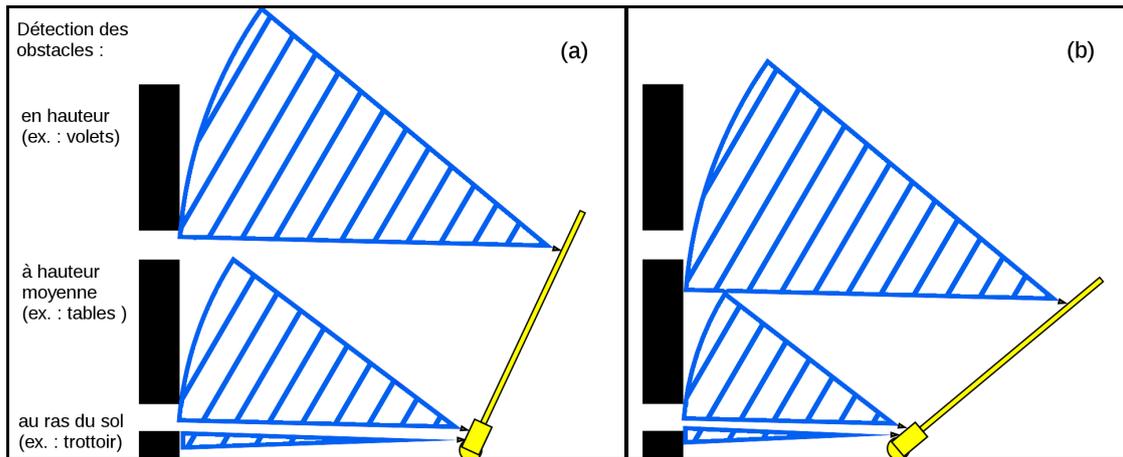


FIGURE 3.14 – Fonctionnement des trois groupes de capteurs d’obstacles du LMSC en fonction de l’inclinaison de la canne : (a) normale, (b) forte.

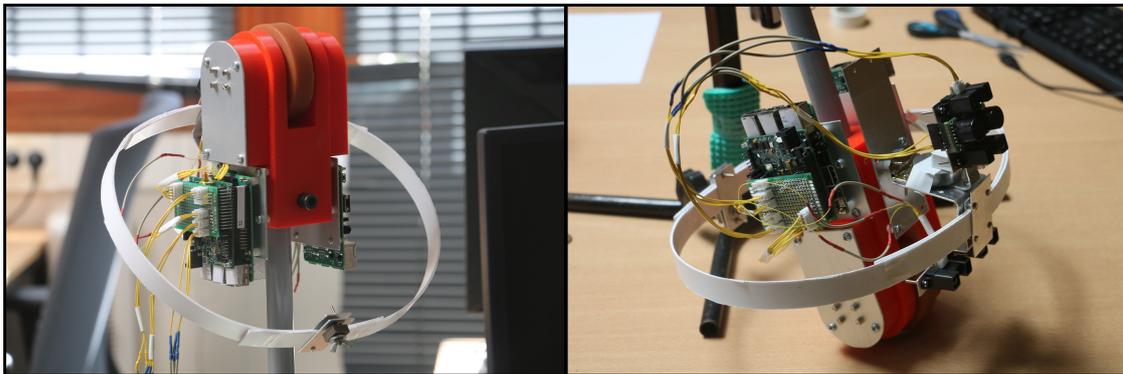


FIGURE 3.15 – Anneau de stabilisation des capteurs d’obstacles.

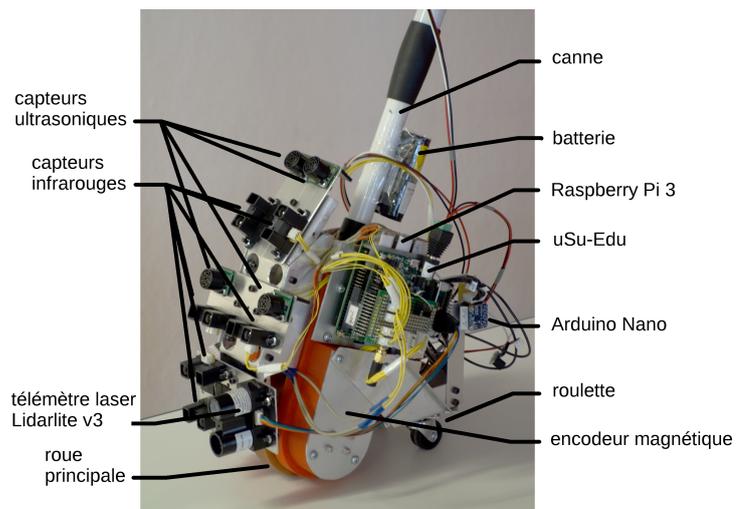


FIGURE 3.16 – Deuxième version du LMSC, avec support de capteurs fixe et capteurs et roues additionnelles.

### 3.3.3 Conception de la plate-forme électronique

Le travail principal de conception de la plate-forme électronique a porté sur la réalisation des objectifs du concept 2SEES : la robustesse et la fiabilité. Pour ce faire, nous avons commencé par examiner l'architecture du SEES (figure 3.3). La nouvelle architecture devait être plus tolérante aux pannes et plus flexible en termes de capteurs.

Les différentes architectures présentées dans cette section sont destinées à implémenter des mécanismes de redondance temporelle plutôt que spatiale (cf. section 3.2).

Le LMSC reprend dans les grandes lignes les caractéristiques du SEE-stick dont il est la continuité. Il est donc le principal support des capteurs qui fournissent les données du système 2SEES<sup>12</sup>.

Pour la fonctionnalité de détection d'obstacles, ces capteurs sont les suivants :

- Les capteurs ultrasoniques Easybotix MaxSonar EZ déjà utilisés sur le SEE-stick ont été conservés [Max, 2015]. Ils consomment environ 2 mA et peuvent être alimentés avec une tension comprise entre 2,5 V et 5,5 V. La canne en utilise deux paires : une paire redondante qui observe les obstacles de hauteur moyenne, et l'autre qui observe les obstacles en hauteur ;
- Six capteurs infrarouges Sharp GP2Y0A02YK sont utilisés en complément des capteurs ultrasonique [Sha]. Ils ont une portée de 150 cm (en intérieur, pas en plein soleil), un faisceau de détection plus étroit que les capteurs ultrasoniques mais une consommation importante (33 mA à 50 mA sur une alimentation 5 V). Ils sont répartis en trois paires redondantes : deux qui accompagnent les deux paires de capteurs ultrasoniques et une autre pour les obstacles au sol ;
- Un capteur laser Garmin Lidar Lite v3 surveille les obstacles au sol [Gar, 2016]. Il a une portée de 40 m, un « cône » de détection d'ouverture nulle (c'est un laser ponctuel) et une précision de  $\pm 2,5$  cm en-dessous de 5 m. Sa consommation est élevée, 105 mA au repos et 135 mA pendant les mesures, pour une alimentation en 5 V ;
- Deux capteurs laser ST Microelectronics VL53L0X ont été positionnés sur les côtés de la canne afin de mesurer la distance aux murs [ST, 2016]. Bien qu'étant également des capteurs lasers, ceux-ci sont très différents du Lidar Lite v3 : ils ont une portée de 2 m environ en intérieur, un cône de détection large (25 °), et une consommation d'environ 20 mA pour une tension de 3 V.

Les capteurs pour la fonctionnalité de navigation sont les suivants :

- La rotation de la roue est mesurée par un encodeur magnétique AMS AS5040 [ams, 2015] ;
- Les accéléromètres et les gyromètres de la puce LSM330DLC [ST, 2012].

---

12. Principal mais pas unique : les capteurs embarqués dans le smartphone (p. ex. : GPS, caméra) peuvent également être utilisés.

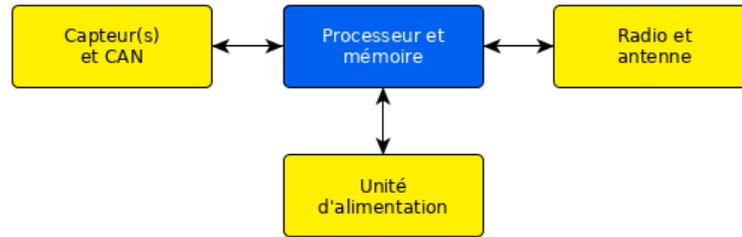


FIGURE 3.17 – Architecture de nœud de réseau de capteurs à simple cœur.

### Architecture

Le LMSC est destiné à collecter des informations sur son environnement, à en traiter une partie et à transmettre le reste au smartphone et au cloud, en consommant le moins d'énergie possible. En cela, c'est un appareil très semblable aux nœuds de réseaux de capteurs sans fil : nous nous sommes donc inspirés de leur architecture pour concevoir celle du LMSC.

La plupart de ces nœuds sont généralement composés de quatre unités : des capteurs accompagnés de leur électronique dédiée, un cœur de calcul, une radio pour l'émission et la réception de données, et une unité d'alimentation (figure 3.17) [Akyildiz et al., 2002]. Les capteurs et leur électronique collectent des données, leur appliquent des pré-traitements simples (p. ex. : une amplification), et convertissent les données analogiques en numérique (CAN) pour leur exploitation par le calculateur. Ce dernier est chargé des traitements complexes sur les données et de faire fonctionner le système d'exploitation, nécessaire passé un certain niveau de complexité, car il simplifie la gestion du matériel, du développement d'applications et des fonctionnalités avancées. La radio connecte le nœud au réseau et l'unité d'alimentation fournit l'énergie nécessaire au fonctionnement du capteur. Cette dernière est essentielle, car elle détermine l'autonomie du système.

Ces architectures à simple cœur sont amenées à être progressivement remplacées par des appareils multi-cœur, car les premiers ne sont pas assez performants [Hamdi et al., 2008, Akyildiz et al., 2008]. En effet, les réseaux de capteurs sans fil utilisent de plus en plus de capteurs par nœud et ces capteurs sont de plus en plus évolués. Des capteurs multimédia comme les caméras sont, par exemple, souvent utilisés. D'autre part, la quantité et la complexité des données imposent des traitements de plus en plus sophistiqués. Enfin, le traitement doit également se faire de plus en plus rapidement, souvent en temps réel. L'architecture d'un nœud multi-cœur est fondamentalement proche de celle d'un nœud simple cœur (figure 3.18) [Munir et al., 2015]. Le LMSC est concerné par tous ces aspects : il doit embarquer des capteurs divers pour ses différentes fonctionnalités, il est susceptible d'embarquer une ou des caméras, et les traitements doivent être dans leur grande majorité faits en temps réel. Nous considérerons dans la suite indistinctement les architectures multi-cœur à proprement parler, dans lesquelles plusieurs cœurs sont réunis dans une même puce, et multi-processeur, dans lesquelles plusieurs puces de calcul sont utilisées.

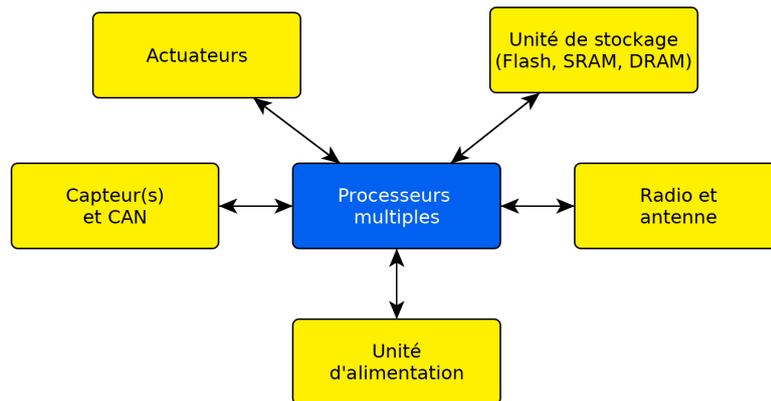


FIGURE 3.18 – Architecture de nœud de réseau de capteurs multi-cœur.

L'intérêt principal d'une architecture multi-cœur est sa flexibilité : le même système peut exécuter des applications demandant un haut niveau de performance tout en économisant de l'énergie [Shi et al., 2011, 2013, Munir et al., 2014]. D'une part, parce que plusieurs cœurs fonctionnant en parallèle consomment moins d'énergie qu'un seul cœur équivalent [Kleihorst et al., 2006], et, d'autre part, parce qu'il est possible d'éteindre ou de mettre en sommeil les cœurs ou les puces non utilisés. Avec la puissance additionnelle, on peut envisager de déporter certains calculs depuis le *back-end* réseau vers le nœud si le rapport entre l'énergie de la transmission économisée et l'énergie du calcul dépensée est favorable. Cela permet aussi, tout simplement, la conservation d'une fonctionnalité en cas d'absence de connectivité réseau, que ce soit à cause d'une panne ou d'une simple absence de couverture. Ces deux cas sont très intéressants pour le LMSC, en raison de l'objectif d'amélioration de la sûreté de fonctionnement : cette architecture confère au système 2SEES une meilleure robustesse dans son fonctionnement. Des cœurs multiples ont d'autres atouts pour la tolérance aux fautes : chaque cœur peut surveiller le fonctionnement correct des autres cœurs (calcul dupliqué, par exemple), et si un cœur tombe en panne, un autre peut prendre le relai. Nous avons donc choisi une architecture multi-cœur pour le LMSC.

Le premier projet d'architecture est visible sur la figure 3.19, qui montre à la fois les processeurs et les tâches qui leur sont affectées. Elle a été conçue pour pouvoir travailler sur la tolérance aux fautes, la puissance et l'autonomie. De nombreux processeurs sont impliqués dans cette architecture : sept microcontrôleurs à très faible consommation et deux microprocesseurs (en principe des Raspberry Pi), plus puissants mais plus consommateurs. L'architecture n'est pas symétrique : la plupart des processeurs y tiennent un rôle particulier – sauf certains qui sont en double pour la tolérance aux fautes. Ces processeurs et leurs rôles sont les suivants :

1. Un groupe de deux microcontrôleurs (MCU 0 sur la figure 3.19) est chargé de la navigation à l'estime. Deux tâches y sont exécutées : l'échantillonnage des capteurs liés – encodeur de roue et gyromètre – et l'intégration numérique de la navigation à l'estime ;
2. Un autre groupe de deux microcontrôleurs occupe un rôle similaire, mais pour la fonctionnalité de détection d'obstacles : ils échantillonnent les cap-

- teurs d'obstacles et évaluent si un obstacle est présent ;
3. Un microcontrôleur, que nous nommons « contrôleur » est chargé de contrôler l'état de tous les autres, en vue de détecter les pannes éventuelles. Cette opération est réalisée par un mécanisme de battements de cœur : chaque puce exécute une tâche consistant à envoyer régulièrement un signal simple (ou battement de cœur) au processeur de contrôle. Quand ce dernier ne reçoit plus le battement de cœur d'une puce, il la redémarre ;
  4. Certaines tâches fonctionnent grâce aux données sauvegardées ; c'est en particulier le cas de la navigation à l'estime, qui construit la trajectoire pas à pas. Pour éviter qu'une panne et un redémarrage n'entraîne la perte de ces données, un processeur est dédié à leur stockage. Quand le contrôleur redémarre un autre processeur, il peut ainsi lui envoyer les données de reprise ;
  5. Un microcontrôleur gère l'interfaçage avec l'extérieur du LMSC : le smartphone et l'utilisateur. Il ordonnance et transmet les messages venant des autres processeurs selon leur priorité. Il est directement connecté à une interface locale (moteur vibrant) pour les messages les plus urgents à transmettre à l'utilisateur, c'est-à-dire les détections d'obstacles.
  6. Enfin, une paire de processeurs puissants – *a priori*, des Raspberry Pi – sont dédiés aux tâches gourmandes en puissance de calcul. Dans le cadre du LMSC, ce type de tâche doit être exécuté périodiquement : en effet, la navigation à l'estime doit être couplée à une méthode de localisation absolue pour être effective. En intérieur et en l'absence d'infrastructure spécifique, la méthode de localisation absolue la plus accessible est, *a priori*, la vision par ordinateur, à savoir la reconnaissance d'éléments de l'environnement qui sont cartographiés. Les Raspberry Pi sont donc affectés à l'exécution de cette tâche gourmande en puissance.

Ce premier projet d'architecture plaçait la redondance à l'intérieur même des fonctionnalités : la détection d'obstacles reposait sur deux puces, et la localisation utilisait deux microcontrôleurs (pour la navigation à l'estime) et deux Raspberry Pi (pour la localisation absolue). Elle nécessitait un circuit imprimé évolué, car destiné à connecter de nombreuses puces entre elles. À cause de ce haut niveau de répartition, le système d'exploitation aurait été nécessairement complexe, et les applications difficiles à développer et surtout à déboguer.

Dans un second temps, l'architecture du LMSC s'est orientée vers la symétrie. Au contraire, la seconde ébauche place la redondance à un plus haut niveau : c'est un système fonctionnel complet, à même de fonctionner seul, qui a été répliqué, ce qui a permis de simplifier grandement la réalisation. Cette version symétrique simplifie la fabrication car des cartes pré-existantes peuvent être réutilisées : comme pour le SEES, les cartes iLive sont adaptées à cela (figure 3.20). Le système d'exploitation peut être globalement centré sur une seule unité complète avec un module de gestion de la tolérance aux fautes, plutôt qu'être intrinsèquement parallèle. Les applications sont également beaucoup plus faciles à développer : elles peuvent fonctionner de la même manière que celles du SEES. En contrepartie, cette architecture est moins puissante. Tout d'abord, elle est moins efficace : en cas de panne d'un seul composant d'une unité complète, il est nécessaire de basculer tout le processus sur l'autre unité. En d'autres termes, une seule panne met hors de service

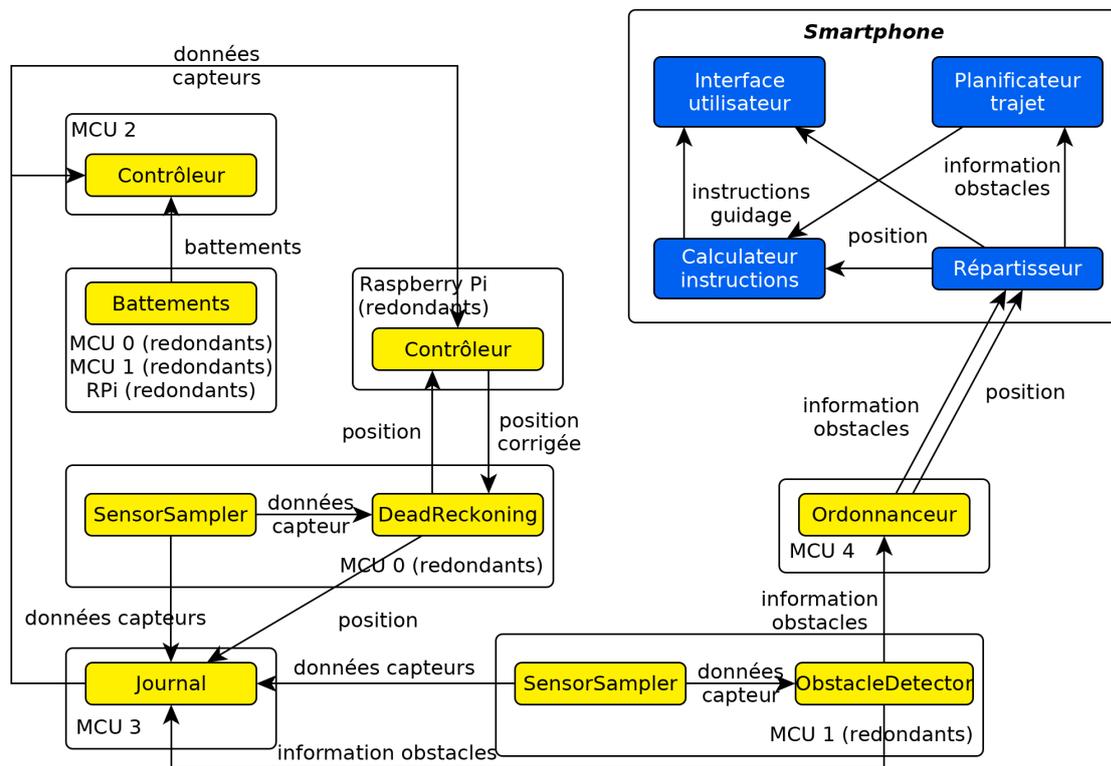


FIGURE 3.19 – Architecture multi-cœurs prévue initialement pour le LMSC, avec l'affectation des tâches pour chaque microcontrôleur (MCU sur la figure) et pour les Raspberry Pi.

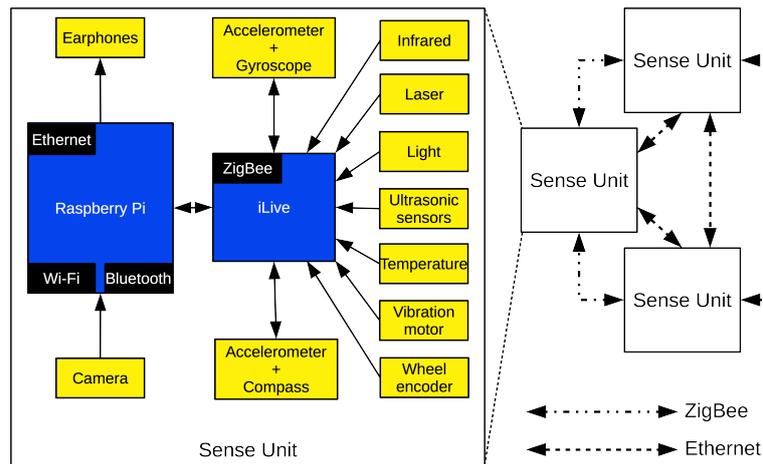


FIGURE 3.20 – Architecture redondante symétrique.

un groupe complet de composants, là où la première architecture peut tolérer des pannes sur plusieurs composants sans souffrir d’une diminution de fonctionnalité, car les composants affectés à des fonctionnalités différentes ne sont pas liés.

Au final, l’architecture choisie et implémentée était proche de la seconde (symétrique) présentée ci-dessus. Les différences principales sont le choix de deux unités répliquées plutôt que trois, le remplacement des cartes iLive par des cartes *uSu-Edu*, et l’ajout de deux Arduino Nano<sup>13</sup> (figure 3.23). La carte *uSu-Edu* (figure 3.21) a été développée par l’équipe SMIR pour être une plate-forme flexible pour l’expérimentation sur les réseaux de capteurs sans fil et les objets intelligents. Elle a été notamment utilisée dans le projet européen Captor<sup>14</sup>. Une carte *uSu-Edu* comprend suffisamment de composants pour être utilisable directement, ne nécessitant qu’une source d’énergie et une application pour être opérationnelle. Elle comprend les éléments suivants :

- Un microcontrôleur Atmel Atmega 128RFA2 AVR 8 bits fonctionnant à 8 MHz ou 16 MHz à faible consommation qui inclut notamment une radio 802.15.4 pour les communications sans fil [Atm, 2014] ;
- Un étage d’alimentation flexible et robuste, capable de fournir de l’énergie à partir d’une large gamme de tensions en entrée, ce qui permet l’utilisation de nombreuses sources d’énergie : batteries lithium-ion, piles alcalines, alimentations secteur, panneaux solaires, etc. ;
- Des capteurs directement embarqués sur la carte : un LSM303DLHC contenant un accéléromètre trois axes, et un magnétomètre trois axes et un LSM330DLC contenant un accéléromètre trois axes et un gyromètre trois axes [ST, 2013, 2012], un capteur de luminosité ambiante Vishay TEPT4400 [Vis, 2014], un thermomètre NXP LM75B [NXP, 2015], et un capteur de pression et de température Bosch Sensortec BMP180 [Bos, 2015]. Ces capteurs permettent de mesurer trois paramètres contextuels et les mouvements de la carte ;

13. Site de la fondation Arduino : <https://www.arduino.cc/> (accès le 3 octobre 2019)

14. Rapport final accessible à l’adresse suivante : [https://www.captor-project.eu/wp-content/uploads/2018/02/D4.5\\_final.pdf](https://www.captor-project.eu/wp-content/uploads/2018/02/D4.5_final.pdf)

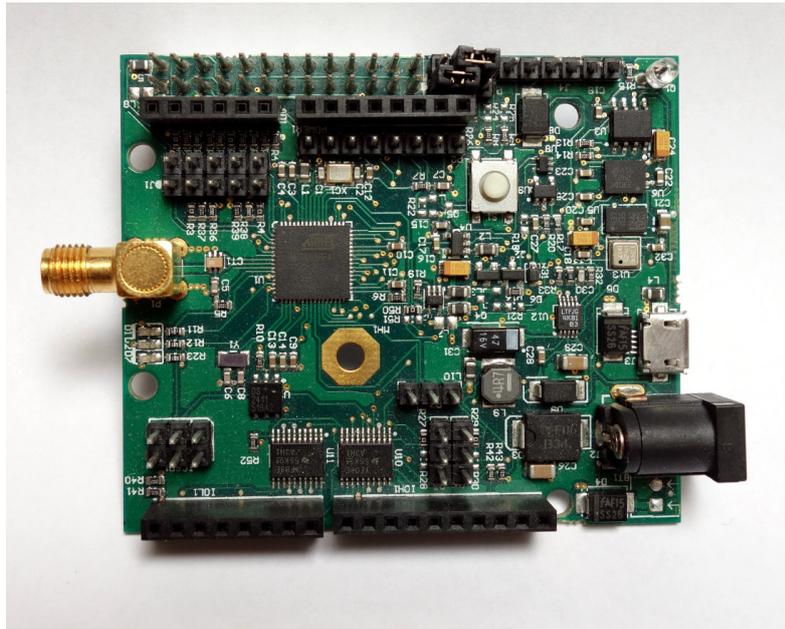


FIGURE 3.21 – La carte uSu-Edu.

- Des connecteurs, dont deux sont particulièrement importants : un port d’extension sur le modèle de celui des Arduino qui permet de connecter la plupart des cartes d’extensions conçues pour Arduino, et un port pour se connecter sur un Raspberry Pi.

Les capteurs Lidar Lite v3 et VL53L0X sont connectés à des cartes Arduino Nano, pour des raisons de facilité : des pilotes spécifiques à ces capteurs sont disponibles sur la plate-forme Arduino, ce qui évite d’avoir à redévelopper des pilotes pour la carte uSu-Edu et le nouveau système d’exploitation. La carte Arduino est elle aussi connectée à un port série du Raspberry Pi 3.

### 3.3.4 Système d’exploitation

Le système d’exploitation fait le lien entre le matériel et le logiciel, fournissant des services basiques aux applications à travers une interface de programmation qui permet de s’abstraire des considérations bas-niveau du matériel. Les systèmes d’exploitation qui font fonctionner les ordinateurs classiques ne sont pas adaptés aux matériels de faible puissance que sont les microcontrôleurs 8 bits et 16 bits. Ces derniers, grâce à leur basse faible consommation d’énergie et leurs entrées sorties intégrées, sont souvent les seules puces qui peuvent être utilisées dans les dispositifs nécessitant de limiter leur consommation au minimum, tels que les nœuds de réseaux de capteurs sans fil ou bien le LMSC. Les systèmes destinés à ces applications peuvent être classés en trois types : *multithread*, événementiels et hybrides. Dans les systèmes *multithread*, une tâche peut être préemptée par le système à tout moment de son exécution pour exécuter une autre tâche à sa place. Ce mécanisme permet de répondre aux contraintes temps réel : les tâches prioritaires peuvent s’exécuter dans un temps pré-déterminable. Par contre, ce type de système implique une forte empreinte mémoire et une certaine complexité en raison

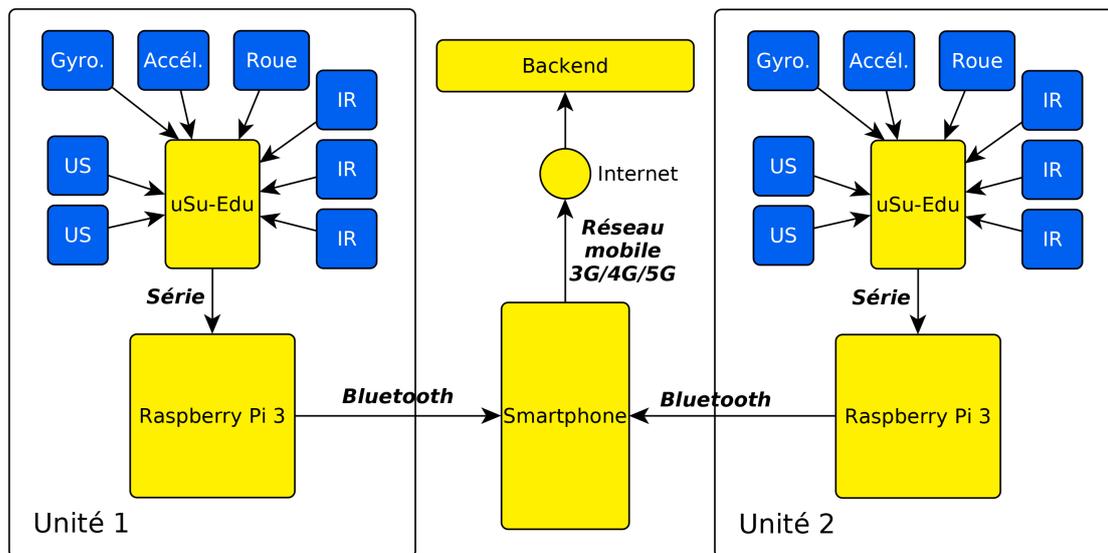


FIGURE 3.22 – Architecture initialement adoptée pour le LMSC.

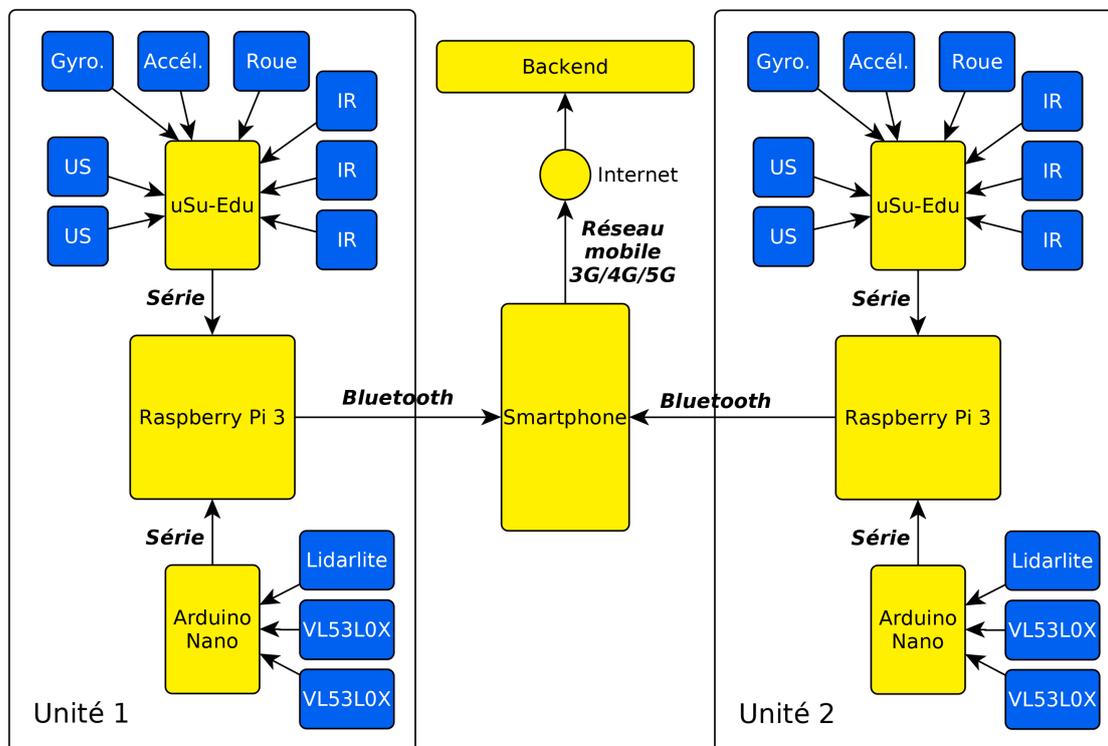


FIGURE 3.23 – Architecture du LMSC après ajout de capteurs.

de la nécessité de sauvegarder l'état des tâches, c'est-à-dire les registres du processeur et la pile d'exécution à chaque changement de contexte. MantisOS, DREAM et SDREAM sont des exemples de systèmes *multithread* destinés aux ordinateurs embarqués [Bhatti et al., 2005, De Vault and HOU, 2002, Zhou et al., 2006].

Dans les systèmes événementiels, les tâches sont interruptibles mais pas préemptibles. Ils sont plus simples à développer, ont une empreinte mémoire plus faible, et sont plus faciles à porter que les systèmes multithread, car ils n'ont pas besoin de gérer les sauvegardes du contexte des threads, qui sont nécessaires en cas de préemption. En contrepartie, ils ne sont pas adaptés aux contraintes temps réel [Duffy et al., 2008] : le système peut se trouver monopolisé par une tâche très longue. TinyOS, Contiki et SOS sont des exemples de systèmes événementiels pour les ordinateurs embarqués [Hill et al., 2000, Dunkels et al., 2004, Han et al., 2005]. Les systèmes d'exploitation hybrides combinent des caractéristiques des deux précédents types. C'est le cas par exemple de TinyOS avec TOSThread, de Contiki *multithread*, et de MIROS [Klues et al., 2009, Dunkels et al., 2004, Liu et al., 2014].

Deux systèmes d'exploitation sont utilisés dans le LMSC : un pour les Raspberry Pi, et un pour les cartes uSu-Edu. Sur les premiers, un système Raspbian classique fonctionne. C'est une distribution GNU/Linux spécialement développée pour Raspberry Pi, basée sur la très stable et répandue distribution Debian GNU/Linux. Le système d'exploitation qui fonctionne sur les cartes uSu-Edu est plus intéressant. Celui-ci est un système événementiel amélioré. Chaque tâche  $y$  est représentée par un ensemble « d'actions » formant un automate. Chaque action comprend en général une ou plusieurs transitions, « immédiates » ou temporisées, vers une ou plusieurs autres actions. Un ordonnanceur s'exécute à la fin de chaque action pour sélectionner dans la liste des actions programmées laquelle doit être exécutée à la suite de la précédente et mettre à jour les délais des actions temporisées. Des priorités peuvent être affectées aux actions à effectuer, influençant la sélection de l'ordonnanceur et permettant ainsi de fournir une meilleure qualité de service pour certaines tâches prioritaires. Ainsi, contrairement à un système événementiel classique, il n'y a pas besoin d'attendre la fin d'une tâche pour passer à une autre tâche prioritaire ; le blocage est limité à la durée de l'action en cours. Ces meilleures performances temps réel sont obtenues en contrepartie d'une légère baisse des performances du système, puisque l'ordonnanceur est appelé à la fin de chaque action.

La granularité de la découpe des tâches en actions est à la charge de la personne qui développe l'application. Plus cette granularité est fine, plus le système s'approche du *multithreading* et de la préemptabilité, qui peuvent être vus comme des actions atomiques, et donc de la satisfaction de contraintes temps réel, au détriment de l'empreinte mémoire et de la complexité de représentation et de gestion des applications. À l'inverse, des actions comprenant beaucoup d'éléments limitent la souplesse du système mais le simplifient. En pratique, le découpage des tâches d'un nœud de réseau de capteurs sans fil ou d'un objet intelligent simple est « naturel » : acquisition de données, traitement de ces données, et action – transmission à un échelon supérieur du réseau, stockage, ou bien actuation physique. Ce découpage sépare généralement les opérations longues ou impliquant des temps d'attente, ce

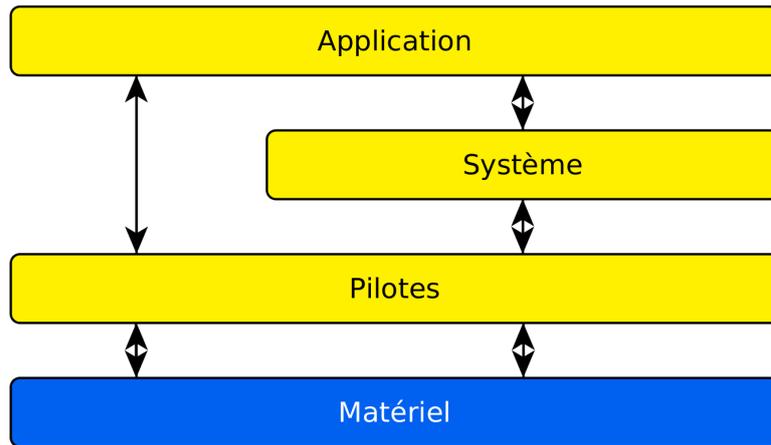


FIGURE 3.24 – Connexions entre l'application, le système, les pilotes, et le matériel.

qui permet de maximiser la performance du système.

L'application, c'est-à-dire l'ensemble des tâches, a accès aux pilotes des composants matériels (capteurs, ports série, etc.) et aux primitives offertes par le système (gestion de la mémoire, gestion des événements, etc.), qui a lui-même accès aux pilotes (figure 3.24). Une tâche est codée par une fonction contenant un *switch*, où chaque action est contenue dans un cas du *switch* (programme 3.2). La fonction a trois paramètres : le numéro de l'action à exécuter, un pointeur sur les données transmises à l'application, et la taille de ces données. Les communications et la synchronisation entre les tâches sont réalisées par des événements logiciels, c'est-à-dire des messages envoyés à l'ordonnanceur pour lui demander d'exécuter une action d'un état. Une action qui crée un événement spécifie la tâche et l'action à exécuter, un pointeur sur des données et la taille des données. On peut attribuer aux événements une priorité supérieure à la normale ou bien un délai avant que l'ordonnanceur ne mette l'événement dans sa liste d'attente d'exécution.

Un simulateur permettant de s'abstraire du matériel (figure 3.25) et de faire fonctionner directement les applications et le système sur un ordinateur doté d'un système d'exploitation classique (Linux, macOS, Windows) a été développé. Cela facilite le débogage, notoirement plus difficile avec un système embarqué, et permet en plus, l'intégration de nœuds virtuels et de nœuds matériels dans un même réseau.

```

1 _uint8 userThread0(_uint8 action_number, _uchar* data,
2   _uint8 size)
3 {
4   switch( action_number )
5   {
6     case 0:
7       {
8         /* code de l'action 1 */
9       }
10    break;

```

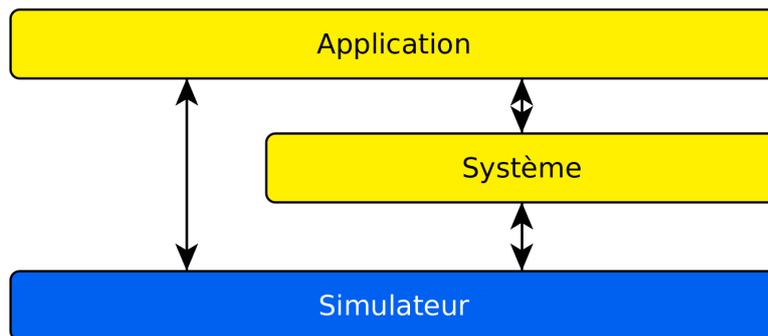


FIGURE 3.25 – Liens entre l’application, le système et le simulateur ; ce dernier simule le matériel et les pilotes.

```

11     case 1:
12     {
13         /* code de l'action 2 */
14     }
15     break;
16 }
17
18 return 0;
19 }
  
```

Listing 3.2 – Structure du code d’une tâche possédant deux actions dans le système d’exploitation.

Le système est doté d’une fonctionnalité de mise à jour à distance. À la compilation, les points d’entrée des tâches sont alignés par le compilateur sur le début des pages mémoire du microcontrôleur. Cela permet de remplacer une tâche entière, à l’intérieur d’une page, sans affecter les adresses mémoires des autres. Si les tâches n’étaient pas alignées sur les pages, les adresses mémoires qu’elles gèrent seraient dépendantes et la modification de l’une des tâches nécessiterait la modification des autres.

Le système d’exploitation peut facilement être équipé de deux mécanismes de tolérance aux pannes : la pseudo-réplication passive et la détection d’erreur avec un compteur de redémarrages [de Vaulx, 2003]. La pseudo-réplication passive est une forme temporelle de redondance, dans laquelle, parmi plusieurs répliques du système, une seule est active à la fois. Quand une erreur est détectée, le système active une autre réplique et bascule dessus ; contrairement à la réplication passive, la reprise est faite à partir d’un état initial prédéfini plutôt que d’un état sauvegardé de la dernière réplique passive. Il n’y a donc pas besoin de sauvegarder régulièrement l’état du système, ce qui simplifie l’implémentation du mécanisme. Il est à noter que la pseudo-réplication passive est applicable dans le cas de la canne 2SEES grâce à la présence de deux unités matérielles identiques.

Le compteur de redémarrages permet au système de reprendre son fonctionnement normal après une erreur affectant partie de la mémoire *flash* où se trouve le code du système. Un compteur, incrémenté par le chargeur d’amorçage à chaque re-

démarrage, permet de détecter une corruption du code du système en vérifiant la fréquence des redémarrages. Pour éviter les défaillances, le microcontrôleur Atmega 128RFA1 qui équipe la carte possède un chien de garde (*watchdog*) qui, s'il n'est pas remis à zéro régulièrement, provoque un redémarrage du système [Atm, 2014]. Dans ce contexte, une fréquence élevée de redémarrages indique une erreur affectant le système lui-même ou le matériel. Une mesure permettant d'éviter la défaillance de cette instance du système, valable dans le cas où l'erreur provient d'une corruption du contenu de la mémoire, est de faire basculer le système sur une autre copie située à un autre endroit dans la mémoire, supposé non corrompu. C'est le chargeur d'amorçage qui effectue cette opération.

Le fait que le système d'exploitation utilise des automates pour représenter les tâches est également intéressant parce qu'il facilite l'implémentation de mécanismes de tolérance aux fautes. En effet, en observant le fil de l'exécution des états des automates-tâches, il est possible de détecter les erreurs dans le système, ce qui permet ensuite l'implémentation d'un mécanisme de redondance temporelle. L'analyse du fil d'exécution constitue une forme d'analyse de signature du programme [Arlat et al., 2006]. Elle nécessite un sous-système de surveillance de l'enchaînement des états, qui peut s'exécuter soit à l'intérieur du système, ou bien dans une autre instance du système, sur un autre microcontrôleur, ou bien encore comme application dans un autre dispositif ; dans notre cas, on peut imaginer laisser la surveillance du système d'exploitation fonctionnant sur la carte uSu-Edu à une application fonctionnant sur le Raspberry Pi. En surveillant le fil d'enchaînement des états, il est tout d'abord possible de repérer les exécutions incorrectes parce qu'enfreignant la structure de l'automate telle qu'elle a été définie, c'est-à-dire les exécutions *syntactiquement* incorrectes. Par exemple, dans l'application décrite sur la figure 3.26, une exécution  $\dots \rightarrow TA_1 \rightarrow TA_2 \rightarrow TA_1 \rightarrow \dots$  nous indique une défaillance, sous l'hypothèse de la justesse de l'automate de représentation de l'application. On peut imaginer une analyse du flux des états à un niveau plus élevé, où les défaillances correspondraient à des enchaînements qui sont valides si l'on se réfère à l'automate. Cette détection est plus difficile, car elle correspond à une défaillance qui n'a pas été décrite formellement par l'automate, et on ne peut donc pas utiliser ce dernier pour la juger. Il est possible qu'elle ne soit pas encodée dans l'automate parce qu'il est impossible d'utiliser une telle représentation pour discriminer les cas valides de ceux ne l'étant pas. Il est également possible qu'une telle représentation soit envisageable, mais qu'elle représente une augmentation de complexité telle que la personne ayant conçu l'application ait choisi de simplifier l'automate de la tâche et l'implémentation, afin, par exemple, d'économiser de l'espace dans la mémoire morte du microcontrôleur (figure 3.27). Dans ce second cas, on peut imaginer que le sous-système d'analyse et l'automate complet soient stockés dans un autre composant (par exemple, pour le cas du LMSC, dans le Raspberry Pi). Le premier cas, où la défaillance n'est pas codable dans la structure de l'automate, est plus complexe. Ce cas nécessite que le sous-système d'analyse possède des connaissances supplémentaires lui permettant de discriminer les cas sains des défaillances. À quoi peut correspondre une telle défaillance ? Il peut s'agir d'un problème de capteur : un capteur se trouve très souvent en désaccord avec les autres capteurs homogènes, ce qui pousse le système à déclencher des phases de re-calibration. Il est probable que ces

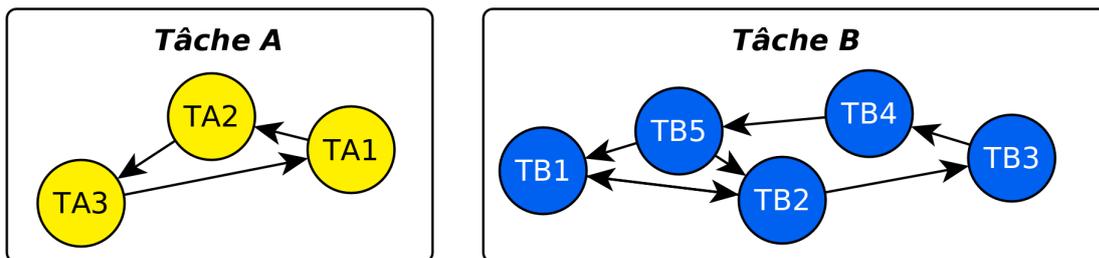


FIGURE 3.26 – Le système d’exploitation de la carte uSu-Edu représente les tâches par des automates.

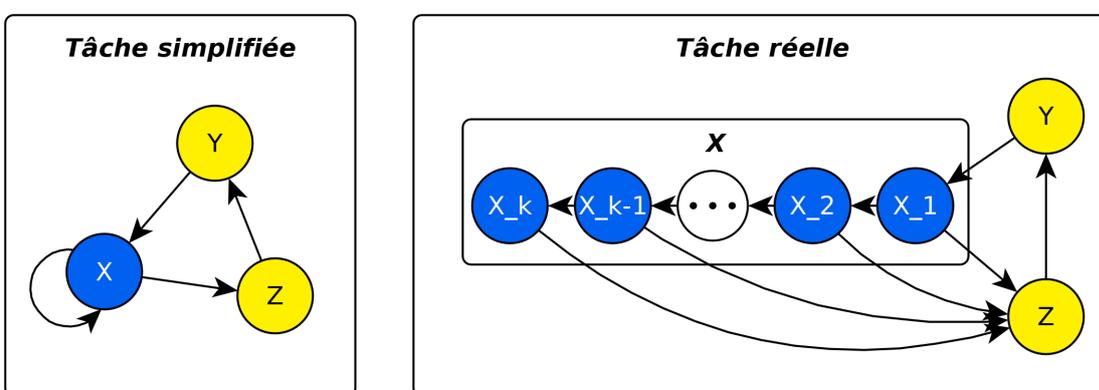


FIGURE 3.27 – L’automate de gauche est la simplification de l’automate de droite, dans lequel la tâche X peut être répétée au plus  $k$  fois alors que dans la version simplifiée, cette limite disparaît.

re-calibrations soient un état de l’automate de détection ; le sous-système d’analyse pourra alors observer dans le fil d’exécution un grand nombre de fois l’état « re-calibration ». S’il revient trop fréquemment, cela signifie très probablement que le capteur en question est défaillant. Pour détecter cette anomalie, le sous-système d’analyse devra s’appuyer sur des indicateurs statistiques : par exemple, si l’automate passe plus de 10% de son temps (discret) en re-calibration, alors on considère le capteur comme défaillant. Ce type de défaillance et les informations permettant la détection sont de niveau *fonctionnel*. Il peut également s’agir d’une défaillance dépendante du *contexte* où les informations ayant trait au contexte en question ne sont pas accessibles à l’application dans laquelle fonctionne l’automate, par exemple, une information accessible au smartphone uniquement. Dans le LMSC, ce cas est tout à fait envisageable étant donné que le smartphone gère les informations de haut niveau comme la cartographie. On a ainsi plusieurs possibilités, correspondant à plusieurs niveaux de validité : syntaxique, fonctionnelle, contextuelle. Plus le niveau est bas, plus les défaillances sont simples à détecter.

## 3.4 Conception et implémentation de la fonctionnalité de navigation

Après la construction du prototype, le premier travail a été la recréation d'une fonctionnalité de navigation.

### 3.4.1 Problématique

Deux grandes méthodes de localisation s'offrent à nous : la navigation à l'estime (*dead-reckoning* en anglais) et les méthodes de positionnement absolu. On a vu (section 3.1.3) des exemples pour chacune de ces méthodes ; la version originale du SEES par exemple, employait les deux : encodeur de roue / gyromètre (navigation à l'estime) et GPS (positionnement absolu). Les méthodes de positionnement absolues sont conceptuellement simples car leur principe est très générique : le dispositif obtient une ou un ensemble d'informations qui lui permettent de connaître sa position. Cela peut être visuel : une caméra détecte un élément, comme par exemple un QR code spécifique, dont la position est précisément connue. Les systèmes de GNSS sont les mécanismes de positionnement absolus les plus utilisés. D'autres systèmes radio peuvent être utilisés : balises Bluetooth, RFID, points d'accès WiFi, etc. Les éléments qui apportent une information de positionnement sont appelés « amers ».

Les méthodes de navigation à l'estime sont en réalité moins des méthodes de localisation que de suivi de trajectoire. Elles consistent à reconstruire une trajectoire « globale » à partir des informations sur le déplacement effectué à chaque instant de cette trajectoire. On en a vu un exemple avec la navigation en intérieur du système SEES, où la trajectoire est reconstituée avec les informations de distance parcourue et de changement de cap<sup>15</sup> mesurés en permanence. La navigation à l'estime ne constitue en fait une méthode de localisation que dans la mesure où, si la position d'origine est connue, la connaissance de la trajectoire permet de calculer la position finale et chacune des positions intermédiaires.

En pratique, il existe une nuance qui amène soit à une troisième méthode ou bien à une généralisation de la méthode de localisation absolue décrite plus haut : les méthodes de positionnement absolues dans un contexte limité. La méthode exposée précédemment a été décrite comme « purement » absolue : où que se trouve le système, sa position sera connue. De telles méthodes existent, mais ne sont pas seules : d'autres fonctionnent de manière similaire mais dans un contexte local. Ainsi, l'information de la proximité de toilettes ne permet pas de savoir où l'on se trouve dans le monde mais le permet probablement à l'intérieur d'un bâtiment. La connaissance positionnelle est alors dépendante d'une connaissance supplémentaire permettant de discriminer le lieu exact : souvent, cela se fera par la connaissance floue du lieu (telle que : « je sais que je suis dans tel bâtiment, mais je ne sais pas où exactement ; la vue de la porte des toilettes m'apporte l'information »).

Systèmes techniques, animaux et humains fonctionnent de la même manière. Le

---

15. Le cap est la direction dans le plan horizontal.

terme « navigation à l'estime » provient de la navigation maritime, dans laquelle les navires estiment depuis longtemps leur position en mesurant régulièrement leur vitesse et leur cap. Chez les animaux et les humains, ce mode de déplacement est appelé l'intégration de chemin. Pour la localisation absolue, ou absolue avec contexte, il est facile de trouver des exemples : le domaine de la navigation maritime est encore concerné (étoiles, phares, etc.) ; une personne qui connaît et voit le Puy-de-Dôme sait immédiatement où elle se trouve.

La problématique de la navigation consiste dans notre cas à trouver un équilibre dans la satisfaction de contraintes conflictuelles. Les contraintes principales sont les suivantes :

1. Précision : la localisation doit être suffisamment précise (l'écart entre position estimée et position réelle doit être inférieur à une trentaine de centimètres environ), et ce, à n'importe quel instant. Cela ne signifie pas que la précision doit être en permanence supérieure à cette valeur, mais qu'elle doit l'être quand c'est nécessaire (à proximité de portes par exemple) ;
2. La consommation d'énergie doit être la plus faible possible, pour augmenter l'autonomie. Pour ce faire, il est nécessaire d'abaisser le niveau de fonctionnement des composants : éteindre au maximum les capteurs, mettre les microcontrôleurs inutilisés en sommeil, etc. suivant le principe de *l'active multi-sensor context awareness* [Yusro et al., 2013a] ;
3. Infrastructures : la dépendance aux infrastructures doit être minimisée. Il faut préférer l'indépendance à toute infrastructure, la dépendance à des infrastructures peu contraignantes, ou bien l'exploitation opportuniste. Nous considérons qu'une infrastructure est peu contraignante si elle est pré-existante, ou prévue pour être déployée dans le futur (cf. chapitre 2). L'exploitation opportuniste consiste à utiliser une ou des infrastructures quand c'est possible et avantageux, tout en étant capable de fournir les fonctionnalités en leur absence. L'utilisation des infrastructures apporte alors une amélioration de caractéristiques non-essentiels : performance énergétique par exemple. La dépendance totale à une infrastructure, surtout contraignante (spécialisée), doit être évitée, car elle limite l'ubiquité potentielle du système.

Les deux grandes méthodes de navigation ne sont pas égales devant ces trois contraintes. Dans le cas de la précision, elle sont globalement opposées. D'un côté, de par son principe même, la navigation à l'estime conduit inmanquablement à une dérive au fur et à mesure de la progression du trajet, voire, ce qui est pire, du temps. Elle est très adaptée à des trajets courts et rapide, sur lesquels cette dérive est petite devant l'information de localisation, mais pas aux trajets longs et aux durées longues. De l'autre côté, les méthodes de positionnement absolu souffrent en général d'une précision relativement faible, même si ce n'est pas toujours le cas ; en revanche, cette précision ne se dégrade pas avec l'évolution du trajet et le temps. C'est plutôt l'inverse qui aurait tendance à se passer : un récepteur GPS classique met par exemple une certaine durée à démarrer car il doit commencer par recevoir les informations de positions des satellites pour calculer sa propre position ; une fois ces données acquises, la localisation devient effective. Un autre exemple d'amélioration peut être observé avec les méthodes absolues dépendantes du contexte : au

fur et à mesure de la progression du trajet, des amers supplémentaires sont susceptibles d'être vus, augmentant la précision de la localisation. Les deux méthodes de navigation sont donc complémentaires : la navigation à l'estime permet de suivre la trajectoire avec précision grâce à des « rafraîchissements » périodiques à l'aide d'une méthode absolue afin d'éviter la dérive.

Sur le point de l'infrastructure, les deux méthodes sont également opposées, ici, à l'avantage général de la navigation à l'estime. Cette dernière est généralement indépendante d'éléments spécifiques de l'environnement, car elle ne dépend que de la mesure du déplacement ou de la vitesse et des rotations. Mis à part le cas des capteurs inertiels, ces deux données ne sont pas réellement complètement indépendantes de l'environnement : la vitesse est très souvent obtenue par mesure du déplacement relatif dans l'environnement (rotation d'une roue en contact avec le sol, vitesse du vent pour un avion, flot optique avec un caméra) ; le cap est souvent obtenu grâce à un magnétomètre. Mais ces éléments ou aspects de l'environnement ne constituent généralement pas une infrastructure spécifiquement déployée, ce sont des caractéristiques naturelles de l'environnement. À l'inverse, dans le cas des systèmes de localisation absolue, il est nécessaire d'obtenir une information précise sur la localisation de son environnement. Plus la spécificité<sup>16</sup> d'un amer est élevée, plus il apporte d'information, et donc, plus il simplifie et augmente la performance de la méthode de navigation. On voit ici l'apport d'une infrastructure : elle offre directement un réseau standardisé d'éléments dont la spécificité est importante (tags RFID, QR codes affichés) ; dans un environnement « naturel », les amers sont généralement moins spécifiques (comme le signe indiquant des toilettes) et, surtout, ils sont rarement conçus pour être détectés par des systèmes électroniques.

En ce qui concerne la consommation d'énergie, les choses sont moins claires. Les capteurs inertiels et les magnétomètres susceptibles d'être embarqués dans nos dispositifs ont des consommations très faibles (environ 1 mA), et celle des encodeurs de roue est moyenne (10 à 20 mA). Par contre, les méthodes de détermination de la vitesse et du cap par vision entraînent nécessairement une consommation non négligeable. Les méthodes absolues ont une consommation qui semble moyenne à élevée, le récepteur GNSS du SEES, par exemple, consomme environ 25 mA, ce qui semble être une valeur commune. Les méthodes utilisant des radios locales ont également une consommation non négligeable. Les méthodes basées sur la détection d'amers par la vision sont de grosses consommatrices d'énergie. Un avantage des systèmes de localisation absolue dans le domaine de la consommation d'énergie par rapport à la navigation à l'estime est qu'ils peuvent être utilisés de façon périodique plutôt que continue, ce qui diminue fortement leur consommation.

Si en extérieur, le problème est relativement simple grâce à la présence de systèmes GNSS (cf. chapitre 2), en extérieur, il en va autrement. Dans le cadre de cette thèse, nous nous sommes concentrés sur la fonctionnalité de navigation en intérieur, et avons donc dû explorer les possibilités offertes par les différentes méthodes et les manières de les combiner.

---

16. Caractère d'être spécial, voire, unique.

### 3.4.2 Filtrage particulière

La méthode implémentée dans la canne est dérivée du filtrage particulière. Deux versions ont été expérimentées, une en simulation avec les données collectées sur le SEE-stick, puis une seconde capable de fonctionner en temps réel sur le LMSC.

#### Principe

Formellement, le problème qu'on cherche à résoudre est celui de l'estimation de la position de la canne intelligente, que nous ne pouvons pas mesurer directement, par l'intermédiaire d'observations bruitées. La position réelle de la canne est un état caché, les données des capteurs qu'on utilise pour la localisation sont des observations de cet état, et la position que nous calculons est l'état estimé. L'estimation revient à calculer la densité de la probabilité de l'état caché – c'est-à-dire la position réelle. On cherche à minimiser l'erreur entre la position réelle et la position estimée. Plus formellement, le système et les observations sont décrits par les équations 3.3 et 3.4, où  $x_k$  est l'état caché,  $v_k$  un bruit aléatoire d'évolution de  $x_k$ ,  $y_k$  est l'observation de l'état caché  $x_k$ ,  $e_k$  le bruit de mesure, et  $k$ , l'indice temporel [Gustafsson, 2010]. Le but du filtrage bayésien est d'approcher la distribution *a posteriori*  $p(x_k|y_{1:k})$ , c'est-à-dire la distribution de probabilité de l'état du système  $x_k$  sachant l'ensemble des observations passées  $y_{1:k}$ .

$$x_{k+1} = f(x_k, v_k) \quad (3.3)$$

$$y_k = h(x_k) + e_k \quad (3.4)$$

La formule de Bayes et les probabilités totales donnent alors les équations 3.5, 3.6, et 3.7 [Gustafsson, 2010]. Les équations 3.5 et 3.6 servent à calculer la densité de probabilité grâce aux observations (mises à jour capteur), et l'équation 3.7 à faire évoluer l'état estimé du système au fur et à mesure des itérations.

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \quad (3.5)$$

$$p(y_k|y_{1:k-1}) = \int_{\mathbb{R}^{n_x}} p(y_k|x_k)p(x_k|y_{1:k-1})dx_k \quad (3.6)$$

$$p(x_{k+1}|y_{1:k}) = \int_{\mathbb{R}^{n_x}} p(x_{k+1}|x_k)p(x_k|y_{1:k})dx_k \quad (3.7)$$

Plusieurs méthodes de calcul se basent sur ces équations pour l'estimation d'un système caché observé. Le filtre de Kalman est la solution optimale dans le cas où les modèles d'évolution et d'observation  $f$  et  $h$  sont linéaires et les bruits  $v$  et  $e$  gaussiens. En pratique, ce n'est pas souvent le cas. Il est alors nécessaire d'employer une autre technique, qui permette d'estimer des systèmes non-linéaires avec de meilleures performances. Les principales sont le filtre de Kalman étendu, le filtre de Kalman *unscented*, les filtres basés sur des grilles et le filtre particulaire (aussi nommé filtre bayésien séquentiel). Le filtre de Kalman étendu est basé sur l'application des formules du filtre de Kalman classique sur une version linéarisée des fonctions  $f$  et  $h$ . La linéarisation est donnée par la formule de Taylor du premier ordre (équation 3.8, avec  $\bar{x}_k$  le point où on effectue la linéarisation). Le filtre de Kalman *unscented* repose sur une linéarisation similaire. Ces deux filtres donnent de meilleurs résultats que le filtre de Kalman sur les systèmes non-linéaires mais ils peuvent diverger, et il est difficile d'analyser les causes de divergences [Gustafsson, 2010].

$$x_{k+1} = f(\bar{x}_k, 0) + (x_k - \bar{x}_k) \left. \frac{\partial f(x_k, v_k)}{\partial x_k} \right|_{x_k \rightarrow \bar{x}_k, v_k=0} + v_k \left. \frac{\partial f(x_k, v_k)}{\partial v_k} \right|_{x_k \rightarrow \bar{x}_k, v_k=0} \quad (3.8)$$

Les filtres basés sur des grilles représentent la densité de probabilité d'état pour l'élément dont la position est estimée directement par l'affectation de probabilités aux cases d'une grille discrète, c'est-à-dire un ensemble de points [Gustafsson, 2010]. Dans le cas de la localisation simple – l'estimation de la position, ils travaillent sur une carte *raster*, et affectent à chaque « pixel » la probabilité que l'élément s'y trouve. Cette représentation est très naturelle mais devient lourde en mémoire et en calculs quand la taille de la grille est importante et surtout, quand l'état à estimer comprend de nombreuses variables, car cela augmente exponentiellement la taille de la grille.

Les filtres particuliers fonctionnent de manière similaire aux filtres basés sur des grilles, la différence essentielle étant qu'ils s'affranchissent de la grille. Plutôt que de conserver la densité de probabilité d'état pour chaque point, ils répartissent la densité sur un nombre fixé de points qui évoluent au cours du temps – ce sont les « particules » qui donnent leur nom à la méthode. Chaque particule représente un état possible du système, dont la probabilité est représentée par un poids qui lui est affecté. La fonction de densité de probabilité est donc encodée par les particules et leurs poids. Comme le nombre de particules est limité, et que la fonction de densité s'étend le plus souvent sur un domaine non borné (p. ex. : distribution normale), cette dernière n'est pas intégralement représentée; le filtre particulaire ne permet que de l'approximer. Afin d'encoder le mieux possible la fonction de densité avec une quantité bornée de particules, le filtrage particulaire cherche à maintenir les particules dans les régions de l'espace des états où la densité de probabilité est non-nulle. Si rien n'est fait, il est probable que toutes les particules, sauf une, convergent au fur et à mesure des étapes vers un poids nul, alors que l'autre particule converge vers une probabilité de 1 [Gustafsson, 2010]; le résultat serait alors une très mauvaise approximation de la fonction de densité. Pour éviter cela,

les particules sont régulièrement régénérées : on augmente le nombre de particules là où les poids calculés sont élevés.

Un filtre particulaire nécessite pour fonctionner une distribution initiale, un modèle d'évolution des particules, et un modèle d'observation. Le modèle d'évolution décrit comment les particules se déplacent ; il représente la fonction  $f$  de l'équation 3.3 et le terme  $p(x_{k+1}|x_k)$  de l'équation 3.7. Le modèle d'observation décrit la probabilité d'avoir l'observation (qui est réelle) avec une particule donnée ; il représente la fonction  $h$  de l'équation 3.4 et le terme  $p(y_k|x_k)$  de l'équation 3.5. Intuitivement, le fonctionnement du filtre particulaire peut être expliqué ainsi :

1. On a un ensemble initial donné de particules ;
2. Les particules se déplacent selon un modèle stochastique pré-défini. Par exemple, une voiture continuera sur sa lancée en faisant peu varier sa vitesse ;
3. Pour chaque particule, on regarde si elle est crédible. Typiquement, on simule ce qu'observeraient les capteurs si le système réel était dans l'état de cette particule et on compare cette simulation avec l'observation réelle des capteurs : plus elles correspondent, plus on augmente le poids de la particule, car on considère qu'elle représente mieux l'état réel du système.

### Implémentation

Notre première implémentation d'un filtre particulaire était une simulation utilisant les données capteurs enregistrées par le SEE-stick tout au long d'un parcours dans le bâtiment E de l'ISIMA. La problématique était d'estimer la position de la canne en intérieur, sans infrastructure ; pour cela, il fallait utiliser exclusivement les capteurs embarqués et des données connues a priori : la carte du bâtiment. En pratique, les premières expérimentations (section 3.1.3) ont fait apparaître que la navigation à l'estime dérivait et devait être corrigée avec une méthode absolue de localisation. Nous avons choisi d'utiliser les capteurs ultrasoniques, car ils apportent une information absolue (contextuelle), en détectant la distance aux murs, tout du moins quand ceux-ci sont proches du SEE-stick. Les deux capteurs situés sur les côtés de la poussette, et orientés orthogonalement par rapport à la direction de la course, étaient en particulier susceptibles d'offrir des données précises de positionnement entre les murs. Les capteurs ultrasoniques étaient les seuls capteurs d'obstacles présents sur le SEE-stick ; avec le prototype de LMSC, d'autres capteurs peuvent être utilisés de la même manière. Les capteurs dont les données ont été utilisées sont l'encodeur de roue, le gyroscope – soit les capteurs utilisés pour la navigation à l'estime – et les capteurs ultrasoniques. La connaissance de la distance aux murs qu'apportent des capteurs de distance présente l'avantage d'être facile à implémenter sous forme de filtre particulaire. Une autre option possible aurait été l'utilisation de la caméra pour la détection de caractéristiques remarquables ou d'amers de l'environnement en complément de la navigation à l'estime. Toutefois, cette méthode présente une plus grande complexité de traitement du signal.

Le filtre particulaire n'a pas été implémenté de la manière la plus simple, ou la plus orthodoxe. Le modèle de déplacement posait problème, car nous n'avions pas beaucoup de données pour en générer un qui soit suffisamment convaincant, et

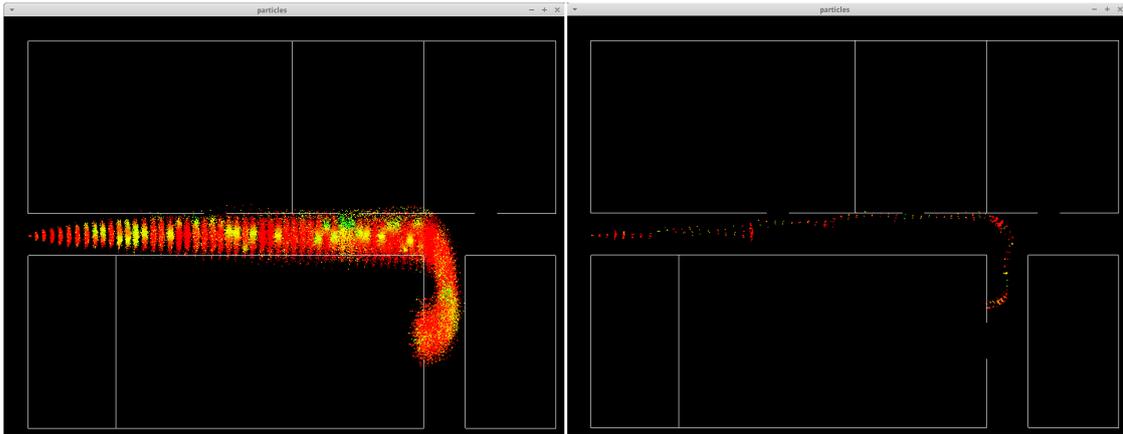


FIGURE 3.28 – Trajectoire estimée par le filtre particulaire sur des données pré-enregistrées avec le SEE-stick. À gauche, la trajectoire sans ré-échantillonnage des particules, permettant de mieux voir l’effet des détections de capteurs ultrasoniques ; à droite, la trajectoire ré-échantillonnée.

parce qu’un modèle de déplacement trop aléatoire nécessite un plus grand nombre de particules pour encoder l’information. En effet, avec un modèle de déplacement se dispersant aléatoirement, seul un petit nombre de particules va réellement porter l’information, alors qu’avec un modèle qui reflète plus fidèlement les déplacements, les particules seront plus concentrées aux endroits où se trouve réellement l’objet. Contrairement à la situation la plus classique en localisation, celle des robots mobiles, la canne n’a que très peu d’influence sur ses propres déplacements. La loi de commande d’un robot détermine le mouvement qu’il devrait effectuer et actionne les moteurs pour le faire, et l’information de la commande peut alors être utilisée comme modèle de déplacement : « le robot devrait se trouver à cet endroit, plus ou moins les imprécisions et glissements ». Alors que dans notre cas, la canne est passive et ce sont ses utilisateurs qui décident ses mouvements. On peut envisager un modèle de mobilité *a priori* sur le mouvement de la canne un peu meilleur que de l’aléatoire équiprobable. Les informations de guidage fournies par la canne à la personne qui l’utilise peuvent par exemple servir à prédire les mouvements de la canne, mais la relation reste très indirecte. La solution la plus efficace nous a paru être l’utilisation des données de navigation à l’estime avec un bruit stochastique additionnel comme modèle de déplacement. Ainsi, plutôt que d’avoir un modèle de déplacement basé sur des données *a priori* et un modèle pour chacun des trois types de capteurs, nous avons choisi d’utiliser les données de navigation à l’estime pour générer dynamiquement le modèle de déplacement, et d’utiliser les données des capteurs ultrasoniques comme seules observations.

### Expérimentation

Ce filtre particulaire a été expérimenté dans le couloir du bâtiment E de l’ISIMA sur un trajet très simple d’une vingtaine de mètres (figure 3.28).

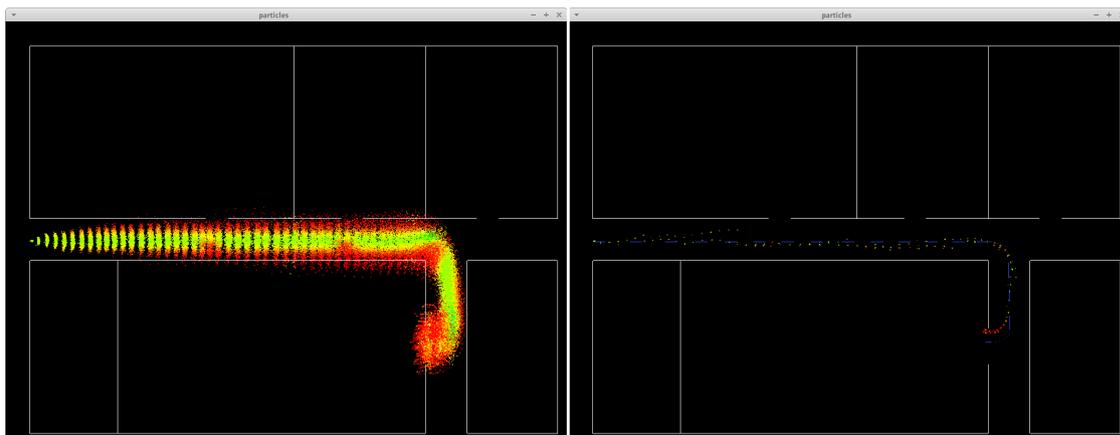


FIGURE 3.29 – Trajectoire estimée par le filtre particulaire sur des données de navigation à l’estime pré-enregistrées avec le SEE-stick et des données de capteurs de distance fixées artificiellement à des valeurs correspondant à la réalité pour mieux voir le fonctionnement du filtre. À gauche, la trajectoire sans ré-échantillonnage des particules, permettant de voir l’effet des détections simulées de capteurs ultrasoniques ; à droite, la trajectoire ré-échantillonnée.

### Discussion

Le filtrage particulaire a globalement fonctionné mais deux problèmes ont été remarqués : le coût en calcul et l’inadaptation des capteurs ultrasoniques à cette tâche.

Le problème des capteurs ultrasoniques a été qu’ils fournissaient souvent des informations erronées, avec une erreur très conséquente : la majorité du temps, quelle que soit la distance réelle entre le capteur et le mur, le capteur indiquait la distance maximale, soit environ 6 m<sup>17</sup>. À cause de cette erreur, le filtre particulaire avait très peu d’informations à utiliser, et cela a impacté l’expérimentation. Afin de vérifier le fonctionnement correct du filtre particulaire, un autre test a été effectué, cette fois en fixant artificiellement les données des capteurs ultrasoniques aux valeurs correspondant aux distances rencontrées dans le couloir (par exemple, 50 centimètres de chaque côté). Le résultat (figure 3.29) montre que les capteurs ultrasoniques n’étaient pas adaptés à cette tâche ou n’ont peut-être pas été employés correctement, mais que le concept d’utilisation d’un filtre particulaire avec des capteurs indiquant la distance aux murs est pertinent.

### 3.4.3 Filtrage particulaire simplifié

La version qui a été implémentée dans le smartphone pour le système 2SEES est une version un peu différente, optimisée et en quelque sorte simplifiée du filtre présenté plus haut. La contrainte principale était que la navigation puisse s’effectuer en temps réel sur le smartphone. Pour cela, nous avons développé une méthode

<sup>17</sup>. Peut-être à cause d’interférences entre eux, d’angle entre les cônes de détection et les murs, ou de réverbérations.

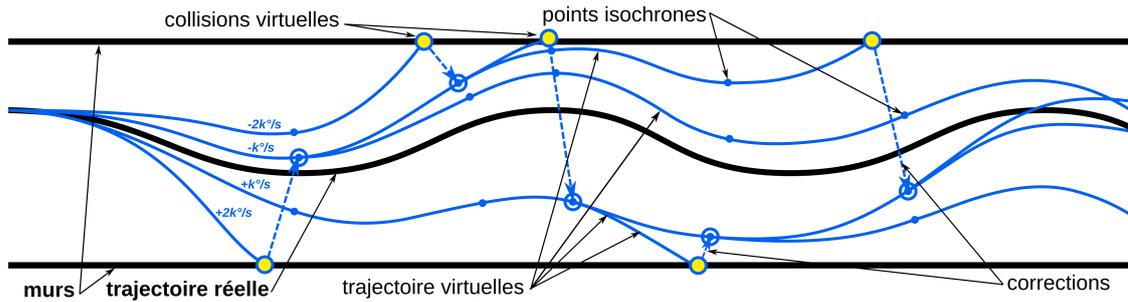


FIGURE 3.30 – Fonctionnement visuel de la méthode de correction des trajectoires par la détection des collisions avec les murs (cartographiés).

originale. Les particules  $y$  représentent des trajectoires plutôt que des états instantanés. Le modèle de déplacement qui les anime est toujours calqué sur les données de navigation à l'estime (encodeur de roue et gyromètre), avec, cette fois, une variation uniquement sur les données du gyromètre, soit la vitesse angulaire : ce modèle réduit la dimensionnalité des particules et donc leur nombre. Pour réduire encore le nombre de particules, les variations ne sont pas stochastiques mais déterministes : on répartit de manière équilibrée le nombre de particules entre deux variations de vitesse angulaire extrêmes. On aura, par exemple, une particule ayant la vitesse angulaire donnée par le gyroscope, deux particules à  $\pm 0,001^\circ/s$ , deux particules à  $\pm 0,002^\circ/s$ , et deux particules à  $\pm 0,003^\circ/s$ . Avec sept particules, on couvre une variété d'erreur de  $\pm 0,003^\circ/s$ . Avec un nombre aussi réduit de particules, il est probable qu'un bruit stochastique ne distribue pas les particules de manière « convenable ». Cet aspect déterministe est la raison pour laquelle nous qualifions le filtre de « simplifié » ; ce n'est plus un filtre particulaire à proprement parler.

Aucun capteur de distance n'est utilisé ; pour simplifier et accélérer le calcul, on utilise directement les collisions avec les murs pour juger de la validité des trajectoires des particules. Si la trajectoire d'une particule rencontre un mur, c'est qu'elle ne représentait pas la réalité, où cet événement est supposé impossible. On repositionne alors la position de la particule en question sur celle d'une autre trajectoire valide (figure 3.30) sélectionnée au hasard. Ce mécanisme imite celui du filtre particulaire ; c'en est, en quelque sorte, une version binaire : soit une trajectoire est valide car elle n'a pas rencontré de mur, soit elle est strictement invalide car elle en a croisé un. Les observations sont uniquement basées sur la cartographie.

### Expérimentation

L'algorithme a été implémenté par XU Shu, ingénieur de recherche au LIMOS. Une optimisation supplémentaire a été mise en place pour éviter que le nombre de vérifications de collisions entre les particules et les murs soit trop important, ce qui constitue l'opération la plus longue du programme. La carte est divisée en régions ; chaque région contient un ensemble de murs. Quand l'algorithme examine les collisions potentielles d'une particule, il vérifie uniquement les murs qui se trouvent à l'intérieur de la zone où se trouve la particule, ou bien, si la particule quitte une zone pour aller dans une autre, des deux zones concernées. Cette optimisation permet

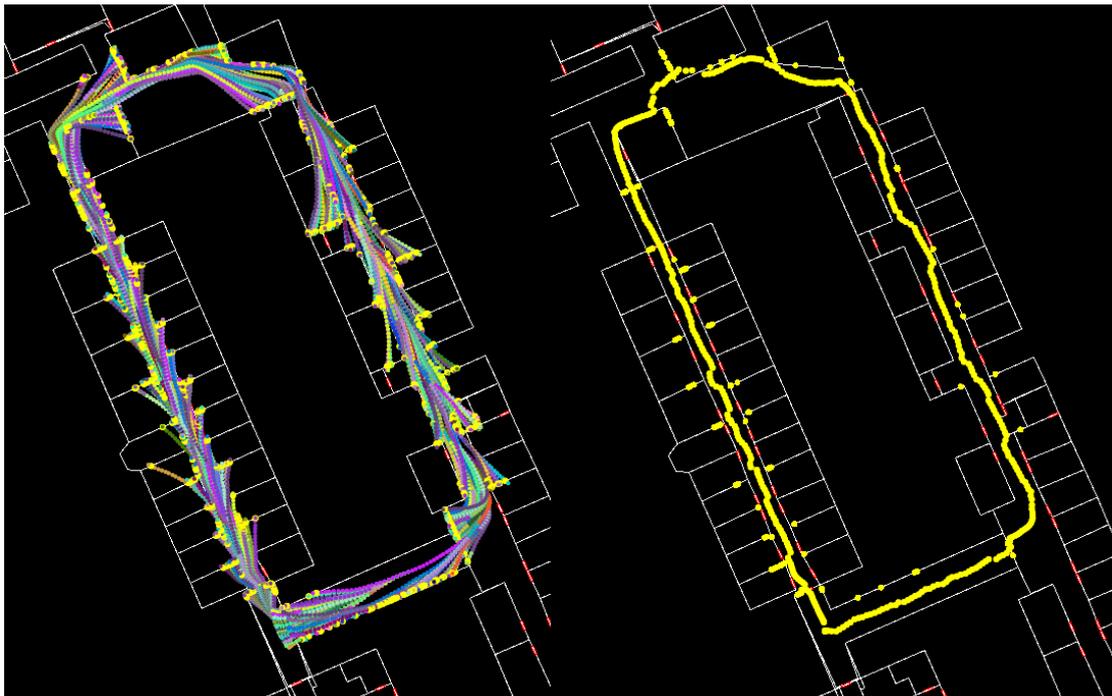


FIGURE 3.31 – Estimation de trajectoire par la méthode particulaire simplifiée – trajectoires. À gauche, toutes les particules sont représentées ; à droite, l'estimation résultante de la moyenne des particules.

une accélération importante des calculs, dont la complexité n'est plus dépendante de la taille de la carte mais seulement du nombre de particules. La contrepartie est l'affectation des murs aux zones, mais cette opération peut s'effectuer en avance du processus de navigation, et sur le serveur.

L'expérimentation a été la suivante : une personne a fait un tour dans les couloirs de l'ISIMA, les yeux non bandés, en essayant de se positionner au milieu du couloir. Ce qui était évalué n'était pas la fonctionnalité de navigation (instructions de guidage, etc.) mais la localisation. Le sujet a effectué une trajectoire en boucle traversant plusieurs couloirs dans les bâtiments du LIMOS. La distance totale parcourue est de  $150 \pm 5$  mètres. L'algorithme fonctionnait avec 50 trajectoires (particules), et une correction incrémentale de 0,001 degré par pas. Les particules avaient donc des corrections échelonnées entre -0,025 degrés par pas et +0,025 degrés par pas. L'algorithme fonctionnait à 50 Hz environ et les corrections variaient donc entre -1,25 degré par seconde et 1,25 degré par seconde. Les trajectoires résultantes sont visibles sur la figure 3.31. Il est à noter que l'amplitude des corrections devait être augmentée pour passer la première porte : en effet, il était difficile de s'aligner précisément avec la direction du couloir.

## Discussion

La figure 3.31 montre que la méthode fonctionne dans une certaine mesure : on sait où l'on se trouve dans le bâtiment, mais avec une précision correspondant environ à une largeur de couloir, donc on ne sait pas où on se trouve dans le couloir. Cette

méthode est pertinente sous l'hypothèse que l'erreur du gyroscope est bornée entre des extrêmes définis par les variations de vitesse angulaire des particules extrêmes, et que suffisamment de trajectoires sont présentes à l'intérieur de ces bornes pour ne pas s'invalider à cause de la carte. En fait, avec suffisamment de particules, l'erreur du gyroscope sera en permanence proche de l'erreur d'une des particules, le problème étant alors les collisions avec les murs lorsque l'erreur du gyroscope change : la prochaine trajectoire correspondant à l'erreur du gyroscope ne sera alors plus à l'emplacement réel de la canne. Toutefois, si la géométrie des bâtiments est favorable, ou plutôt non défavorable, les trajectoires resteront concentrées dans un espace valide restreint, et l'estimation sera correcte. Cette méthode est donc très sensible à la configuration spatiale des lieux : elle fonctionne bien dans les couloirs et les lieux complexes et contraints, où les trajectoires sont forcées de prendre le chemin correct. Par contre, dans les grands espaces ouverts et les lieux « auto-similaires », comme par exemple des couloirs parallèles très proches, la localisation ne sera pas très bonne, car les trajectoires incorrectes ne seront pas éliminées par des collisions avec les murs.

L'algorithme souffre d'un autre problème, handicapant même dans les espaces réduits : quand il n'y a pas assez de particules, et surtout, quand une erreur franche est présente, typiquement, une erreur sur l'orientation initiale de la canne, les particules peuvent se retrouver toutes bloquées dans un angle. Dans ces cas là, il n'y a plus d'estimation de position, car les particules ne peuvent plus être repositionnées ailleurs qu'à l'endroit où elles sont bloquées. Pour résoudre ce problème, on peut envisager un mécanisme supplémentaire : quand on détecte que toutes les particules sont bloquées, c'est-à-dire qu'elles ont toutes rencontré un mur, on régénère le filtre à partir d'un état antérieur, mais avec plus de particules et/ou, éventuellement, une plus large gamme de variations entre elles. On rembobine en quelque sorte le filtre, pour que les particules ou variations supplémentaires soient à même de franchir la zone concernée. Le suivi est interrompu pendant cette reprise mais ce n'est que temporaire et cela permet de retrouver la localisation correcte au bout d'un moment.

## 3.5 Conception, implémentation et test de la fonctionnalité de détection d'obstacles

Le LMSC n'est pas utilisable comme une canne classique (en balayant devant soi), et n'est pas à même d'assurer une protection physique contre les chocs comme le SEE-stick. Une fonctionnalité de détection d'obstacles a donc été développée et implémentée en vue de réaliser des expérimentations avec des personnes déficientes visuelles. Un deuxième bénéfice de la détection d'obstacle est qu'elle nous a permis de travailler sur un aspect intéressant de la robustesse de la canne : la redondance au niveau informationnel (tableau 3.1). En effet, les capteurs d'obstacles ont des caractéristiques très variées et des affinités variables avec les différents contextes (cf. section 2.2.1). Afin de maximiser les conditions d'efficacité du LMSC, nous devons utiliser plusieurs types de capteurs ; une telle configuration a été adoptée

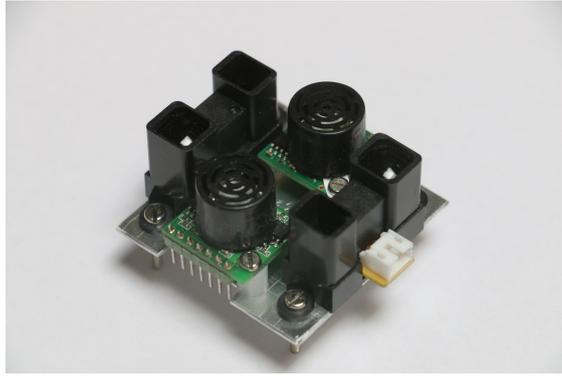


FIGURE 3.32 – Une cellule de détection composée de deux capteurs infrarouges et deux capteurs ultrasoniques, pour assurer à la fois redondance et adaptabilité à l’environnement.

pour chaque zone de détection (figure 3.32).

Nous avons commencé par déterminer globalement les affinités des capteurs d’obstacles avec les différents contextes. Pour que nous puissions envisager d’utiliser des informations contextuelles dans le système, il faut qu’elles soient effectivement accessibles, ce qui est plus ou moins simple, voire plus ou moins envisageable, selon l’aspect du contexte dont il s’agit. Pour ce qui est des capteurs d’obstacles, nous avons principalement identifié quatre éléments de contextes pertinents :

- La luminosité ambiante ;
- Les caractéristiques matérielles des obstacles ;
- Le degré d’encombrement de l’environnement ;
- Les interférences entre les capteurs ;

La luminosité ambiante, intéressante pour son influence sur les capteurs optiques, est très facilement mesurable et le LMSC embarque un capteur de ce type. On pourrait éventuellement obtenir cette information grâce à une localisation suffisamment précise et à la météo locale, mais cela augmenterait considérablement le niveau de complexité.

Les caractéristiques matérielles des obstacles sont également susceptibles d’influencer les capteurs, mais sont plus difficiles à estimer. L’utilisation d’une caméra est probablement la solution qui vient en premier à l’esprit. Un autre moyen serait d’inclure ces informations dans les données cartographiques que le système manipule. D’un côté, cela augmenterait le coût et complexifierait la cartographie, mais de l’autre côté, cela permettrait un accès simple, direct, et fiable aux informations en phase d’utilisation.

Plus le degré d’encombrement est élevé, plus un capteur avec un faisceau de détection large sera désavantagé, car il masquera la complexité de l’environnement. Or, dans un tel environnement, il est nécessaire que les informations soient suffisamment riches et donc complexes pour que la personne puisse y évoluer. Un capteur à faisceau très large ne permettra par exemple pas ou difficilement de distinguer une ouverture de porte plus étroite que son faisceau. Pour la mesure du degré d’encombrement, la situation est similaire aux caractéristiques matérielles : on peut l’estimer avec une caméra, voire une caméra 3D, ou bien on utiliser des données

### 3.5. CONCEPTION, IMPLÉMENTATION ET TEST DE LA FONCTIONNALITÉ DE DÉTECT

cartographiques, soit en encodant directement l'encombrement, soit en utilisant un modèle général pour l'estimer à partir des informations disponibles sur les lieux (p. ex. : cette pièce est un *open-space*, elle est probablement encombrée; cette pièce est un couloir, elle est probablement vide). Même si on peut aller loin dans le développement du modèle, utiliser des méthodes d'apprentissages, etc., leur utilisation ainsi que celle de la cartographie directe semblent peu fiables.

La plupart des capteurs de distance émettent un signal et analysent le signal réverbéré par l'environnement. Si plusieurs capteurs de même nature se rencontrent, ce qui est envisageable ne serait-ce que pour le cas où plusieurs personnes déficientes visuelles équipées se rencontrent, il est très probable que les signaux des capteurs se perturbent mutuellement. L'estimation de la probabilité de telles interférences est dans certains cas potentiellement très robuste et dans d'autres, très difficile. Si elle concerne un ensemble de systèmes 2SEES, alors elle est simple : le *back-end*, qui fait le lien entre tous les systèmes 2SEES, connaît la position de chacun d'entre eux, et a donc connaissance des risques de perturbations et peut en informer les systèmes. Les perturbations externes au système 2SEES sont beaucoup plus difficiles à détecter et à dissocier d'autres problèmes.

Le 2SEES est principalement destiné aux expérimentations en intérieur où le niveau de luminosité était en général limité. L'encombrement variable était, lui, bien présent, mais son estimation pose des problèmes pratiques qu'on a rapidement présentés. Nous avons donc travaillé sur les deux autres aspects du contexte, c'est-à-dire les caractéristiques matérielles et les perturbations.

#### 3.5.1 Test des affinités entre capteurs et matériaux

La première partie de l'étude a consisté à mesurer les affinités existant entre les différents capteurs et les différents matériaux ou surfaces. Les trois types de capteurs placés à l'avant du LMSC ont été testés :

- Capteur infrarouge Sharp GP2Y0A02YK ;
- Capteur ultrasonique Maxbotix EZ-0 ;
- Télémètre laser Garmin Lidar Lite v3.

Ils ont été testés sur six cibles différentes :

- Carton fort recouvert de papier blanc (désigné par la lettre *W* pour *White*) ;
- Carton fort recouvert de papier noir (*B* pour *Black*) ;
- Carton fort recouvert d'un film plastique réfléchissant (*R* pour *Reflective*) ;
- Carton fort recouvert de tissu de laine d'une épaisseur de 1 mm environ (*C* pour *Cloth*) ;
- Mousse polyéthylène épaisse rigide (*F* pour *Foam*) ;
- Plaque en verre transparent de 3 mm (*G* pour *Glass*).

Toutes ces surfaces étaient placées orthogonalement par rapport aux cônes de détection des capteurs.

Une configuration supplémentaire a été testée où le film plastique réfléchissant de la cible *R* était orienté d'un angle de 45 degrés environ avec les faisceaux des capteurs (*R/A* pour *Reflective, angled*). Les mesures ont été faites à 50 cm, 100 cm, et 150 cm  $\pm$  5 cm (figure 3.33). Les moyennes et écarts-types des mesures ont été calculés

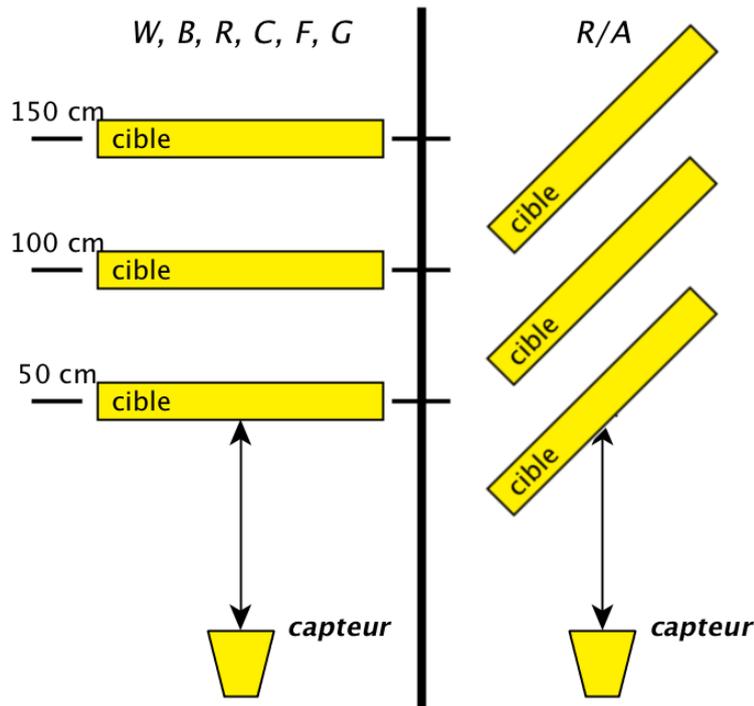


FIGURE 3.33 – Configuration pour le test des affinités capteur – matériau.

sur 1000 échantillons collectés séquentiellement à 50 Hz pour chaque combinaison des trois paramètres : capteur, matière et distance de la cible (tableau 3.3). De plus, l'effet des perturbations a été testé sur les capteurs ultrasoniques, dans la configuration présentée sur la figure 3.34, dont les résultats sont donnés dans le tableau 3.4.

Un algorithme simple de détection d'obstacles a été développé en utilisant les résultats collectés dans les tests de sensibilité au contexte. Le principe de l'algorithme est d'affecter un poids entre 0 et 1 aux estimations des capteurs. Ce poids représente la confiance que le système accorde à l'estimation. Ces confiances forment un paramètre de l'algorithme ; pour l'instant, ces confiances sont fixées manuellement, dérivées des mesures de sensibilité des capteurs aux différents matériaux, mais, dans le futur, elles pourront être calculées automatiquement à partir des données capteurs et d'informations sur le contexte mesurées (la seconde problématique). L'algorithme est spécifié à partir de deux cas extrêmes : quand il n'y a pas d'information sur le contexte, tous les capteurs sont également dignes de confiance (dans notre modèle) ; quand le contexte rend un seul type de capteur pertinent, le système ne doit utiliser que ce type de capteur. Dans la première situation, n'importe quel capteur détectant un obstacle devrait déclencher une détection du système. Les capteurs ne fonctionnent pas toujours et nous préférons les faux positifs aux accidents. Nous avons choisi d'implémenter cet algorithme naïf de manière binaire : le programme prend en entrée les estimations individuelles binaires des capteurs et les confiances binaires, et donne en sortie son estimation de la présence d'un obstacle. Cette estimation est calculée en vérifiant si le produit scalaire des deux

### 3.5. CONCEPTION, IMPLÉMENTATION ET TEST DE LA FONCTIONNALITÉ DE DÉTECT

TABLE 3.3 – Affinités capteurs – matériaux à 50 cm, 100 cm, 150 cm. Les valeurs dans le tableau sont les niveau moyens de sortie des capteurs dans l'unité native de la sortie, soit des millivolts pour les capteurs infrarouges et ultrasoniques, et directement des centimètres pour le Lidar Lite v3, avec entre parenthèses les écarts-types. Les valeurs en gras sont remarquables car la sortie du capteur y est clairement faussée.

IR	W	B	R	R/A	C	F	G
50	1095 (15)	1096 (14)	<b>426 (163)</b>	<b>1327 (19)</b>	1126 (11)	1052 (15)	<b>1469 (13)</b>
100	558 (10)	569 (16)	615 (20)	<b>19 (26)</b>	563 (18)	565 (17)	<b>814 (14)</b>
150	359 (18)	<b>289 (39)</b>	<b>493 (86)</b>	<b>18 (8)</b>	363 (19)	376 (19)	<b>410 (23)</b>
US	W	B	R	R/A	C	F	G
50	481 (7)	481 (7)	479 (7)	475 (7)	479 (7)	472 (7)	481 (7)
100	635 (7)	628 (7)	627 (6)	651 (10)	635 (6)	627 (6)	637 (8)
150	785 (7)	783 (8)	785 (8)	<b>895 (61)</b>	785 (8)	784 (10)	787 (7)
Las.	W	B	R	R/A	C	F	G
50	60 (1)	63 (2)	57 (1)	63 (5)	62 (1)	64 (1)	<b>4 (2)</b>
100	111 (0)	123 (3)	111 (1)	112 (0)	106 (1)	108 (1)	<b>207 (1)</b>
150	164 (1)	168 (2)	167 (8)	165 (1)	163 (1)	168 (1)	<b>213 (1)</b>

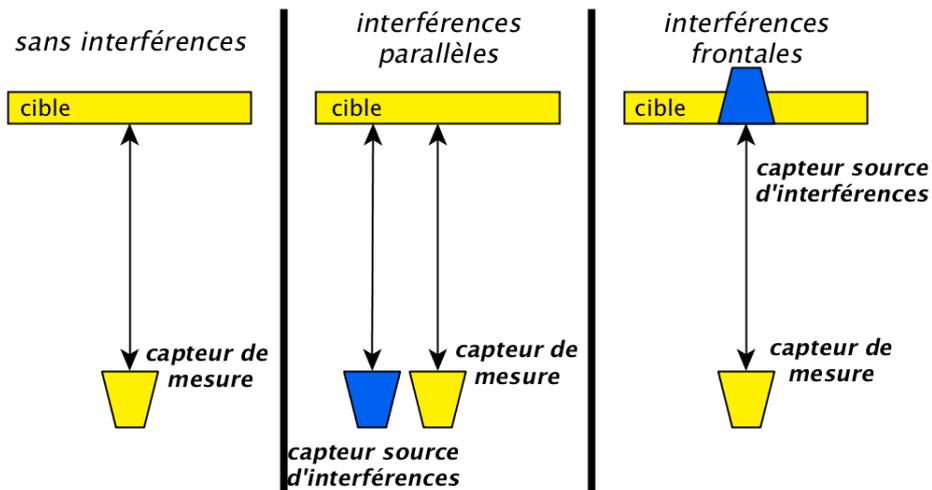


FIGURE 3.34 – Configuration du test visant à déterminer les effets des interférences sur les capteurs de distance.

TABLE 3.4 – Effet des interférences sur les capteurs ultrasoniques. Les valeurs sont en millivolts, avec l'écart-type entre parenthèses.

US	Pas d'interférence	Interférence parallèle	Interférence frontale
50 cm	491 (6)	486 (13)	487 (15)
100 cm	647 (6)	620 (54)	632 (48)
150 cm	795 (6)	736 (103)	747 (96)

vecteurs en entrée est supérieur à 1 (équation 3.9).

$$estimation(x) = \left\{ \sum_i confiance_i(x) * estimation_i(x) \geq 1 \right\} \quad (3.9)$$

Cet algorithme a été évalué sur trois des configurations : papier blanc (W), surface réfléchissante orientée (R/A), et verre (G). Les taux de positifs corrects (PC) et de négatifs corrects (NC) ont été comparés. Le tableau 3.5 montre les différences entre les résultats qui sont obtenus dans deux situations extrêmes : sans information sur le contexte (tous les capteurs utilisés, toutes les confiances à 1), et dans les situations où le contexte est entièrement déterminé (un seul capteur actif).

TABLE 3.5 – Évaluation du taux de détection selon le capteur avec la formule 3.9. Les taux de positifs corrects (PC) et de négatifs corrects (NC) sont comparés.

		<b>W</b>	<b>R/A</b>	<b>G</b>
<b>IR, US, Laser</b>	<b>PC</b>	100.0	96.8	99.4
	<b>NC</b>	99.9	100.0	99.9
<b>IR, US</b>	<b>PC</b>	100.0	96.5	99.4
	<b>NC</b>	99.9	100.0	99.9
<b>IR, Laser</b>	<b>PC</b>	100.0	<b>3.1</b>	<b>0.0</b>
	<b>NC</b>	99.9	100.0	99.9
<b>US, Laser</b>	<b>PC</b>	100.0	96.8	99.4
	<b>NC</b>	100.0	100.0	100.0
<b>IR</b>	<b>PC</b>	99.8	<b>0.0</b>	<b>0.0</b>
	<b>NC</b>	99.9	100	100
<b>US</b>	<b>PC</b>	100.0	96.5	99.4
	<b>NC</b>	100.0	100.0	100.0
<b>Laser</b>	<b>PC</b>	97.0	<b>3.1</b>	<b>0.0</b>
	<b>NC</b>	100.0	100.0	100.0

### 3.5.2 Discussion

Les résultats des tests d'affinité entre capteurs et matériaux ont montré que chaque type a ses propres faiblesses (en gras dans les tableaux 3.3 et 3.4). Le capteur infrarouge en a plusieurs : à 150 cm, une surface noire augmente sensiblement l'écart-type des mesures ; le verre et les deux configurations avec des surfaces réfléchissantes augmentent aussi l'écart-type et, ce qui est plus grave, génèrent un biais de grande ampleur. Le capteur ultrasonique est affecté par l'orientation de la cible et les interférences d'autres capteurs. Le Lidar Lite fournit des mesures précises dans tous les cas, sauf avec la cible en verre.

Il y a trois problématiques distinctes qui doivent être résolues pour effectuer une détection robuste au contexte en temps réel :

- Connaître l'influence des contextes sur les mesures des capteurs ;

### 3.5. CONCEPTION, IMPLÉMENTATION ET TEST DE LA FONCTIONNALITÉ DE DÉTECT

- Savoir identifier le contexte actuel ;
- Savoir comment améliorer les algorithmes en prenant en compte le contexte.

L'expérimentation sur les affinités porte sur le premier point, tandis que l'algorithme simple de détection présenté porte sur la troisième problématique.

Pour être robuste, le système doit donc prendre en compte la variabilité de la qualité des mesures des capteurs en fonction du contexte. Ces tests ne sont pas exhaustifs et visaient à montrer qu'un seul type de capteur n'est pas apte à fonctionner dans toutes les situations. La sensibilité au matériau de surface de l'objet cible n'est qu'un exemple, destiné à mettre en lumière l'importance de la connaissance du contexte. En pratique, avec plusieurs types de capteurs installés, deux stratégies de détection sont possibles. La première est la plus simple : tous les types de capteurs fonctionnent en même temps, et on combine leurs informations pour juger de ce qui se trouve devant la canne.

La seconde stratégie consiste à allumer ou éteindre les capteurs selon la pertinence de leurs informations. Une manière de voir cela est la suivante : en l'absence de capteur adapté à certaines situations (p. ex. : une surface transparente) il est inutile d'utiliser d'autres capteurs. En revanche, si le capteur adapté est présent, on peut éteindre les autres et économiser de l'énergie sans perdre en précision. Par exemple, devant une porte en verre, éteindre le télémètre laser qui consomme 105 mA et le capteur infrarouge qui consomme 33 mA représente un gain appréciable.

Cette simulation d'algorithme de détection d'obstacles montre que la détermination du contexte est une partie primordiale du système. Afin d'y parvenir, nous pouvons utiliser les capteurs embarqués dans le LMSC : capteur de luminosité, de température, d'humidité, caméra du smartphone. Nous pouvons aussi enrichir la carte avec des informations difficiles à déterminer à partir des seules données capteurs, comme la texture des murs, du sol, ou la couleur des portes. Grâce à ces données, notre algorithme devrait être capable de fournir des estimations plus robustes en temps réel, sans utiliser la totalité des capteurs en permanence, et ainsi minimiser la consommation d'énergie du système pour améliorer sa durée de fonctionnement.

La fusion des informations peut prendre des formes diverses. La solution la plus simple, possible quel que soit le nombre de capteurs, est de faire la moyenne de leurs estimations. Elle est peu utile : tant que les capteurs fournissent des estimations proches, la moyenne fournit une information *a priori* pertinente, mais dès lors qu'un cas plus difficile se présente, dans lequel un capteur mesure une distance à l'obstacle radicalement différente des autres, alors l'effet de la moyenne sera de supprimer l'information fournie par les capteurs en donnant une position intermédiaire *a priori* absurde. Si on se trouve dans le cas où certains capteurs donnent une mesure correcte et d'autres une mesure incorrecte, ce qui est très probable, la moyenne sera alors systématiquement incorrecte, et masquera l'information correcte fournie par certains capteurs. Ainsi, si un capteur indique un obstacle à un mètre de distance et un autre à cinq mètres, il est probable qu'un obstacle se trouve soit à un mètre, soit à cinq, mais très peu probable qu'il se trouve à trois mètres comme l'indique la moyenne. Il se trouve que ce type de désaccord est précisément celui où l'on cherche à rendre la détection plus robuste par l'utilisation de capteurs hétérogènes, il faut donc que la méthode de fusion choisie soit pertinente.

Si trois types de capteurs sont présents, on peut utiliser la médiane, c'est-à-dire un vote à la majorité, soit deux capteurs. Cette méthode de fusion est très classique pour la redondance et la tolérance aux fautes, dans toutes les situations où trois composants (p. ex. : des calculateurs) sont présents. Elle ne tient en revanche pas compte des connaissances sur le contexte.

Enfin, on peut utiliser une méthode plus puissante de fusion de données : la théorie de Dempster-Shafer. Celle-ci fonctionne grâce à la prise en compte des combinaisons d'hypothèses atomiques [Wu, 2004]. Ces hypothèses atomiques sont les éléments minimaux indivisibles pouvant être observés. Si un obstacle peut se trouver à 1 m, 3 m, ou 5 m, les combinaisons possibles (ensemble des parties de l'ensemble des hypothèses atomiques) seront par exemple les hypothèses suivantes :

$$\{\{1m\}, \{3m\}, \{5m\}, \{1m \text{ ou } 3m\}, \{1m \text{ ou } 5m\}, \{3m \text{ ou } 5m\}, \{1m \text{ ou } 3m \text{ ou } 5m\}, \emptyset\} \quad (3.10)$$

Pour chaque capteur ou observateur, on peut alors calculer la *croyance* et la *plausibilité* de cet observateur en chaque combinaison d'hypothèses atomiques. La *croyance* correspond à la somme des poids que l'observateur donne à toutes les sous-hypothèses incluses dans l'hypothèse en question, et la *plausibilité* correspond au complémentaire de la somme des poids des hypothèses n'incluant aucune sous-hypothèse de l'hypothèse en question. La probabilité de l'hypothèse se trouve entre la croyance et la plausibilité, qui forment un intervalle de confiance. Des formules permettent ensuite de combiner les jugements de plusieurs observateurs, et de calculer la croyance et la plausibilité du système.

### 3.5.3 Fonctionnement global du 2SEES

La carte uSu-Edu échantillonne tous les capteurs à une fréquence d'environ 50 Hz et envoie ces données par port série au Raspberry Pi. Ce dernier calcule les déplacements par navigation à l'estime toutes les 10 trames, soit 5 Hz. Cela permet de diviser par dix les coûts de communication avec le smartphone et de réduire la fréquence à laquelle l'algorithme de localisation doit fonctionner. Au démarrage de l'application smartphone, celle-ci télécharge la carte depuis un serveur distant et l'affiche sur l'écran<sup>18</sup>. L'utilisateur peut ensuite se connecter en Bluetooth au Raspberry Pi 3 : une fois la connexion établie, le smartphone reçoit les trames des déplacements. À la première trame, l'algorithme de localisation génère des particules-trajectoires et commence le suivi.

Il est possible d'activer le programme de détection d'obstacles en parallèle. Ce dernier extrait les données des capteurs d'obstacles des trames reçues de la carte uSu-Edu et juge de la présence d'obstacles avec la formule 3.9. Un moteur vibrant est fixé à la poignée du LMSC et connecté au Raspberry Pi, qui l'active quand un obstacle est détecté.

Le chapitre suivant est consacré à la troisième fonctionnalité du système 2SEES : la perception environnementale par les objets intelligents.

18. Cette interface n'est pas encore adaptée aux personnes déficientes visuelles, cf. figure 3.35.

### 3.5. CONCEPTION, IMPLÉMENTATION ET TEST DE LA FONCTIONNALITÉ DE DÉTECT

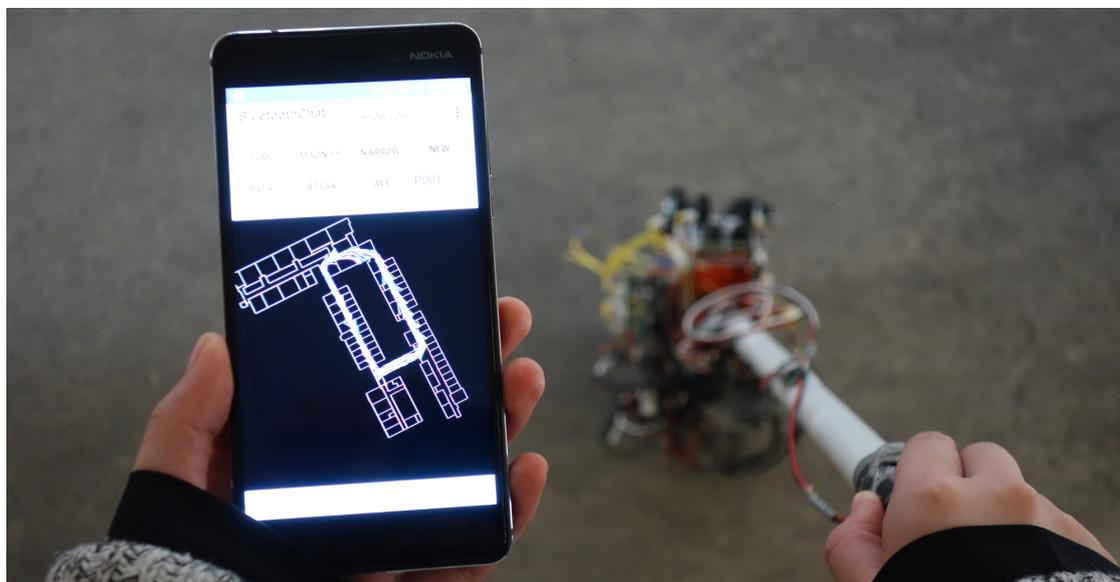


FIGURE 3.35 – L'application smartphone du 2SEES.



# Chapitre 4

## Objets intelligents pour la perception : SO2SEES

SO2SEES, la fonctionnalité présentée dans ce chapitre, est destinée à permettre aux personnes déficientes visuelles d'interagir avec les objets intelligents qui les entourent. Des concepts préliminaires, utiles pour la compréhension du concept et du fonctionnement de SO2SEES, sont d'abord présentés dans la section 4.1. La place de la fonctionnalité à l'intérieur du système SO2SEES y est d'abord décrite. Suit une introduction générale du substrat sur lequel elle s'appuie, c'est-à-dire les objets intelligents, et une présentation plus précise de la question centrale à laquelle nous avons été confrontés : l'interopérabilité. Le concept et l'architecture de la fonctionnalité sont ensuite exposés en détail dans la section 4.2. Celle-ci est assez complexe, car elle implique plusieurs acteurs à la fois : système 2SEES et gestionnaires d'objets intelligents. Le composant essentiel de la fonctionnalité, son interface, est réalisé par des requêtes, qui sont soumises à des contraintes fortes de concision. Deux scénarios ont été conçus pour tester la fonctionnalité. La section 4.3 décrit les objets intelligents que nous avons développés, l'implémentation logicielle de chacun des acteurs, et le déroulement à proprement parler des deux scénarios. Enfin, nous discutons des apports et des questions autour de cette fonctionnalité dans la section 4.4. Celle-ci décrit la manière dont nous nous sommes confrontés aux nombreuses problématiques impliquées : relation au contexte, intersection des bases de connaissances, gestion des erreurs, interopérabilité. Les apports perceptifs espérés de la fonctionnalité y sont également discutés.

### 4.1 Introduction et concepts préliminaires

#### 4.1.1 Assistance à la perception de l'environnement

Le système 2SEES possède trois fonctionnalités : détection d'obstacles, localisation/navigation, et assistance à la perception de l'environnement. Les travaux réalisés sur les deux premières ayant été présentés dans le chapitre précédent, celui-ci

s’attache à exposer ceux qui ont été réalisés sur la fonctionnalité d’assistance à la perception de l’environnement.

Ce concept recouvre une importante diversité d’interactions : tout ce qui est externe à la personne forme son environnement, et la perception environnementale représente la presque totalité de la perception. Nous utilisons volontairement ce terme générique pour désigner l’assistance à la perception et aux interactions avec les objets qui ne rentrent pas dans le cadre des deux autres fonctionnalités. Plusieurs objets et types d’interactions envisageables sont visibles sur la figure 4.1 :

- Interactions avec les systèmes de transports, représentés par une voiture, un tramway, et son arrêt. Les sous-fonctionnalités associées peuvent être la prévention des collisions et la communication des destinations et des horaires des transports en commun ;
- Un passage piéton, qui signale son état ;
- Reconnaissance de signalisation traditionnelle : texte et panneau signalétique visuel d’un commerce (sur la figure, une pharmacie) ;
- Une porte affichant les horaires d’ouverture de la pharmacie ;
- Détection de personnes ;
- Une boîte aux lettres facile à repérer et indiquant les horaires de levée du courrier et le bureau de poste le plus proche.

Tous ces éléments et interactions présentent un point commun : ils sont naturels ou faciles à utiliser pour la plupart des gens<sup>1</sup>, mais sont des tâches non-triviales pour les personnes déficientes visuelles, générant une charge mentale pouvant être importante.

Certains de ces cas ont déjà été étudiés, ainsi que de nombreux autres qui, dans le cadre du découpage en fonctionnalités du 2SEES, relèveraient de la fonctionnalité d’assistance à la perception. On a vu au chapitre 1 des applications de reconnaissance du texte et des panneaux d’indication [Yi and Tian, 2015, Wang and Tian, 2011]. On a aussi mentionné plusieurs types d’appareils d’assistance à la traversée des passages piétons : les applications embarquées par la personne, typiquement sur un smartphone et reposant sur la vision par ordinateur, d’une part [Ivanenko et al., 2008, 2009, Shen and Coughlan, 2012, Coughlan and Shen, 2012, 2013], et les balises sonores sans fil déclenchées par la personne, d’autre part [Alexandre et al., 2018].

Dans ce chapitre, nous nous sommes concentrés sur un aspect peu étudié jusqu’à présent de cet ensemble : les interactions entre la personne et les objets intelligents, c’est-à-dire dotés de capacités de calcul et de communication. Un objet « simple », par défaut non-adapté à l’utilisation par des personnes déficientes visuelles, peut le devenir une fois équipé de l’électronique nécessaire à en faire un objet intelligent. Dans le cadre de la figure 4.1, cela peut s’appliquer à tous les objets. Les informations normalement écrites, même à distance, peuvent être communiquées directement à la personne de manière plus fiable qu’avec une caméra et un logiciel de reconnaissance de texte. Cela est valable pour le panneau de la pharmacie, les horaires sur la porte, les informations à l’arrêt de tramway, celles sur une boîte aux lettres. Les objets peuvent également se signaler : c’est le rôle du signe extérieur

---

1. Sauf peut-être la gestion des horaires de transports en commun.

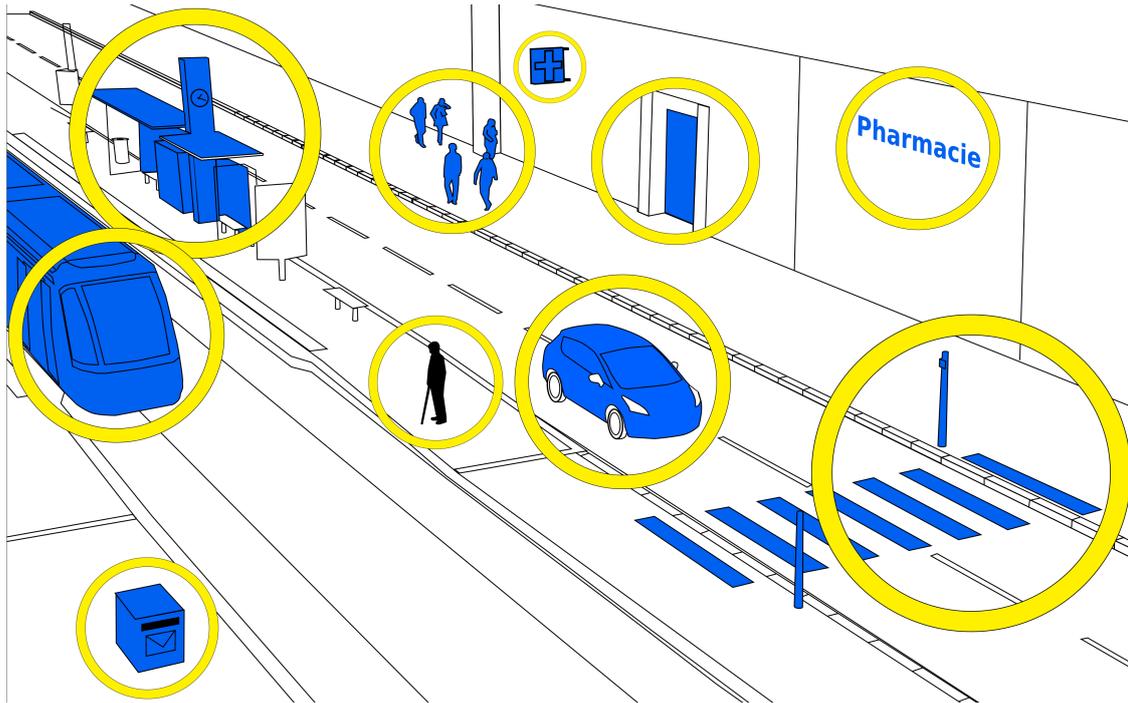


FIGURE 4.1 – Aperçu de la gamme d’interactions possibles par la fonction d’assistance à la perception.

d’une pharmacie et de la forme et de la couleur jaune des boîtes aux lettres.

#### 4.1.2 Objets connectés

L’idée derrière l’*Internet of Things* ou *IoT* (« Internet des objets » en français) est d’étendre la connectivité d’Internet au domaine des objets physiques.

Dans les premières esquisses du concept qui deviendra par la suite l’*Internet of Things*, les *things*<sup>2</sup> en question étaient des objets standards équipés d’un tag RFID pour permettre leur identification [Atzori et al., 2010]. L’aspect dominant du paradigme était alors l’identification des objets.

Un autre concept pionnier a été celui du *pervasive computing* (ou « informatique omniprésente » en français) : des ordinateurs sont présents, invisibles, en grand nombre dans tous les environnements et les objets. L’accent était alors mis sur le développement de systèmes d’exploitation et de logiciels à même d’exploiter la mise en réseau de toutes ces machines [Ibarra-Esquer et al., 2017].

Peu à peu, les conditions techniques ont permis l’émergence réelle de dispositifs tels qu’envisagés par les inventeurs du concept d’informatique omniprésente. La technologie de gravure des processeurs a fait d’énorme progrès, permettant de fabriquer des puces à la fois petites, économes en énergie, et puissantes. La qualité des autres composants a suivi un cheminement similaire ; les écrans, par exemple, sont plus

2. Que nous traduirons par « objets », même si les deux mots ne véhiculent pas les mêmes nuances.

économiques en énergie, plus fins, et d'une qualité plus élevée qu'il y a ne serait-ce que cinq ans. L'informatique omniprésente est extrêmement dépendante de l'autonomie énergétique des appareils, qui conditionne leur aptitude à la mobilité. Un autre aspect de l'informatique omniprésente implique des appareils peu mobiles, voire fixes, déployés dans l'environnement. Là aussi, l'autonomie est importante : dans de nombreux cas, il n'est pas envisageable de déployer une infrastructure énergétique spécifique pour les appareils. Pour envisager une « omniprésence », il faut que ce soit les appareils qui s'adaptent à l'environnement et non l'inverse. Cela passe, entre autres, par l'autonomie énergétique des appareils. Les batteries constituent donc le maillon essentiel de la chaîne énergétique des appareils indépendants ; leurs technologies ont également connu des progrès au cours des vingt dernières années, même si l'intensité du phénomène est loin de ce qu'a vu l'électronique.

Des ordinateurs réellement mobiles sont apparus : les smartphones. Connectés en permanence à Internet, et possédés par une grande partie de la population, ils représentent une première forme d'informatique omniprésente. Les tablettes, sont, en regard des objets connectés, de la même catégorie, bien qu'elles soient généralement utilisées de manière plus sédentaire. Chez les particuliers, smartphones et tablettes ont été suivis par les objets traditionnellement non informatisés, fonctionnellement complexes et de valeur : voitures, télévisions, et électroménager. Les voitures passent souvent par l'intermédiaire du smartphone pour leur connexion au réseau d'un côté et à l'utilisateur de l'autre. Télévisions connectées et électroménager connecté, eux, se rattachent aux réseaux WiFi des foyers. La prochaine étape du processus devrait être celle de la connexion et de l'intelligence des objets non électrifiés et des consommables. Certains en sont déjà équipés : il existe des pneus équipés de capteurs et connectés au smartphone<sup>3</sup>. Les pneus sont certes des consommables complexes et de prix élevé, et la gamme concernée recouvre des pneus spécifiques (sports mécaniques et véhicules industriels) – incomparables avec, par exemple, des couverts jetables. Mais on peut s'attendre à ce que le principe soit prolongé dans un futur proche aux pneus plus « standards » et à d'autres domaines et objets.

En conséquence, la définition communément admise de ce qu'est un *objet connecté* (au sens d'un objet de l'Internet des objets) tend à pencher de plus en plus du côté de la définition des *objets intelligents* (*smart objects* en anglais). Les objets intelligents sont des objets autonomes qui embarquent des capteurs et des actionneurs<sup>4</sup> afin d'interagir avec leur environnement [Fortino et al., 2012]. Leur caractéristique centrale, qui leur donne droit à l'épithète « intelligent », est leur capacité à raisonner : ils embarquent des processeurs et des algorithmes destinés à prendre des décisions en se basant sur les informations qu'ils manipulent. Cette capacité de décision n'est pas une évidence ; les réseaux de capteurs sans fil « traditionnels » suivent une architecture où les rôles sont clairement répartis : les nœuds-capteurs sont chargés de collecter les données et de les transmettre (via, en général, un ou plusieurs échelons intermédiaires de propagation) à une machine qui, elle, raisonne

---

3. Pneus Michelin Track Connect : <https://www.michelin.fr/auto/accueil-auto/michelin-track-connect> (accès le 31 octobre 2019).

4. Les actionneurs sont des dispositifs générant un phénomène physique en réponse à un signal : moteurs, freins, composants émettant de la lumière ou du son, etc.

sur les données et prend les décisions. Bien entendu, pour participer à *l'Internet* des objets, il est nécessaire que les objets intelligents possèdent des capacités réseau et soient connectés à Internet.

Dans notre travail, nous avons choisi de ne pas utiliser une définition spécifique de l'Internet des objets. Nous avons plutôt adopté une approche *de facto* : pour les aspects en lien avec l'Internet des objets, nous avons choisi de nous reposer sur une caractéristique centrale et largement partagée des environnements de l'Internet des objets, à savoir l'utilisation d'une plate-forme *back-end cloud* pour centraliser les données générées par les objets.

Il est à noter que le nombre d'objets est en augmentation rapide depuis plusieurs années et qu'il est probable que cette tendance se poursuive [Gubbi et al., 2013].

### Métadonnées et ontologies

Les métadonnées sont les données décrivant des caractéristiques d'autres données. La date de publication d'un livre, la taille d'un fichier, la langue dans est écrit un fichier (ou un livre) sont des métadonnées. Quand des systèmes se transmettent des données, ces informations sont extrêmement utiles : comment peuvent-ils fonctionner, par exemple, s'ils ne savent pas quel type de données est contenu dans un fichier ? Ils peuvent essayer de reconnaître le format, mais cela est peu fiable et complexe ; ils peuvent connaître le format *a priori* – typiquement, s'ils ne se transmettent que des fichiers d'un seul format, ils n'ont pas à s'attendre à un autre format – mais cela est peu souple. La présence de métadonnées associées aux données offre un mécanisme fiable, souple et efficace pour résoudre ce type de problématique. Elles sont donc un élément central de l'interconnexion des systèmes. De nombreuses métadonnées sont utilisées dans les systèmes informatiques : négociation dans les protocoles réseaux, types MIME, données EXIF, etc. Le web sémantique est l'extension du web classique avec des métadonnées sémantiques dans des formats standardisés, permettant l'utilisation des informations par les machines et les algorithmes. C'est un concept très lié au web des données (*semantic web* en anglais), qui est la mise en réseau des données, l'interconnexion des silos<sup>5</sup> de données existants.

Afin de permettre et de faciliter l'intégration avec les écosystèmes existants et futurs de l'Internet des objets, nous nous sommes efforcés d'adopter des standards répandus. En particulier, nous utilisons les standards du web sémantique du W3C comme base pour la représentation, l'intégration et l'extraction d'informations et de connaissances : RDF, RDFs et OWL.

RDF, le *Resource Description Framework*, constitue la brique de base du W3C pour la construction du web sémantique [Wood et al., 2014]. RDF propose un modèle abstrait de relation générique entre des éléments nommé « triplet ». Un triplet RDF est constitué d'un sujet, d'un prédicat et d'un objet :  $\{\textit{sujet}, \textit{prédicat}, \textit{objet}\}$ . Le prédicat représente le lien entre le sujet et l'objet, c'est une propriété

---

5. Un silo informationnel est un système d'information dans lequel il n'y a pas d'interaction avec l'extérieur.

qui les lie. Ces triplets s'assemblent entre eux pour former un graphe décrivant les relations qui existent dans un ensemble de concepts. Pour la définition de triplets RDF en pratique, plusieurs syntaxes existent : les principales sont RDF/XML, Notation3, Turtle, N-triples, et JSON-LD dont la standardisation se stabilise. Dans nos travaux, nous avons principalement utilisé la notation Turtle pour nos données RDF, pour sa facilité d'usage.

Un point essentiel de RDF pour sa contribution au web des données et au web sémantique est l'utilisation d'identifiants *Internationalized Resource Identifiers* (IRI) ou *Uniform Resource Identifiers* (URI)<sup>6</sup> et d'espaces de nommage. Les IRI sont des identifiants qui peuvent être utilisés pour des ressources arbitraires [Berners-Lee et al., 2005, Duerst and Suignard, 2005]. Elles peuvent être des *noms*, des *locateurs*, c'est-à-dire des *Uniform Resource Locators* ou URL (la forme classique d'IRI rencontrée sur le web), ou bien les deux à la fois [Berners-Lee et al., 2005]. Elles sont utilisées par RDF pour identifier les concepts<sup>7</sup>. Cela permet de faciliter la connexion du web sémantique avec le web classique, qui repose sur les URL, et fournit un moyen simple pour lier les données entre elles : un concept est accessible en suivant son IRI.

*RDF Schema* ou RDFS est un vocabulaire établi par le W3C définissant des concepts génériques de modélisation des données : classes, sous-classes, valeurs littérales (chaînes de caractères, nombres, etc.) et types de données, propriétés, sous-propriétés, etc.

Le langage *Web Ontology Language* ou OWL, également du W3C, est une surcouche de RDFS et RDF qui permet de définir des ontologies plus riches que RDFS seul [W3C OWL Working Group, 2012]. OWL étend les possibilités de RDFS en définissant des concepts supplémentaires : équivalence entre des classes ou des propriétés, symétrie des relations, égalité entre des ressources, etc. OWL autorise la création d'ontologies expressives (et complexes) et permet à des systèmes de déduction de faire des raisonnements plus poussés sur les ontologies.

Les espaces de nommage (*namespaces* en anglais) sont le mécanisme qui permet à des vocabulaires de définir les mêmes concepts de manière différente sans que cela induise une incompatibilité. Ainsi, les espaces de nommages autorisent le développement d'un grand nombre de vocabulaires de manière décentralisée. Ils sont généralement notés par un identifiant court suivi de deux points, préfixant l'identifiant des concepts manipulés : `rdf`, `rdfs:`, `owl:`, `foaf:` (ontologie *Friend of a Friend*), `dc:` (ontologie *Dublin Core*) sont des exemples d'espaces de nommages très communs. Ainsi, il est possible d'avoir un concept `rdfs:Class` coexistant avec un concept (imaginaire) `dc:Class` : les deux sont discernables. Les identifiants comme ceux que nous venons de lister sont des versions courtes ; les espaces de noms complets ou réels sont des IRIs qui identifient des vocabulaires, et qui sont

---

6. Dans ce travail, nous utilisons des IRIs ; nous ne mentionnerons plus les URIs mais ce qui s'applique aux IRIs peut aussi s'appliquer aux URIs, les premiers étant essentiellement des versions des seconds prenant qui prennent en compte les caractères internationaux.

7. D'autres concepts sont des littéraux, c'est-à-dire des valeurs ayant un type. Il est également possible de ne pas identifier certains concepts, qui s'appellent alors des « nœuds vides » ou *blank nodes*.

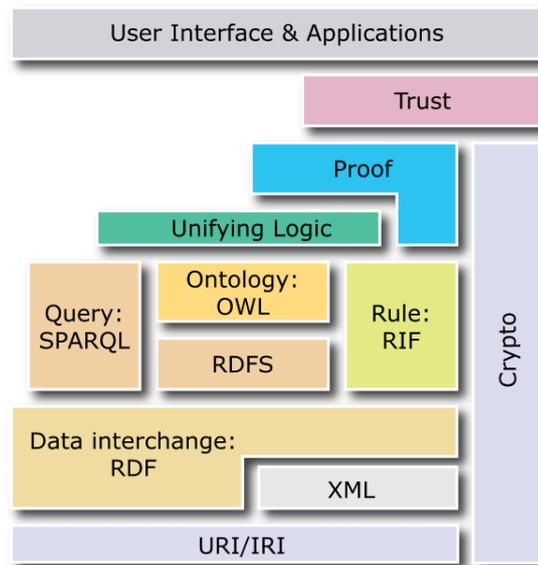


FIGURE 4.2 – Pile « classique » des standards et langages du web sémantique du W3C. Source : Marobil (CC0) ([https://en.wikipedia.org/wiki/File:Semantic\\_web\\_stack.svg](https://en.wikipedia.org/wiki/File:Semantic_web_stack.svg))

presque toujours des URLs pointant vers des documents RDFS ou OWL définissant le vocabulaire [Powers, 2003]. Les URLs des exemples précédents sont les suivantes<sup>8</sup> :

- rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> ;
- rdfs: <http://www.w3.org/2000/01/rdf-schema#> ;
- owl: <http://www.w3.org/2002/07/owl#> ;
- foaf: <http://xmlns.com/foaf/0.1/> ;
- dc: <http://purl.org/dc/elements/1.1/>.

Il est possible de créer un espace de nommage. Dans la suite, nous désignerons par **externe:** les espaces de nommages hypothétiquement définis par des acteurs externes, et par **2sees:** celui du système 2SEES.

Le langage *SPARQL Protocol and RDF Query Language* ou SPARQL permet de faire des requêtes et de manipuler des graphes RDF (cela inclut donc les langages basés sur RDF comme RDFS et OWL) : à la manière du SQL pour les bases de données relationnelles, SPARQL permet l'extraction de données, l'ajout, etc. La pile « classique » des standards du web sémantique est visible sur la figure 4.2. La pile que nous utilisons en pratique pour la fonctionnalité SO2SEES est visible sur la figure 4.3.

Un composant essentiel de la fonctionnalité SO2SEES est également basé sur un standard du W3C, non encore finalisé celui-là : le standard *Web of Things – Things Description* (WoT-TD), que l'on pourrait traduire par « web des objets – descrip-

8. À partir d'un navigateur web, on accède souvent à une définition « pour humains » des ontologies plutôt qu'à un fichier RDF sérialisé, car un mécanisme de redirection sur le serveur HTTP qui écoute les requêtes de ces URL peut parfois proposer une version textuelle adaptée aux humains quand ils détectent que l'accès vient d'un navigateur web.

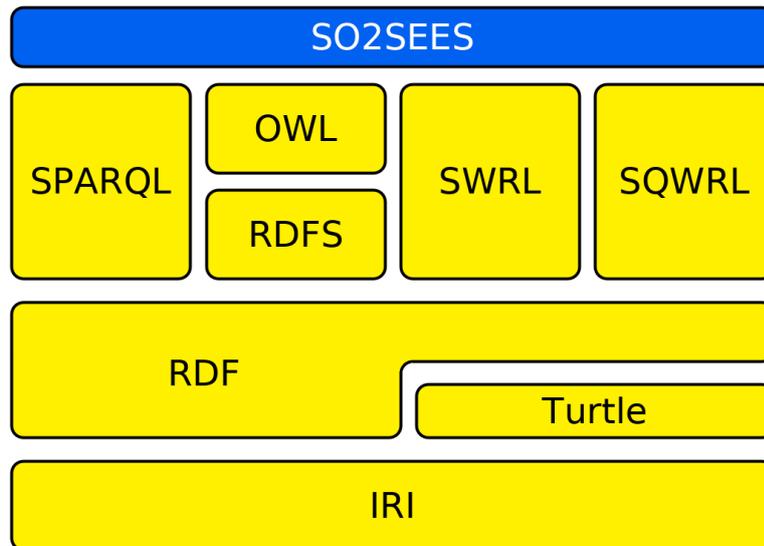


FIGURE 4.3 – Standards du web sémantique utilisés par la fonctionnalité SO2SEES.

tion d’objet » [Kaebisch et al., 2019a]. Comme son nom l’indique, ce standard définit un modèle formel pour la description des objets. Il est développé pour faciliter les interactions entre les objets intelligents, et les descriptions qu’il fournit sont donc principalement faites du point de vue des applications logicielles. En cela, WoT-TD est très différent des ontologies pré-existantes de description des objets connectés du W3C, SSN et son extension SOSA [Haller et al., 2017]. Ces deux ontologies sont destinées à la description des réseaux de capteurs sans fil. Elles sont orientées vers la facilitation de l’exploitation des données recueillies par les réseaux de capteurs, en fournissant le plus précisément possible les caractéristiques des capteurs et leurs interactions avec leur environnement. Elles permettent de modéliser quelle caractéristique de quel phénomène est observée par le capteur (p. ex. : la taille d’un arbre), la manière dont les observations sont faites, à quelle fréquence d’échantillonnage, etc. Ces deux vocabulaires ne sont pas limités aux seuls capteurs, ils permettent aussi de prendre en compte les actionneurs avec leurs actions possibles, et les effets de leurs actions sur l’environnement. Des propriétés de qualité des données sont également incluses (précision, latence, dérive, etc.) ainsi que des outils pour travailler sur le cycle de vie des objets (maintenance, changement des batteries, etc.). Ces deux ontologies sont très complètes.

WoT-TD a une approche très différente. Premièrement, son approche est d’abord celle d’un standard de description des objets intelligents avant d’être une ontologie ; sa définition préférentielle est en JSON et non RDF/XML ou Turtle. Ensuite, elle n’est pas focalisée sur les capteurs environnementaux et les observations scientifiques, mais couvre plutôt les objets connectés dans des domaines très variés : applications grand public, industrielles, transports, agriculture, surveillance de l’environnement, etc. [Kovatsch et al., 2019]. Son vocabulaire reflète cet état de fait : seule une minorité est dévolue aux interactions avec l’environnement. La majorité est dédiée à la définition des interfaces « virtuelles » de l’objet : comment peut-on interagir informatiquement avec l’objet ? Quelles données expose-t-il, et quels formulaires, quels protocoles (HTTP, CoAP) permettent d’y accéder ? Quel est le

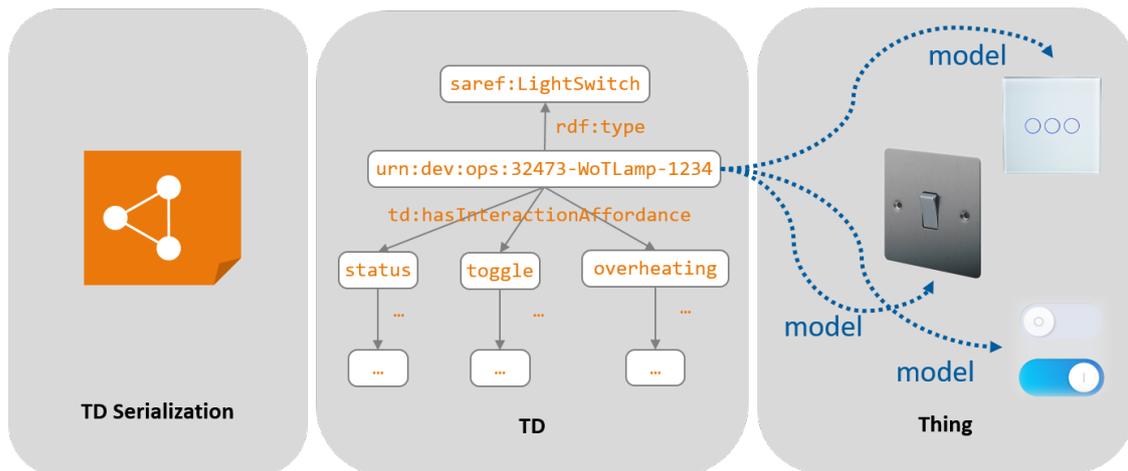


FIGURE 4.4 – Intégration de WoT-TD avec une ontologie plus large (source : [Kaebisch et al., 2019a]; Copyright © [2019] World Wide Web Consortium, (MIT, ERCIM, Keio, Beihang). <http://www.w3.org/Consortium/Legal/2015/doc-license>).

mécanisme de sécurité protégeant l’objet et ses données ? La description de ce que fait l’objet n’est pas formellement couverte par le vocabulaire de WoT-TD en profondeur. SOSA et SSN permettent de modéliser comment un objet s’intègre dans son environnement physique et matériel, tandis que WoT-TD permet d’expliquer comment un objet connecté s’intègre dans un environnement virtuel.

WoT-TD est n’importe comment capable de s’intégrer dans une ontologie plus large qui modélise l’environnement [Kaebisch et al., 2019a]. Cela peut être SOSA et SSN, cela peut être SAREF comme décrit dans l’exemple du W3C (figure 4.4), ou bien peut-être une autre ontologie destinée à une application spécifique, comme la nôtre. Le standard WoT-TD correspond bien à notre application : orienté vers les objets et les usages « de la vie de tous les jours », il est construit sur les standards W3C de web sémantique. Nous avons besoin de découvrir les objets environnants et de pouvoir interagir avec eux, et le second point est précisément ce que propose WoT-TD. De plus, il ne s’occupe pas des considérations bas niveau, ce qui évite de poser des contraintes additionnelles sur, par exemple, les protocoles réseaux utilisés, ou les ressources matérielles des objets impliqués (processeur, mémoire).

## 4.2 SO2SEES

### 4.2.1 Concept

Le concept de la fonctionnalité SO2SEES est de servir d’interface entre une personne déficiente visuelle et les objets intelligents qui l’entourent. Cela devrait permettre d’apporter à la première une partie des informations des seconds. Idéalement, ces informations devraient constituer une assistance à la perception : la fonctionnalité fournit une assistance à la perception *des* objets intelligents, qui, nous

l'espérons, peut devenir une assistance à la perception *par* les objets intelligents.

### 4.2.2 Architecture globale et fonctionnement

L'architecture du système SO2SEES implique plusieurs acteurs, chacun regroupant plusieurs composantes (figure 4.5). Ils ont d'un côté, un échelon<sup>9</sup> local, et de l'autre, un *back-end*, fonctionnant typiquement sur une plate-forme dans le nuage (en anglais *cloud computing*). L'acteur central du SO2SEES est le système 2SEES ; son échelon local inclut le LMSC et un smartphone exécutant une application spécifique qui fait office à la fois d'interface utilisateur, de connexion à Internet et au *back-end cloud* du 2SEES, et de plate-forme de traitement des données. Les autres acteurs sont les fabricants, vendeurs et gestionnaires d'objets intelligents. Ce sont toutes les entités sous la responsabilité desquelles se trouvent les divers aspects des objets intelligents disséminés dans l'environnement. Les concepteurs et fabricants déterminent le fonctionnement des objets, les « gestionnaires » sont ceux qui déploient les objets, les cartographient, et souvent en exploitent les données. Les vendeurs peuvent remplir une partie de ces rôles et avoir une influence sur les objets et sont donc inclus dans la même catégorie. Nous ne nous soucions pas de la manière, certainement variable, dont ces rôles sont répartis sur les divers acteurs impliqués ; nous les cataloguons simplement comme « acteurs externes » et interagissons avec leurs services.

Pour chacun d'entre eux, on supposera une structure *verticale, ouverte et standardisée*. L'aspect vertical vient de ce que chaque acteur repose sur ses propres protocoles, depuis les protocoles bas-niveau jusqu'au *cloud*. Le principe opposé, un écosystème *l'horizontal*, supposerait que les objets d'un fabricant soient, par exemple, être intégrables dans l'infrastructure d'un autre ; on ne fera pas cette hypothèse ici, car elle risquerait de contraindre fortement les possibilités de réalisation concrète. Un seul étage est supposé interopérable, formant un « ruban horizontal » : le niveau des données stockées dans le *cloud*. Nous faisons donc l'hypothèse qu'au moins une partie des données générées et manipulées par les objets intelligents et stockées dans le *cloud* sont accessibles par des mécanismes publics, voire standardisés. Nous avons choisi de nous baser sur les *IRIs*. De là viennent les qualificatifs de *standard* et *ouvert* appliqués à la structure des acteurs externes. Les IRIs ont été choisies car elles sont un standard déjà largement adopté, en raison de leur rôle prépondérant sur le web.

Une autre contrainte imposée par le système SO2SEES sur les acteurs externes est la possibilité pour l'échelon local du 2SEES (LMSC et smartphone) d'obtenir des descriptions formelles des objets environnants (figure 4.6). Cette découverte peut s'effectuer au travers d'un processus local, ou bien au travers de connaissances publiques. Dans le premier cas, les appareils voisins s'échangent directement des messages réseaux, tandis que dans le second, un annuaire géographique référence les objets. Le mécanisme exact a peu d'importance ; il faut seulement que le système 2SEES ait connaissance des objets connectés se trouvant à proximité de la personne

---

9. Nous utiliserons le mot « échelon » pour exprimer le sens du mot anglais *tier*, qui désigne une couche logicielle ou un composant de niveau architectural.

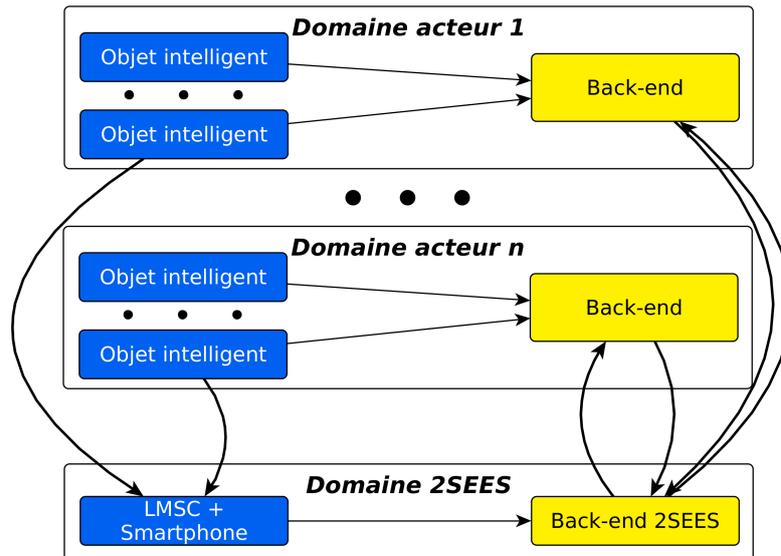


FIGURE 4.5 – Acteurs et plates-formes impliquées dans la fonctionnalité SO2SEES.

qui l'utilise, puisse accéder à leurs identités, et éventuellement, à leurs données. Pour les descriptions formelles, nous avons choisi le standard WoT-TD décrit plus haut, car son vocabulaire correspond bien à notre cas d'utilisation, permettant de décrire quelles données sont accessibles et les moyens d'y accéder.

Lors de la navigation, l'échelon local du 2SEES va rencontrer des objets intelligents et découvrir les IRIs qui leur sont rattachées. Le smartphone va alors envoyer ces IRIs au *back-end cloud* 2SEES, qui s'en servira pour obtenir les descriptions formelles des objets. C'est à ça que servent les IRIs dans le SO2SEES et c'est là qu'intervient la contrainte de relation entre les IRIs et les descriptions formelles : d'une manière ou d'une autre, chaque IRI doit permettre l'accès à une description. Cela peut être par déréréfencement direct, à la manière des URL web, où l'IRI est un hyperlien HTTP qui pointe sur la description. Cela peut passer par une étape préliminaire de traitement de l'IRI, par exemple pour en extraire une URL pointant sur la description. On peut également envisager que l'accès se fasse par l'intermédiaire d'une recherche dans une base de données.

Une fois que le *cloud* 2SEES a obtenu la description des objets, il peut les analyser afin d'extraire les indices contextuels qui vont permettre de déterminer quelles requêtes sont pertinentes par rapport au *contexte global*. Ce que nous entendons par contexte global est l'union de plusieurs sous-contextes :

- Un sous-contexte est extrait des objets intelligents environnants : les éléments et les concepts auxquels les objets sont reliés sont extraits de leurs descriptions formelles ;
- Un autre sous-contexte est composé des habitudes générales de la personne qui utilise le système SO2SEES, que le système aura analysé, et de ses préférences, qu'elle aura explicitées ;
- Enfin, une troisième partie du contexte est formée par la situation actuelle de la personne : que fait-elle ? que veut-elle ? où va-t-elle ? etc.

Une liste de requêtes pertinentes doit être générée à partir de ces éléments. Cela

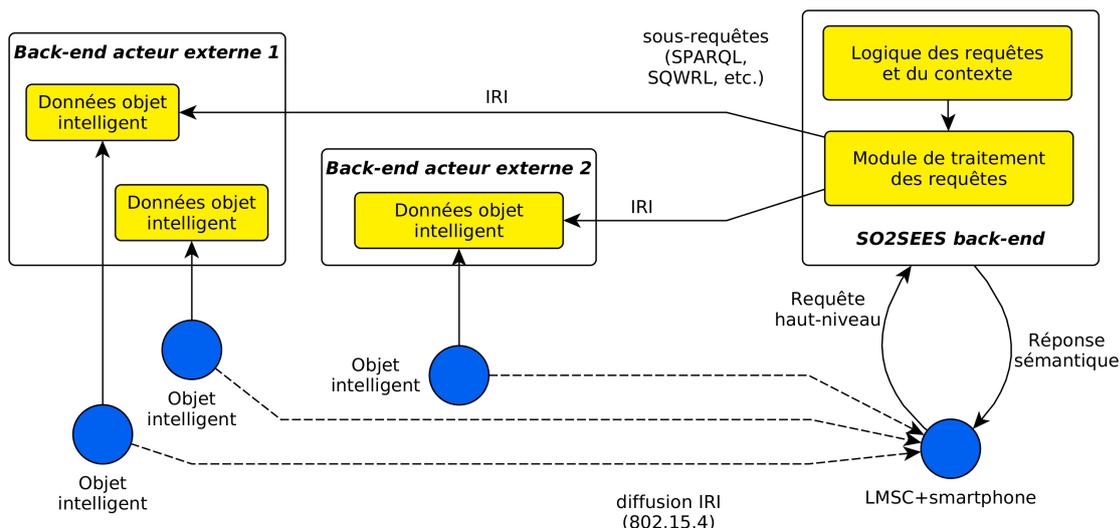


FIGURE 4.6 – Modèle des interactions entre les composantes intervenant dans la fonctionnalité SO2SEES.

peut se faire dans une approche montante, avec des méthodes d'intelligence artificielle, de linguistique, et de sémantique. Plus simplement, une approche descendante peut être employée : le programme d'analyse sélectionne les requêtes parmi une liste pré-existante dont il évalue la pertinence par rapport au contexte.

Une fois la liste de requêtes générée par le *back-end*, celui-ci les envoie à l'interface utilisateur qui s'exécute sur le smartphone. La personne utilisant le système est alors informée de la disponibilité de requêtes et peut les écouter, en choisir une, et l'exécuter. Lorsqu'une requête est choisie, le smartphone transmet le signal au *back-end*, qui va alors exécuter la procédure de la requête. Cette procédure est déjà définie, soit parce que la requête était pré-définie, soit parce qu'elle vient d'être générée dynamiquement par le *back-end*. Typiquement, l'exécution de la procédure d'une requête impliquera de faire des sous-requêtes aux plates-formes *back-end* des autres acteurs, qui gèrent les objets connectés, si les informations nécessaires ne sont pas disponibles dans les descriptions formelles desdits objets. La section suivante explore les questions liées aux requêtes et à l'évaluation du contexte.

### 4.2.3 Fonctionnalités, requêtes, et règles

Les informations obtenues grâce aux objets intelligents permettent à la fois au système 2SEES d'améliorer ses fonctionnalités de base et d'en ajouter de nouvelles.

#### Objets intelligents pour la perception spatiale

La perception spatiale est comme on l'a vu, un élément central de la mobilité des personnes déficientes visuelles. Les objets intelligents environnants peuvent apporter des informations spatiales à une personne de plusieurs manières. Ils peuvent

tout d'abord participer au mécanisme de localisation du LMSC. Le positionnement radio est une méthode de localisation très largement utilisée, et puisque les objets intelligents sont la plupart du temps équipés d'émetteurs-récepteurs radio, ils peuvent être utilisés comme balises radios. On l'a vu, les systèmes satellitaires (Galileo, GLONASS, GPS, etc.) ont une couverture parfois mauvaise et souvent absente à l'intérieur des bâtiments. La présence d'objets intelligents dans un bâtiment fournit une infrastructure à *destination externe* et *déjà installée* (cf. classification donnée dans le chapitre 2, section 2.3.6), en somme idéale car gratuite. L'avantage des objets intelligents est qu'étant donné leur popularité de plus en plus grande, il est probable qu'ils forment dans un futur proche un réseau dense de nœuds. L'exploitation de stations radio, et particulièrement de points d'accès WiFi, est une méthode répandue dans la localisation en intérieur [Lymberopoulos et al., 2015]. L'utilisation d'objets intelligents nécessite tout de même plusieurs aménagements par rapport à un usage « naturel » :

- La position des objets intelligents doit être connue. Cela implique une entrée de la part des utilisateurs ou, alternativement, une auto-localisation par rapport à d'autres points connus, cette seconde solution étant techniquement plus complexe et reposant, à un moment, sur la localisation de certains éléments ;
- Cette position doit être transmise au système 2SEES. Cela peut passer par un référencement cartographique, mais est peu pertinent pour certains objets, très mobiles. Plus logiquement, la position des objets devrait faire partie de leur description formelle ;
- Pour implémenter des mécanismes de positionnement plus précis, des technologies additionnelles peuvent être nécessaires : temps de vol, angle d'arrivée, etc. On en a un aperçu dans l'article de Lymberopoulos et al. : le WiFi est souvent complété [Lymberopoulos et al., 2015]. Les solutions, efficaces, reposant sur la cartographie des signaux Wi-Fi (*fingerprinting* en anglais), nécessitent elles aussi un travail supplémentaire de cartographie.

La cartographie représente en tout cas une contrainte importante pour la localisation en intérieur [Youssef, 2015].

Deuxièmement, les objets intelligents peuvent être utilisés pour l'amélioration directe de la représentation spatiale mentale des personnes, et ce, à travers plusieurs mécanismes. La fonction d'un lieu est très étroitement liée aux objets qui s'y trouvent, ce qui rend le lien entre les objets et les lieux facile à retenir. Par exemple, une des composantes de l'identification des lieux est leur fonction sociale [Declerck and Lenay, 2018] : une pièce peut par exemple être une cuisine, un bureau, ou bien une boutique. Mentalement, l'association des lieux avec leurs rôles aide grandement à s'en souvenir, en particulier si on peut se relier soi-même au lieu et à la pièce. Les objets intelligents peuvent donner cette information et donc être utilisés comme indicateurs du rôle des lieux.

Par exemple, une cuisine occidentale standard comporte un réfrigérateur, une plaque de brûleurs ou un plaque chauffante et un four micro-ondes ; on trouvera, dans un bureau, un ou plusieurs ordinateurs et peut-être une imprimante. Les objets qui se trouvent dans la cuisine ont peu de chances d'être installés dans un bureau, et vice-versa. Si de tels objets sont détectés par le système 2SEES, ce der-

nier peut déduire le rôle de la pièce par l'entremise d'un algorithme cognitif simple. La détection des objets peut se faire de manière optique, avec une caméra et un algorithme de reconnaissance d'objets, mais cette méthode est lourde en calcul, comparativement peu fiable, et pas toujours possible car les objets ne sont pas toujours dans le champ de vision de la caméra. En comparaison, la généralisation des objets intelligents et leur identification par la fonctionnalité SO2SEES permettrait une identification certaine, avec un coût énergétique très faible et peu de contraintes physiques. Il n'est pas nécessaire que l'intégralité des objets d'une pièce soient des objets intelligents, il suffit que deux ou trois le soient pour que le système puisse déterminer le rôle de la pièce. Par contre, ce qui est nécessaire, c'est la présence d'un moyen pour le système de distinguer les espaces et de déterminer comment les objets intelligents leurs sont liés, c'est-à-dire identifier dans quel espace se trouve chaque objet intelligent.

Outre la mise en valeur du rôle des lieux, les objets intelligents peuvent être utilisés comme repères aux points de décision. Naturellement, les humains utilisent les objets qu'ils détectent aux endroits où ils doivent prendre une décision de navigation, c'est-à-dire les endroits où il est nécessaire de faire un choix entre plusieurs chemins possibles, dont l'exemple typique est celui des intersections [Janzen and Van Turenout, 2004]. Les associations entre objets et points de décision peuvent émerger de manière inconsciente, à travers l'implication du gyrus parahippocampique [Janzen and Van Turenout, 2004]. Elles peuvent également émerger de manière consciente, à travers l'usage du langage, par exemple : « Je dois tourner à gauche après l'arrêt de bus ». Ce type d'assistance présuppose une perception de la position spatiale de l'objet plus précise que la simple identification de la présence dans une pièce – en tout cas pour la plupart des situations. Plusieurs niveaux de localisation des objets, absolue et relative à la personne sont envisageables :

- L'identification de la position la plus basique est celle de la *proximité* : on sait que l'objet est dans le voisinage de la personne, sans savoir précisément à quelle distance. Cette connaissance est accessible par défaut : si l'objet est détecté, c'est qu'il est à proximité. Selon le protocole radio, on peut estimer que l'objet est à moins d'une certaine distance ;
- L'utilisation du *RSSI* (*Received signal strength indicator* ou indicateur de puissance du signal reçu) permet d'affiner un peu l'estimation de distance. En effet, la puissance de signal reçue dépend de la distance entre les deux stations. Malheureusement, elle dépend également d'autres facteurs : le type et l'orientation des antennes, les obstacles se trouvant entre les deux stations, les rebonds du signal sur les murs (transmission multi-chemins), etc. Les protocoles de communication modernes, qui utilisent plusieurs antennes rendent encore moins claire la part qu'occupe la distance dans la valeur du RSSI ;
- Un degré plus élevé de précision peut être atteint par l'encodage direct de la position de l'objet intelligent dans les métadonnées qui lui sont attachées, dans notre cas, par exemple, dans sa description formelle . On peut envisager au moins deux stratégies pour l'encodage de cette position dans les métadonnées. La première est plutôt orientée vers l'interfaçage avec des systèmes électroniques et informatisés : elle nécessite la connaissance formelle de la po-

sition de l'objet. Cette position peut avoir été obtenue par auto-localisation de l'objet dans son environnement radio – solution évoquée plus haut, qui nécessite un protocole d'auto-localisation adéquat et un réseau dense et déjà localisé – ou bien par le relevé positionnel de l'objet lors de son installation – solution qui augmente considérablement le déploiement d'objets intelligents. La deuxième stratégie, elle, est plutôt orientée vers l'interfaçage avec des personnes. Elle repose en effet sur la description linguistique de la position de l'objet intelligent. Ainsi, un thermostat aurait inscrit dans ses métadonnées qu'il se trouve « fixé sur le mur du fond de la pièce, à 1,4 m de l'angle avec le mur de gauche ». Ce mode de définition de la position a deux avantages qui découlent de la facilité d'interaction qu'il apporte pour les personnes : il est directement exploitable par la personne à qui est communiquée l'information, et il est facile à définir pour celle qui installe l'objet. À grande échelle, cela implique très probablement une standardisation du vocabulaire spatial employé pour garantir la clarté et l'uniformité des descriptions, mais cette solution reste plus simple que la localisation absolue. Chacune de ces deux stratégies favorise un type d'interfaçage au détriment de l'autre : pour rendre une position cartésienne intelligible pour une personne, il faut soit un module automatisé de description sémantique de la position, qui fonctionne avec la carte, soit une interface haptique adaptée, dont une réalisation efficace est très délicate. De l'autre côté, la transformation d'une localisation encodée en langage naturel dans un système formel de coordonnées est également une tâche complexe. Dans l'idéal, la position serait encodée dans les deux modes, permettant ainsi l'utilisation à la fois par des humains et par des systèmes informatisés. Le surcoût de cette double approche est potentiellement peu élevé par rapport au coût de base de l'une ou de l'autre des stratégies : le coût marginal de la mobilisation d'une personne pour la cartographie *et* la description naturelle de la position étant par exemple probablement négligeable par rapport à celui de la mobilisation de la même personne pour une seule de ces tâches ;

- La dernière solution est un compromis : elle permet une bonne précision dans la localisation, mais n'est pas fiable à 100 % ; elle ne nécessite pas d'intervention spécifique de calcul ou de description de la position lors de l'installation de l'objet, mais n'est pas sans contraintes car elle repose sur une base de donnée définie *a priori* par le fabricant référençant les objets, et présuppose un accès standardisé à cette base – cette contrainte pouvant cependant être vue comme une extension de la contrainte que nous présupposons dans le cadre du SO2SEES de la disponibilité et de l'accessibilité de descriptions formelles des objets. La méthode est la suivante : lorsque le système 2SEES (ou autre) détecte la présence d'un objet et l'identifie, il va chercher dans la base de données publique en question les critères permettant d'identifier visuellement l'objet dans l'environnement. Ces critères peuvent être simplement constitués par l'apparence de l'objet, stockée sous forme d'images. L'apparence de l'objet connue, le 2SEES sera alors capable de rechercher une occurrence de l'objet dans les images de la caméra, sans avoir à passer par des considérations d'apparence générique de l'objet<sup>10</sup>. La localisation

---

10. Pour clarifier, considérons l'alternative : pour rechercher un thermostat, le système pourrait

relative de l'objet par rapport à la canne pourra alors être réalisée à partir de la connaissance de la position et de l'orientation de cette dernière. Un standard d'identification reposant sur les QR codes apposés sur les objets rendrait toute cette procédure bien plus efficiente, en raison de la robustesse de décodage qu'offrent les codes par rapport à la reconnaissance d'objet, qui est plus vulnérable aux variations de lumière ambiante et à l'orientation, et plus intensive en calculs. Un tel standard permettrait en fait de se passer de l'aspect connecté des objets pour leur localisation ; mais la connexion au réseau apporte des possibilités importantes d'économies d'énergie. Dans l'esprit de l'*active multi-sensor context awareness*, la connexion réseau permet de n'allumer la caméra, génératrice d'une importante consommation d'énergie, que lorsque des objets pertinents ont été détectés.

Même avec de tels aménagements pour augmenter la précision de la localisation, le cadre et la situation autour des personnes déficientes visuelles et des objets intelligents reste cependant différent de celui des personnes voyantes et de l'identification visuelle des objets. La perception de la position d'un objet intelligent par une personne déficiente visuelle à travers le SO2SEES est bien moins précise spatialement que la vision de la position d'un objet.

Si leurs positions sont connues, les objets intelligents peuvent être utilisés pour fournir une information spatiale détaillée de l'organisation d'une pièce à une personne déficiente visuelle. Lorsqu'elle entre dans une pièce, celle-ci pourrait profiter de la connaissance de l'organisation de la pièce offerte par la liste des objets intelligents les plus pertinents spatialement. En première approche, cette pertinence spatiale peut être estimée par un examen de la géométrie : les bureaux, tables, chaises et, plus généralement, les meubles participent à l'organisation d'une pièce – les autres éléments importants étant la forme de la pièce et la disposition des portes et des fenêtres. Le nom des objets, accompagnés de leurs positions relatives par rapport à la personne, permettent de constituer une liste informative : « Il y a une armoire à un mètre à votre gauche, une table à deux mètres devant, et un bureau à deux mètres à droite. ». On se trouve évidemment ici aussi en prise avec les problématiques de localisation des objets intelligents, de choix de stratégies orientées vers l'interaction avec les humains ou bien les systèmes informatisés et de conversion entre les différentes représentations.

### **Objets intelligents pour la perception non-spatiale**

Une possibilité intéressante que nous offrent les interactions avec les objets intelligents est celle de soumettre des requêtes de haut niveau, similaires à celles que permettent les assistants virtuels comme Siri d'Apple, Cortana de Microsoft ou Alexa d'Amazon. La personne qui utilise le système pose une question en langage naturel, et celui-ci y répond. Ainsi, par exemple, l'utilisateur peut ressentir du froid et poser la question « Pourquoi fait-il froid ? » au système. Celui-ci va alors vérifier la liste des objets intelligents pertinents au regard de cette question, c'est-à-dire

---

alternativement se baser sur des critères *a priori* valables pour l'ensemble des thermostats : ils sont de couleur unie, avec un écran, de moins de vingt centimètres de diamètre et fixés sur un mur ; ce type de recherche est moins fiable que ce qui est proposé.

les objets en lien avec la température ambiante : les thermomètres connectés qui mesurent la température peuvent donner une mesure objective ou confirmer une tendance, les systèmes de chauffage ou de climatisation peuvent donner leur état (allumé ou éteint) ou prendre une nouvelle consigne (maintien de la température ambiante à 20 °C par exemple). On peut aussi imaginer des portes intelligentes et des fenêtres intelligentes conscientes de leur état (ouverture ou fermeture). Le système 2SEES, au travers de sa fonctionnalité SO2SEES, peut alors compiler ces informations et les transmettre à la personne : il prend de cette manière le rôle d'interface entre les objets intelligents et la personne qui les utilise. En cela, il y a une grande similitude de principe entre la fonctionnalité SO2SEES et les assistants personnels précités, qui, au départ, interfaces entre Internet et les personnes qui les utilisent, cherchent à devenir le nœud central des réseaux d'objets intelligents domestiques (cf. le micro-ondes Alexa). La différence essentielle tient à la prise en compte de problématiques spécifiques aux personnes déficientes visuelles par le système 2SEES : la localisation des objets intelligents n'est par exemple pas un sujet d'intérêt pour les assistants personnels. On peut tirer parti du développement de ces assistants personnels, et s'en servir comme d'une infrastructure simplifiant l'interfaçage avec les objets intelligents, puisque le déploiement d'un assistant fixe implique l'appairage avec les équipements connectés locaux ; ainsi résolue, la problématique n'est plus à la charge du système 2SEES. Il peut également être très profitable d'exploiter les capacités d'analyse des requêtes en langage naturel que possèdent les assistants. En effet, ceux-ci, développés par les grandes entreprises du numérique que sont Google, Apple, Microsoft et Amazon, sont une application évoluée de l'intelligence artificielle. Leurs réseaux sont capables de « comprendre » des requêtes complexes, d'en extraire la sémantique et d'y répondre ; ce qui représente un problème difficile à résoudre. Plutôt que de répliquer cette fonctionnalité, on peut donc envisager de l'intégrer dans la fonctionnalité SO2SEES. Cela pose bien entendu la question de la confidentialité et de la diffusion de données privées ; plutôt que contenues dans le système 2SEES, les requêtes seraient directement traitées par les plates-formes citées.

Un système d'assistance peut faire autre chose (« faire plus » est probablement optimiste) que remplacer les capacités absentes ou perdues à cause du handicap : il peut offrir de nouvelles possibilités d'interactions avec l'environnement. C'est particulièrement le cas de l'environnement des objets intelligents, qui constitue un nouveau domaine à explorer : les modes d'interactions n'ont pas tous été envisagés. Traditionnellement, les interactions sont principalement basées sur la vision ; c'est également le cas dans l'informatique, où les interfaces sont avant tout visuelles, hors de certains cas spécifiques (dont justement les interfaces adaptées). Les objets intelligents représentent une opportunité pour le développement de nouveaux modes d'interactions, en particulier dans le cadre du handicap visuel, pour au moins deux raisons :

- Ils ont une présence duale : à la fois dans le monde physique et dans le monde virtuel, ils ont la capacité de servir de passerelle – dans une certaine mesure – pour les personnes ayant des difficultés d'interactions avec le monde physique telles que les personnes déficientes visuelles<sup>11</sup> ;

---

11. Cela peut également être valable pour celles qui sont atteintes par d'autres handicaps.

- Les informations qu'ils manipulent ne sont pas limitées à des quantités physiques. Tout l'intérêt de l'*intelligence* de ces objets est en effet de réaliser des jugements sur ces niveaux de quantités physiques pour en extraire des connaissances d'ordre symbolique, de haut niveau, concentrant plus d'information que les grandeurs scalaires et donc plus à même d'être transmises par des canaux non-visuels, dont la bande passante est limitée (cf. section 1.3).

Plus qu'une simple interface entre les objets intelligents et la personne qui l'utilise – c'est-à-dire un moyen d'accès aux informations relatives aux objets intelligents et générées par eux – le 2SEES peut combiner ces informations de manière holistique afin de fournir des connaissances de plus haut niveau à la personne qui l'utilise. Pour ce faire, il est nécessaire que les requêtes possibles ne soient pas limitées à la simple énumération d'informations cohérentes des objets pertinents, mais également de nature à agréger ces informations.

Génériquement, le cycle de vie d'une requête est le suivant (figure 4.7) : la requête est d'abord dictée en langage naturel, puis analysée par le système qui la transcrit dans un langage formel de haut niveau, qui est elle-même ensuite découpée en sous-requêtes plus simples destinées aux différents acteurs à interroger, c'est-à-dire les plates-formes *back-end* des gestionnaires d'objets intelligents. Ces sous-requêtes sont exécutées, puis leurs résultats sont agrégés pour répondre à la requête formelle de haut niveau. Développons le cas d'une requête précédemment prise comme exemple : « Pourquoi fait-il froid ? ». Pour y répondre, le système 2SEES va devoir :

- Identifier les objets intelligents qui possèdent un lien logique avec la température locale : thermomètres, radiateurs, fenêtres. D'autres sources d'informations peuvent aussi être considérées, comme les services de données météorologiques ;
- Le lien logique en question qui relie chacun des objets avec le concept au coeur de la requête, ici la température ambiante dans la pièce, doit être découvert. Le thermomètre « mesure » la température, le radiateur cherche à *fixer* la température selon une valeur de consigne (s'il est équipé d'un thermostat), et la fenêtre représente un facteur de *circulation* d'énergie entre l'intérieur et l'extérieur (différente selon qu'elle est fermée ou ouverte). Idéalement, tout ceci devrait être intégré dans un modèle décrivant les interactions des éléments et phénomènes auxquels on s'intéresse ; ici, ce serait le modèle physique décrivant la température de la pièce avec les objets cités parmi les paramètres ;
- Les objets et sources d'informations d'intérêt doivent ensuite être interrogés afin d'obtenir les informations qui intéressent le système. La température actuelle, voire un historique de la température, doit être obtenu du thermomètre, de même que les états respectifs de consigne et d'ouverture doivent être obtenus des radiateurs et des fenêtres ;
- Le système doit ensuite intégrer les données acquises dans le modèle utilisé à la seconde étape, puis utiliser ce dernier pour en déduire les informations qui serviront à répondre à la requête. Dans notre exemple, s'il fait froid dehors, la cause de la basse température de la pièce peut être une fenêtre ouverte

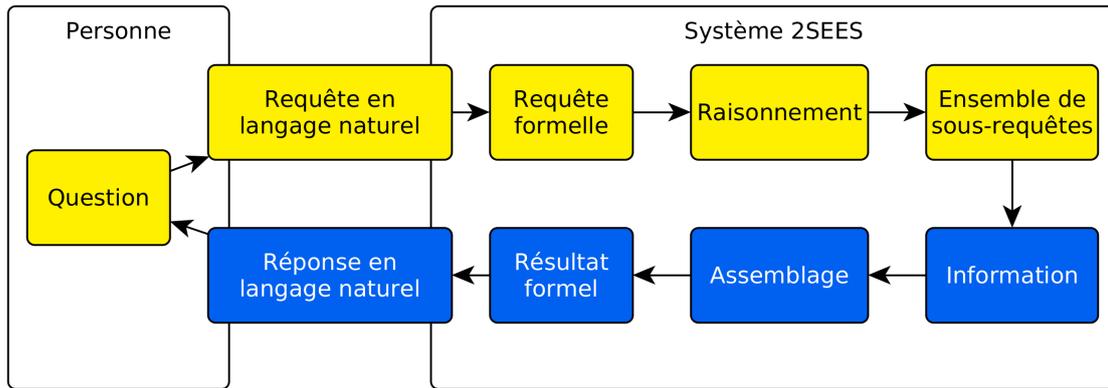


FIGURE 4.7 – Cheminement simplifié d’une requête et de la réponse délivrée.

ou bien le chauffage éteint.

Avec le rôle de passerelle entre la personne qui l’utilise et les objets intelligents environnants que remplit le système 2SEES, il peut être tentant de l’utiliser pour transmettre une grande quantité d’informations des seconds vers la première. C’est justement ce qui doit être évité car dommageable et contre-productif. La plupart des dispositifs d’assistances qui sont parvenus à être adoptés par une large population sont justement ceux qui présentent une quantité d’information limitée à la personne (cf. chapitre 2). La bande passante spatiale des sens autres que la vue est modeste, ce qui limite automatiquement le débit d’informations brutes pouvant être transmises (cf. chapitre 1.3). Plus le débit d’information est élevé, plus la charge cognitive augmente. Dans le cas du 2SEES les transmissions d’informations se font par le biais d’un encodage en langage naturel : au delà d’un certain débit de parole, il est clair que l’appareil sera inutilisable car trop fatiguant.

Pour parvenir à limiter la quantité de données transmises, il est nécessaire de filtrer les informations les plus porteuses de sens, intelligibles et concises. Cela souligne l’importance d’avoir accès à des requêtes de haut niveau telles que décrites précédemment plutôt que des requêtes uniquement bas niveau. Ici, la distinction se joue dans le traitement qu’implique chaque type de requête : celles que nous qualifions de « bas niveau » représentent des actions de collecte de données, alors que celles de « haut niveau » effectuent un traitement sur les données, les retravaillent afin d’en raffiner encore les informations. Au départ, des données brutes sont collectées par les capteurs des objets ; ces données sont traitées et un premier niveau de raisonnement leur est appliqué par les objets *intelligents* et leurs plates-formes de *back-end*. Le système 2SEES vient se greffer au-dessus de ce traitement en combinant ensemble plusieurs sources d’informations et en ajoutant des modèles externes pour en déduire des connaissances utiles. Sous l’hypothèse de la validité des modèles et des méthodes d’agrégation employés par le 2SEES, le système est, du point de vue de ses utilisateurs, plus efficace.

Un outil très important pour le filtrage de l’information par la pertinence est le contexte. Celui que nous examinons ici ne correspond pas à ce qui a été décrit dans la section 4.2.2 : il ne comprend pas les objets intelligents environnants, qui sont maintenant les sujets centraux dont les informations doivent être à traiter. Les ha-

bitudes générales de la personne, ses préférences, et les tenants et aboutissants de sa situation actuelle font partie du contexte. On leur ajoute les règles linguistiques, explicites comme implicites, qui vont structurer les requêtes. Enfin, on y inclut un élément qui est peut-être le plus important, en tout cas le plus simple en principe à déterminer : la position actuelle de la personne. La localisation a été très souvent utilisée comme information contextuelle, voire comme seule composante du contexte, au point que les deux ont fini par devenir presque synonymes dans une partie de la littérature [Schmidt et al., 1999]. Plus qu'en termes de coordonnées dans un repère, on s'intéresse ici à la fonction sociale du lieu : quelle type de pièce est-ce ? Se trouve-t-on sur une voie publique ou dans un bâtiment privé ? La caractérisation du lieu peut être très riche. La localisation est étroitement liée aux objets intelligents, qui sont en quelque sorte des matérialisations du monde virtuel dans des lieux physiques : mesures, activations d'actionneurs, interfaces avec les personnes.

La localisation n'est évidemment pas le seul élément pertinent du contexte : dans un magasin d'électroménager, la requête « Où trouver un four ? » aura des significations différentes selon que la personne qui la prononce est un employé de ce magasin qui cherche de quoi réchauffer son repas pendant sa pause de midi ou bien un client potentiel qui cherche à acheter un four. Un avantage de l'utilisation de la localisation comme information contextuelle est qu'elle est facile à utiliser. D'un côté, c'est une fonctionnalité basique du système, qui a besoin d'être calculée avec ou sans fonctionnalité SO2SEES de communication avec les objets intelligents. La fonctionnalité de navigation dépend de la fonctionnalité de localisation. De l'autre côté, elle est facile à utiliser : le filtrage spatial est facile à implémenter et permet de s'affranchir d'une très grande quantité d'informations qui ne sont pas pertinentes. La connectivité ou la découverte des objets intelligents *locaux* est en réalité une forme de filtrage spatial : on ne considère que les éléments proches du système.

Les sujets à propos desquels les utilisateurs vont faire des requêtes sont très probablement influencés par une partie ou une autre du contexte : par exemple, s'ils sont à un arrêt de tramway, ils pourraient se demander quand le prochain tramway va arriver à l'arrêt particulier où ils se trouvent. Un modèle du contexte à l'intérieur du module d'évaluation des requêtes est donc nécessaire pour comprendre cette requête, savoir de quel arrêt de tramway il faut vérifier les horaires. La figure 4.8 représente le cheminement ou l'évolution générique d'une requête tenant compte du contexte.

L'évaluation des requêtes dépend également de la traduction correcte de ce qu'exprime la personne qui soumet la requête du langage naturel vers le langage formel impliquant des symboles compréhensibles par le système. Cette étape de traduction repose sur les éléments de contexte que sont les règles du langage. Par exemple, la requête « Pourquoi fait-il froid ? » implique au moins deux règles, l'une implicite, l'autre explicite. Dans cette requête, le fait que le mot « il » n'exprime pas un sujet ou un objet particulier mais soit attaché au mode de description du temps qu'il fait constitue une première règle, explicite car on peut la trouver formulée dans un livre de grammaire. Le fait que cette requête interroge sur la température à l'endroit où se trouve la personne et pas à un quelconque autre endroit, c'est-à-dire qu'elle soit

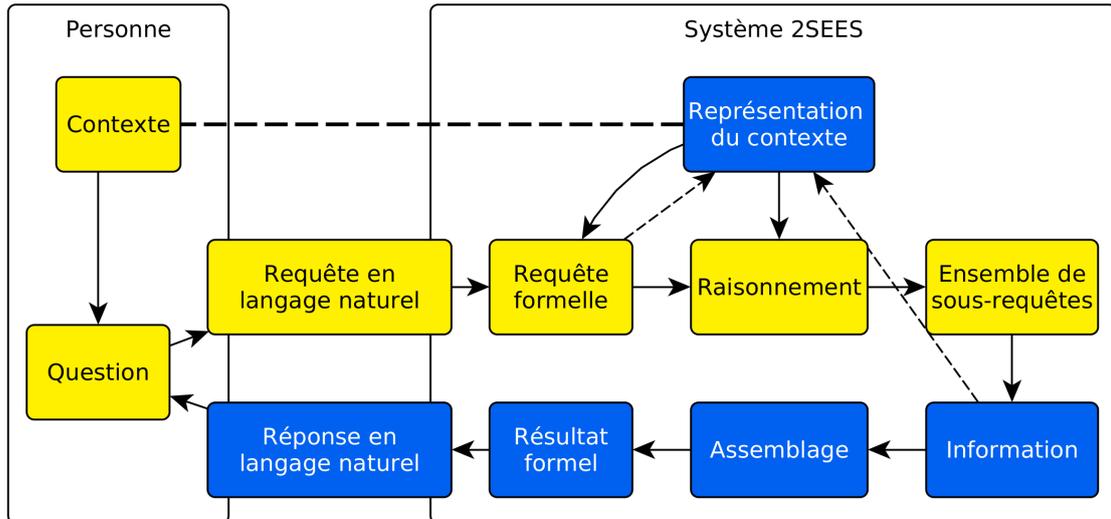


FIGURE 4.8 – Cheminement d’une requête prenant en compte le contexte.

équivalente à « Pourquoi fait-il froid *ici*? » constitue une autre règle de langage, implicite cette fois. Pour traduire correctement la requête, il est nécessaire que le module d’évaluation des requêtes prenne en compte ces deux règles.

Pour implémenter ce module d’évaluation des requêtes, nous avons le choix entre deux grandes approches : une intelligence artificielle entraînée, qui découvre les règles de contexte elle-même, ou bien une sorte de système expert utilisant des règles pré-définies. Nous avons choisi la seconde approche pour sa plus grande simplicité et accessibilité, et parce qu’il existe un modèle de règles compatible avec les standards du web sémantique : le *Semantic Web Rule Language* ou SWRL du W3C [Horrocks et al., 2004]. Le langage SWRL étend la logique de l’OWL avec des règles : chaque règle est composée d’un *corps* (l’antécédent) et d’une *tête* (la conséquence), chacune de ces deux parties étant elle-même composée d’une conjonction d’*atomes*. Ces atomes sont les faits vérifiables élémentaires sur des variables, tels que « cette variable appartient à cette classe » ou encore « ces deux variables sont liées par telle propriété ». Une règle est évaluée de la manière simple suivante : si l’antécédent est vérifié, alors la conséquence est vérifiée. Un exemple canonique de règle de cette forme est donné dans l’équation 4.1 (où  $?x$ ,  $?y$ , et  $?z$  sont des variables représentant des personnes) :

$$aPourParent(?x, ?y) \wedge aPourFrère(?y, ?z) \rightarrow aPourOncle(?x, ?z) \quad (4.1)$$

On peut formuler cette règle en langage naturel par : « si  $x$  a pour parent  $y$  et que  $y$  a pour frère  $z$ , alors  $x$  a pour oncle  $z$  ». Dans le système SO2SEES, de telles règles font partie d’une base de connaissances qui comprend également les connaissances descriptives encodées en OWL et RDFs. En fait, les règles forment une large partie de la quantité de connaissances requise pour répondre aux requêtes. Le système a besoin d’une certaine quantité de connaissances sémantiques et logiques pour répondre aux requêtes – des connaissances qui modélisent le monde. Cette logique

peut se trouver soit du côté du module d'évaluation des requêtes, soit du côté de la base de connaissances, mais elle est incompressible. L'usage de SWRL incite à placer la logique du côté de la base de connaissance. Cette alternative a l'avantage de permettre le partage de la logique des règles entre toutes les requêtes, ce qui facilite l'extensibilité et l'adaptation du système à des requêtes plus nombreuses et expressives. Un autre avantage est le déplacement du fardeau du lien logique entre les concepts depuis le domaine du SO2SEES vers le domaine des fabricants et gestionnaires d'objets intelligents, ainsi que vers les développeurs d'ontologies spécialisées ou généralistes. De plus, dans le cas où une partie de la logique manque pour implémenter des requêtes spécifiques, des règles peuvent être ajoutées par les développeurs et gestionnaires du système SO2SEES : la base de connaissances sur laquelle les raisonnements induits par les requêtes sont effectués se trouve sur le *cloud* du système 2SEES.

Dans un système qui vient d'être mis en service, le problème de la réponse aux requêtes est dynamique : à chaque nouvelle requête, le système doit comprendre la logique sous-jacente et s'en servir pour interroger les bases de connaissances – sa propre base et celles des autres acteurs. La situation d'un système déjà installé est par contre différente. Avec suffisamment de requêtes dans sa base de connaissances, le problème de la réponse aux requêtes devient un problème statique : le système doit trouver la correspondance entre la requête et la ou les règles qui y correspondent. En quelque sorte, toutes ou presque toutes les requêtes envisageables ont déjà été soumises au système, et l'opération de traduction des requêtes est désormais remplacée par une opération de reconnaissance. La première opération est beaucoup plus difficile à résoudre que la seconde : c'est un problème de correspondance entre un langage naturel et des concepts symboliques, qui a certes bien progressé avec les techniques modernes d'intelligence artificielle (apprentissage profond ou *deep learning*), mais reste hasardeux, surtout avec l'importance du contexte que revêt notre cas d'usage.

Notre choix de faire fonctionner le système SO2SEES avec des requêtes pré-définies plutôt qu'un véritable module d'évaluation des requêtes représente, en quelque sorte, un repositionnement sur cet avantage. Mécaniquement, on arrive dans la seconde situation : la personne qui utilise le système sélectionne quelle requête elle souhaite soumettre plutôt que d'en créer une nouvelle. Ainsi, tous les aspects difficiles de la chaîne d'interprétation, c'est-à-dire les aspects où la logique des requêtes est extraite, se trouve déplacée en phase de conception des requêtes par les gestionnaires du système. Cela a l'avantage auxiliaire de rendre l'exécution très rapide, aucun traitement logique n'étant à effectuer en temps réel.

Le fait d'encoder la logique des règles sous forme de règles formelles dans la base de connaissances est aussi intéressant pour le cas où les règles sont générées par une intelligence artificielle plutôt que par des experts. On peut envisager un tel fonctionnement avec, par exemple, du minage de conversations pour extraire les relations logiques dans un ensemble de concepts. La représentation des connaissances générées par l'intelligence artificielle par des règles formelles a l'avantage d'offrir aux humains une représentation accessible de la logique qui les gouverne. Cette caractéristique est importante : nous travaillons en effet sur une application

qui doit être la plus sûre possible, et cela passe par des analyses et une compréhension du fonctionnement du système. L'utilisation « naïve » des réseaux de neurones profonds est peu adaptée à ce type de contraintes, car le processus de décision y est en pratique une boîte noire, difficile à vérifier<sup>12</sup>. Des travaux de recherche essayent d'améliorer les caractéristiques d'interprétabilité et d'explicabilité des réseaux de neurones profonds [Amarasinghe et al., 2018, Choo and Liu, 2018].

Une implémentation partielle de la fonctionnalité SO2SEES a été réalisée et un scénario de test développé. Ils sont décrits dans la section suivante.

### 4.3 Scénarios de test

Dans cette section, nous décrivons un scénario dont l'objectif est de valider le concept de fonctionnalité d'interaction avec les objets intelligents décrit à la section précédente. Pour atteindre cet objectif, il est nécessaire de démontrer à la fois la faisabilité d'une telle fonctionnalité et son intérêt pratique. Le second objectif implique partiellement le premier : un système ne peut pas être testé s'il n'est pas fonctionnel, au moins sur certains de ses aspects ; c'est encore plus marqué pour ce qui est de l'évaluation de l'intérêt. Conséquemment, les expérimentations sont focalisées sur la démonstration de l'utilité de la fonction de perception environnementale par l'intermédiaire des objets intelligents.

Pour établir la plus-value apportée en terme de *perception spatiale*, nous pouvons observer l'effet de l'utilisation du système dans les tâches reposant sur cette forme de perception. C'est le cas des tâches requérant l'orientation. Le premier critère du test est la performance objectivement mesurée : capacité et durée pour accomplir une tâche spécifique, nombre de collisions relevées avec les obstacles, etc. Il est également important de recueillir les avis subjectifs des personnes déficientes visuelles qui essayent la canne. La constitution d'une base commune pour l'expression des avis peut se faire avec l'aide d'un questionnaire. Il est à noter que les bénéfices qu'offre la perception des objets intelligents au système 2SEES pour la réalisation de ses deux autres grandes fonctions (détection d'obstacles et navigation) ne sont pas évalués par ce scénario. Pour cette raison, l'évaluation du scénario ne nécessite pas l'utilisation du LMSC. En effet, dans le cadre de la fonctionnalité SO2SEES, ce dernier n'occupe qu'un rôle d'interface réseau entre les objets intelligents et le smartphone (qui exécute l'application de transmission de requêtes et présente les réponses à la personne), et peut donc être remplacé par un appareil plus simple jouant ce rôle de passerelle réseau. Cette absence du LMSC offre l'avantage de découpler l'évaluation de la fonctionnalité SO2SEES de celle du LMSC. Cette dernière est susceptible de former une source importante de bruit informationnel<sup>13</sup> lors du test de la fonctionnalité SO2SEES, en raison de son format inhabituel et assez peu pratique. Tester la fonctionnalité SO2SEES en isolation signifie que le

---

12. De là viennent d'ailleurs de nombreux débats sur la sûreté de fonctionnement des véhicules autonomes ; il est difficile d'analyser *a priori* les situations qui seront problématiques pour ces systèmes.

13. Autrement dit, une augmentation du nombre de variables dans l'expérimentation, et donc une plus grande complexité.

sujet d'expérimentation est libre d'utiliser le dispositif d'assistance à la mobilité « classique » de son choix, en l'occurrence une canne blanche.

L'intérêt et l'utilité de l'aspect assistance à la perception *non-spatiale* est plus difficile à évaluer que l'aspect spatial, à cause du manque de mesures objectives de la performance du système. En effet, le principe est nouveau, comme les interactions qu'il permet. On peut envisager de le comparer à ce qui est le plus proche : d'une part, à l'utilisation d'un assistant personnel, et d'autre part, l'assistance d'une personne valide. Il est certain que pour cet aspect du système, le caractère subjectif revêtira une grande importance, et la rédaction d'un questionnaire d'évaluation pertinent sera essentielle. La fonctionnalité est destinée à donner accès à des informations qui sont normalement et généralement inaccessibles.

### 4.3.1 Matériel utilisé

Afin de tester et de valider le concept proposé, quatre objets intelligents ont été déployés dans un bureau du LIMOS. Concernant les objets, deux alternatives se présentaient à nous : l'utilisation d'objets du commerce, ou bien le développement d'objets spécifiques. La première possibilité présentait l'intérêt de rendre le scénario plus proche de la réalité, au prix d'une importante complexité dans l'implémentation du système. Il aurait été nécessaire de développer des mécanismes de communication réseau de bas niveau entre l'échelon local du 2SEES et chacun des objets intelligents, ce qui est loin d'être évident car les fabricants tendent à imposer leurs propres passerelles. Comparativement, l'utilisation de « maquettes » d'objets intelligents développées en interne était plus simple. Cette solution, qui a été retenue, est facilitée par l'utilisation de la carte uSu-Edu, qui intègre l'ensemble des composants nécessaires à l'implémentation d'un prototype d'objet intelligent : microcontrôleur basse consommation, émetteur-récepteur (IEEE 802.15.4) et capteurs (accéléromètres, gyromètres, et capteurs de température ambiante, humidité relative de l'air et luminosité ambiante). Le système d'exploitation (cf. section 3.3.4) était également intéressant, car il offre des primitives de communication sans fil faciles à utiliser.

Quatre maquettes ou « simulations » d'objets intelligents ont été implémentées : un thermomètre, un radiateur, une porte et une fenêtre instrumentées. Le thermomètre mesurait simplement la température ambiante et la mettait à disposition dans une base de données – il fonctionnait réellement, grâce au capteur de température de la carte uSu-Edu. Le radiateur simulait un système de chauffage équipé d'un thermostat, essayant de maintenir la température ambiante à une consigne donnée. La maquette de radiateur n'était en pratique qu'une simulation : elle n'était pas liée au système de chauffage réel de la pièce. La porte instrumentée (figure 4.9) détectait et rapportait son état (ouverte ou fermée). La détection était réalisée avec les données du gyromètre et un simple automate à nombre fini d'états. La fenêtre intelligente fonctionnait de la même manière, observant l'état d'ouverture du battant.

Les objets réalisés avec des cartes uSu-Edu étaient reliés à Internet par l'intermédiaire d'une passerelle que nous désignons sous le nom de « serveur local ». Ce



FIGURE 4.9 – Maquettes d’objets : à gauche, la fenêtre intelligente ; à droite, la porte intelligente.

serveur local était lui-même construit autour d’une carte uSu-Edu, connectée à un Raspberry Pi 3 (figure 4.10). Cela lui permettait d’offrir plusieurs interfaces de communication : IEEE 802.15.4 à travers la carte uSu-Edu, et WiFi, Bluetooth et Ethernet à travers le Raspberry Pi. Il est possible d’ajouter un module modem sur port USB pour bénéficier d’une connectivité 3G/4G, ce qui lui permettait de se connecter à Internet même en l’absence de réseau local. Jusqu’à vingt nœuds ou objets intelligents peuvent être connectés à un seul serveur local. Une portée de deux cent mètres en terrain libre est possible avec une antenne omnidirectionnelle ; avec une antenne directionnelle, de plus grandes portées sont envisageables. Dans notre scénario, les objets intelligents étaient déployés avec une topologie en étoile dont le serveur local était le centre. Le serveur local est connecté à un ordinateur de bureau, sur lequel fonctionnait le *back-end* des objets connectés.

Le système 2SEES était constitué par un ordinateur sous FreeBSD faisant fonctionner les logiciels du *back-end*, un smartphone avec l’application 2SEES ainsi que du LMSC ou bien d’un appareil constitué d’une carte uSu-Edu et d’un Raspberry Pi 3 pour le remplacer.

### 4.3.2 Implémentation logicielle

#### *Back-end* des objets intelligents

Le *back-end* des objets intelligents était principalement constitué par un *triplestore* stockant leurs données. Un *triplestore*<sup>14</sup> est une base de donnée dédiée au stockage

14. Le « triple » dans le nom vient du format des énoncés RDF : les triplets (*sujet, prédicat, objet*).



```

    ],
    [ td:mediaType "text/turtle" ;
      td:href "set_thermostat_setpoint"
    ] ;
  td:description "Thermostat setpoint of the radiator,\
    adjusted from software or hardware knob."@en ;
  td:description "Consigne du thermostat du radiateur,\
    choisie logiciellement ou par le bouton."@fr ;
  td:writable "true"^^xsd:boolean ;
  td:observable "false"^^xsd:boolean
] ;
td:description "Demo radiator"^^xsd:string .

```

### ***Back-end* du 2SEES**

Le *back-end* du 2SEES fonctionnait sur un autre ordinateur que celui des objets intelligents ; les deux faisaient partie du même réseau local. Trois applications fonctionnaient sur cet ordinateur : un *triplestore* et deux applications web. Le *triplestore* stockait les données RDF gérées par l'application : logique des requêtes et base de connaissances permettant d'y répondre. La première application web était chargée de répondre aux requêtes ; elle est écrite en Python avec le cadriciel (*framework* en anglais) Flask. La seconde application, qui est également écrite en Python/Flask, est une interface graphique pour permettre l'édition de requêtes (figure 4.17).

### **Application smartphone 2SEES**

L'application smartphone qui permet d'utiliser la fonctionnalité SO2SEES est la même que l'application de navigation ; une surcouche a été ajoutée pour les requêtes et les informations sur les objets intelligents (figure 4.11). Ce choix a été fait pour faciliter l'exploitation ultérieure de synergies entre les deux : la localisation et les cartes sont gérées par l'application de navigation et peuvent être très utiles pour la fonctionnalité SO2SEES ; inversement, les informations des objets intelligents peuvent servir à enrichir la fonctionnalité de navigation.

La surcouche est très simple. Elle consiste en un écran supplémentaire où, par défaut, un seul bouton est disponible, qui active le scan des objets connectés, c'est-à-dire, en pratique l'envoi des identifiants IRI des objets locaux au serveur afin que celui-ci propose les requêtes associées. Quand celles-ci sont proposées, elles s'affichent sous la forme de boutons supplémentaires. Les boutons sont relativement accessibles aux personnes déficientes grâce à la simplicité visuelle de l'interface (des boutons alignés verticalement) et à la fonctionnalité d'accessibilité d'Android, qui lit le contenu de l'écran. Quand une requête est envoyée, la réponse que fournit le serveur est lue par l'application.

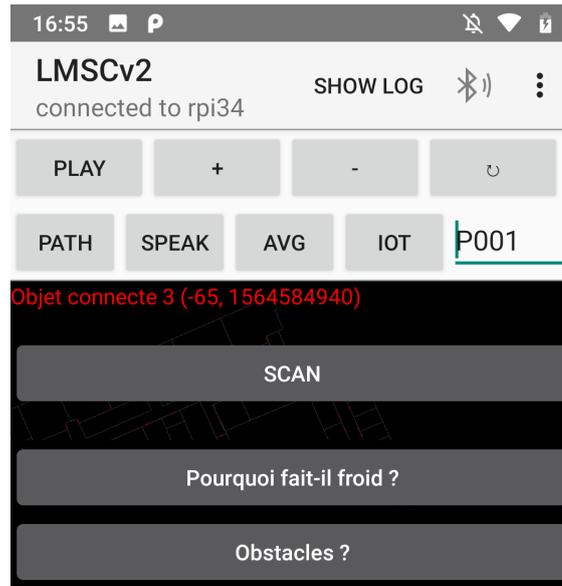


FIGURE 4.11 – Capture d’écran de l’application mobile 2SEES et l’interface SO2SEES proposant deux requêtes.

### 4.3.3 Scénarios de test

Deux scénarios ont été conçus pour permettre une évaluation de la fonctionnalité SO2SEES. Le premier se concentrait sur l’aspect non-spatiale de l’assistance à la perception offerte par la fonctionnalité SO2SEES, tandis que le second visait à explorer les apports possibles en termes de perception spatiale.

#### Premier scénario

Le premier scénario se déroule de la manière suivante : une personne équipée du système 2SEES entre dans la pièce, qui contient quatre objets intelligents : le thermomètre, le radiateur, la porte, et la fenêtre (figure 4.12). Le mécanisme de découverte des objets locaux entre alors en fonction. Une variation par rapport à la conceptualisation exposée plus haut de la fonctionnalité SO2SEES y a été introduite. Pour rappel, deux mécanismes potentiels étaient décrits :

1. Une découverte locale réelle, reposant sur des messages diffusés par les objets intelligents sur le médium de communication partagé entre eux et l’échelon local du 2SEES ;
2. Un registre ou annuaire associant à chaque position géographique les objets qui s’y trouvent. Ce registre serait public et alimenté par les gestionnaires d’objets connectés.

Pour l’implémentation concrète du scénario, c’est finalement une troisième voie qui a été adoptée. Conceptuellement, elle se trouve en quelque sorte entre les deux précitées. En effet, comme la première, elle repose sur des messages diffusés sur le médium local, mais, comme la seconde, elle n’impose pas de contrainte particulière

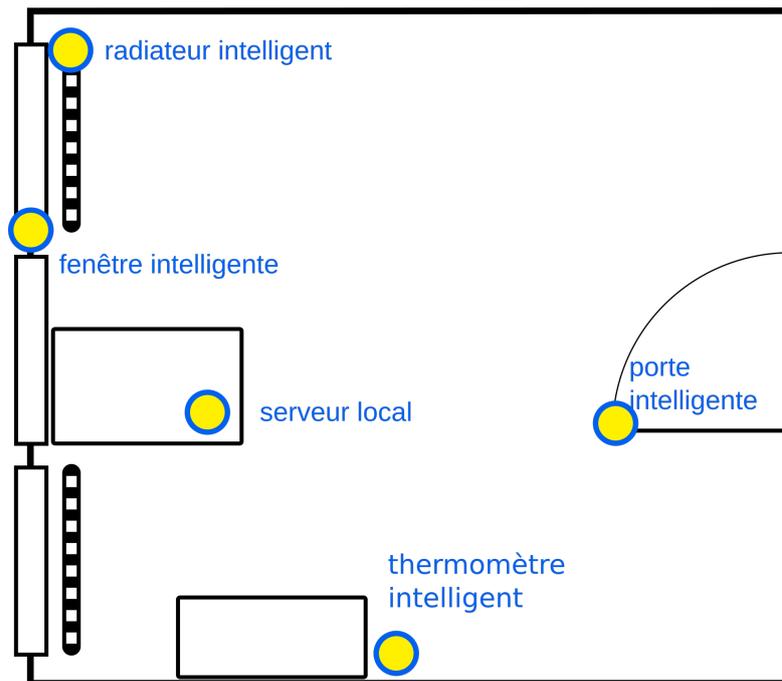


FIGURE 4.12 – Plan de la pièce dans le premier scénario d’expérimentation montrant les objets inclus.

aux objets intelligents eux-mêmes. Cette troisième voie est un changement de responsabilité de la diffusion des IRIs : celle-ci est transférée des objets au serveur local, ce qui apporte des avantages décrits en détails dans la section 4.4.5).

Une fois la personne entrée dans la pièce, les IRIs diffusées par le serveur local sont captées par le LMSC. Ce dernier les transmet à l’application smartphone, qui, à son tour, les envoie au *back-end*. Chaque IRI pointe vers un fichier accessible contenant la description formelle, au standard WoT-TD, d’un objet ; ces fichiers sont hébergés sur le *back-end* des objets intelligents. Celui du système 2SEES, quand il reçoit les IRIs envoyées par l’échelon local, télécharge tous ces fichiers et les analyse afin de déterminer quelles requêtes sont pertinentes dans le contexte que forment les objets intelligents environnants.

Une fois que les requêtes pertinentes ont été identifiées par le serveur, elles sont envoyées à l’application smartphone, qui les énonce à la personne utilisatrice. Deux requêtes sont disponibles dans ce scénario, en lien avec les objets inclus. La première est « Quelle est la température actuelle ? » ; la seconde « Pourquoi fait-il froid ? ». La personne peut sélectionner une requête et la soumettre au système. Cette opération provoque l’émission d’un message contenant l’identifiant de la requête au serveur de *back-end*. Le programme qui fonctionne sur ce serveur va alors chercher le code qui correspond à la requête et l’exécuter. Un programme d’exécution est défini pour chaque requête lors de sa création. Dans notre scénario, le programme de la première requête (« Quelle est la température actuelle ? ») est conceptuellement simple : il se contente d’interroger tous les objets possédant un moyen de mesure de la température. Dans le cadre du scénario, la requête interrogera le thermomètre intelligent et le radiateur, qui possèdent chacun un capteur de température. Les

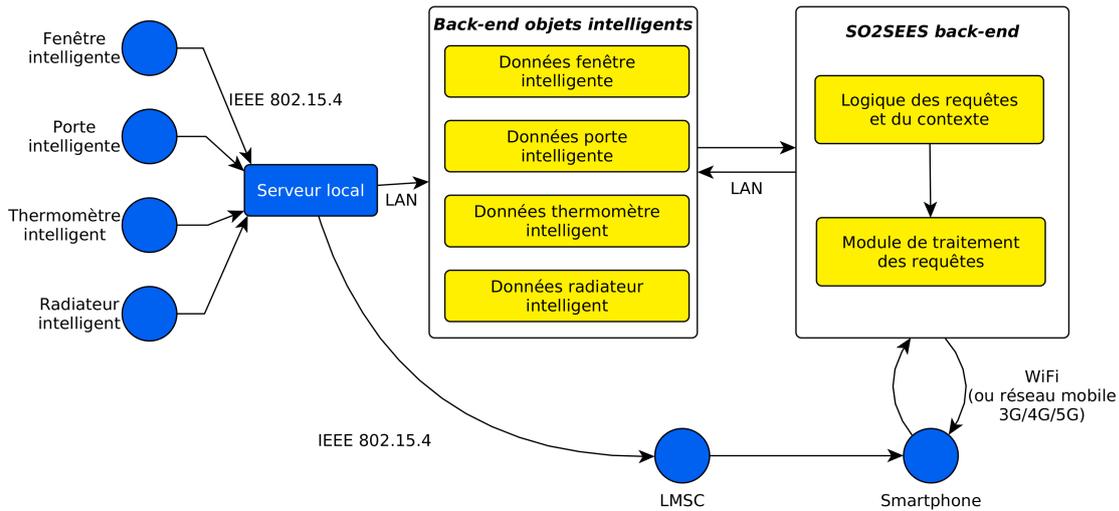


FIGURE 4.13 – Canaux de communication dans le premier scénario.

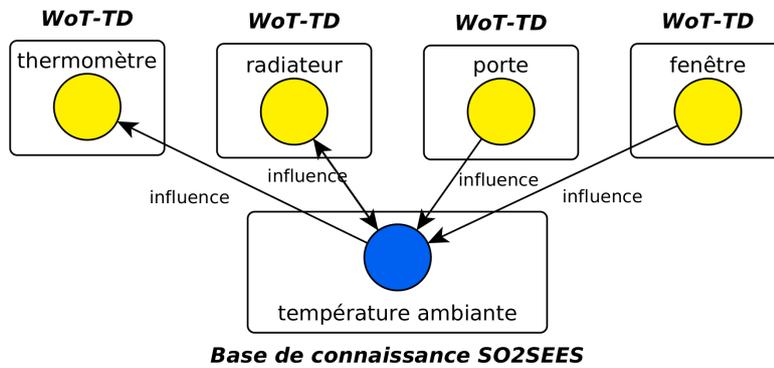


FIGURE 4.14 – Relations d'influence entre les objets et la température.

canaux de communication sont montrés par la figure 4.13.

Le programme de la seconde requête est un peu plus complexe. En effet, celle-ci doit chercher et lister les sources potentielles de froid. On négligera ici la température extérieure, mais en pratique il faudrait la connaître pour avoir un programme de requête réellement pertinent ; elle peut être obtenue grâce à un capteur placé à l'extérieur du bâtiment ou bien tout simplement par l'intermédiaire d'un service mettant à disposition des données météorologiques. En réalité, le modèle qui permet de réellement répondre à la requête est de nature physique, et difficile à représenter avec la logique du web sémantique. Dans un premier temps, la requête peut se contenter de lister à la personne l'état de tous les objets ayant une influence sur la température, c'est-à-dire, dans le scénario, la porte, la fenêtre et le radiateur (figure 4.14). On peut envisager une requête plus évoluée ne donnant l'état de chaque objet que si celui-ci a une influence positive sur le froid (ne citant par exemple le radiateur que si celui-ci est éteint), mais cela n'est pas nécessaire pour que la requête soit compréhensible et efficace.

Dans la première requête, la condition d'éligibilité des objets est facile à implémenter : il suffit de regarder le graphe biparti. Dans le cas de la seconde requête,

c'est plus problématique : où trouver, en effet, l'information concernant *l'influence des objets sur la température*? En principe, elle n'existe pas dans la description formelle de l'objet ; il faut donc que nous l'ajoutions. Cela peut être fait dans un « contexte » ou « corpus de connaissances » attaché à la requête – les liens causaux entre objets et température seraient alors manuellement définis lors de la création de la requête, et attachés à celle-ci. Mais il paraît bien plus élégant et efficient d'inscrire cette information dans une base de connaissances plus large, commune à l'ensemble des requêtes. Encore une fois, le partage permet la réutilisation de la logique descriptive, le développement plus aisé de nouvelles requêtes, et une expressivité accrue de celles-ci grâce à l'accès à un plus grand corpus de description de l'environnement. Cette base de connaissances n'est pas autre chose qu'une ontologie : la définition formelle de termes et concepts et de leurs interactions. Comme les requêtes, cette ontologie pourra être construite progressivement.

Dans l'exemple du scénario, il faut que la base de connaissances contienne les relations entre les objets intelligents et la température (figure 4.14). Afin que le vocabulaire de l'ontologie créée pour certaines requêtes soit le plus apte à être réutilisé par d'autres, il faut essayer de définir les relations aussi génériquement que possible. Par exemple, il vaut mieux lier le concept de *température* au concept *capteur de température* (par la relation *mesure*) qu'au concept de *LM75B*, qui est un capteur de température particulier. Le premier choix étendra la compatibilité avec tous les capteurs de température, tandis que le second limitera l'effet aux objets intelligents utilisant un type de capteur en particulier. Dans le deuxième cas, il faudra créer de nombreuses relations liant chaque modèle de capteur à la température, ce qui est inefficent et redondant en termes informationnels.

Il est ensuite nécessaire de définir les programmes régissant l'exécution des requêtes. On peut simplement utiliser un langage impératif classique pour faciliter l'intégration avec le *back-end* : dans notre implémentation, celui-ci est écrit en Python ; il serait facile de définir chaque requête comme une fonction ou une classe Python chargée d'interroger les acteurs externes et la base de connaissances, avant de calculer un résultat.

On peut également intégrer la plus grande partie possible de la logique des requêtes à l'intérieur même de la base de connaissances. Le langage *Semantic Web Rule Markup Language* ou SWRL (décrit dans la section 4.2.3) est un outil très pratique pour réaliser cet objectif. La logique de la première requête est triviale. Pour la seconde, on peut améliorer la requête précédente en ne prenant cette fois en compte *que* les objets dont l'état particulier est susceptible d'entraîner la sensation de froid. Pour cela, nous devons ajouter le concept de *cause possible* à la base de connaissances :

- Le radiateur est une cause possible de froid uniquement s'il est éteint ou bien réglé avec un point de fonctionnement bas, 18 °C par exemple ;
- La fenêtre est une cause possible de froid seulement si elle est ouverte ;
- Même chose pour la porte ;
- Le thermomètre n'est pas une cause possible de froid.

La logique de la requête pour le radiateur peut être encodée par deux règles : une pour vérifier si le radiateur est éteint, et une pour vérifier si son point de fonction-

nement est inférieur à une température donnée, synonyme de froid (équations 4.2 et 4.3). Celle de la fenêtre peut être encodée dans une seule (équation 4.4). Dans ces règles, les concepts sont séparés entre l'espace du nommage du gestionnaire de l'objet, que nous nommons **externe:**, et l'espace de nommage du système 2SEES, nommé **2sees:**.

$$\begin{aligned}
 & \textit{externe} : \textit{radiateur} (?x) \\
 & \wedge \textit{externe} : \textit{etatRadiateur} (?x, ?etat) \\
 & \wedge \textit{swrlb} : \textit{stringEqualIgnoreCase} (?etat, "éteint") \\
 & \rightarrow \textit{2sees} : \textit{causePossible} (?x)
 \end{aligned} \tag{4.2}$$

$$\begin{aligned}
 & \textit{externe} : \textit{radiateur} (?x) \\
 & \wedge \textit{externe} : \textit{parametrageRadiateur} (?x, ?temperature) \\
 & \wedge \textit{swrlb} : \textit{lessThanOrEqual} (?temperature, 18) \\
 & \rightarrow \textit{2sees} : \textit{causePossible} (?x)
 \end{aligned} \tag{4.3}$$

$$\begin{aligned}
 & \textit{externe} : \textit{fenetre} (?x) \\
 & \wedge \textit{externe} : \textit{estDansEtat} (?x, ?etat) \\
 & \wedge \textit{swrlb} : \textit{stringEqualIgnoreCase} (?etat, "ouverte") \\
 & \rightarrow \textit{2sees} : \textit{causePossible} (?x)
 \end{aligned} \tag{4.4}$$

Ces règles sont en pratique implémentées de manière un peu différente pour correspondre avec l'utilisation de WoT-TD. Nous avons utilisé dans les règles 4.2, 4.3, et 4.4 des concepts correspondants aux objets : **externe:radiateur**, **externe:fenêtre**, **externe:estDansÉtat**, etc. Les objets sont ici définis du côté des acteurs externes, comme l'indique le préfixe **externe:**. Ceci est un choix que nous avons fait ; WoT-TD n'impose pas la création de concepts correspondant à chaque type d'objet. De tels concepts sont très utiles pour la clarté, et nous aurions pû pour cette raison les définir dans la base de connaissances du 2SEES, mais dans l'absolu, le système peut fonctionner sans qu'ils n'existent. WoT-TD est avant tout conçu pour une syntaxe JSON relativement éloignée des considérations du web sémantique et très souple. Ceci vise à favoriser son adoption par les fabricants d'objets, en abaissant la barrière qu'est la complexité d'implémentation et le travail de mise en relation avec les ontologies existantes. Si on se passe de concepts spécifique aux objets en se concentrant sur du WoT-TD « pur », la requête 4.3 deviendra par exemple la requête 4.5 (dans la quelle nous utilisons le préfixe **td:** pour signifier <https://www.w3.org/2019/wot/td>) : on a remplacé **externe:fenetre(?x)**

et `externe:estDansEtat(?x,?etat)` par leurs équivalents en WoT-TD.

$$\begin{aligned}
& td : Thing(?objet) \\
& \wedge td : title(?objet, ?titre) \\
& \wedge swrlb : stringEqualIgnoreCase(?titre, "radiateur") \\
& \wedge td : hasPropertyAffordance(?objet, ?p\_parametre) \\
& \quad \wedge td : type(?p\_parametre, ?valeur) \\
& \quad \wedge swrlb : lessThanOrEqual(?valeur, 18) \\
& \quad \rightarrow 2sees : causePossible(?objet)
\end{aligned} \tag{4.5}$$

Si on souhaite intégrer les concepts correspondant aux objets dans la base de connaissances du système 2SEES – la troisième approche –, il est possible de définir une règle permettant de déduire l'appartenance à la classe radiateur à partir des informations fournies par WoT-TD (équation 4.6). En procédant ainsi pour les autres objets, cela enrichira la base de connaissances du 2SEES avec des concepts d'objets faciles à manipuler.

$$\begin{aligned}
& td : Thing(?objet) \\
& \wedge td : hasPropertyAffordance(?objet, ?p\_type) \\
& \quad \wedge td : type(?p\_type, ?valeur) \\
& \wedge swrlb : stringEqualIgnoreCase(?valeur, "radiateur") \\
& \quad \rightarrow 2sees :: radiateur(?objet)
\end{aligned} \tag{4.6}$$

## Deuxième scénario

Le deuxième scénario est orienté vers le second aspect de la fonctionnalité, la perception spatiale. Il se déroule de la manière suivante : une personne équipée du système 2SEES entre dans la pièce, qui contient quatre objets : la porte intelligente et trois bureaux (meubles) intelligents (figure 4.15). Comme dans le précédent scénario, leurs IRI sont diffusées par le serveur local quand celui-ci détecte une utilisation de la porte. Elles sont ensuite captées par la carte uSu-Edu qui équipe la personne, transmises au smartphone puis au *back-end*, qui détecte les bureaux et propose une requête associée « Quels sont certains des obstacles dans la pièce ? ».

Le *back-end* utilise une règle très simple (équation 4.7) afin de donner les obstacles qui se trouvent dans la pièce. Les « obstacles » utilisés dans la règle ne sont pas définis dans description formelle des objets qui sont des obstacles, ici, des bureaux. Ce serait une lourde contrainte : chaque fabricant ou gestionnaire devrait définir quel objet constitue un obstacle pour les personnes déficientes visuelles – ce point menant naturellement à bien d'autres. Il est plus efficient d'encoder le lien entre les obstacles et chaque type d'objet en une fois, via un ensemble de règles ou de triplets RDFs. Elles peuvent faire partie de la base de connaissances du vendeur ou bien de celle du 2SEES ; nous avons choisi la seconde solution. Dans notre exemple,

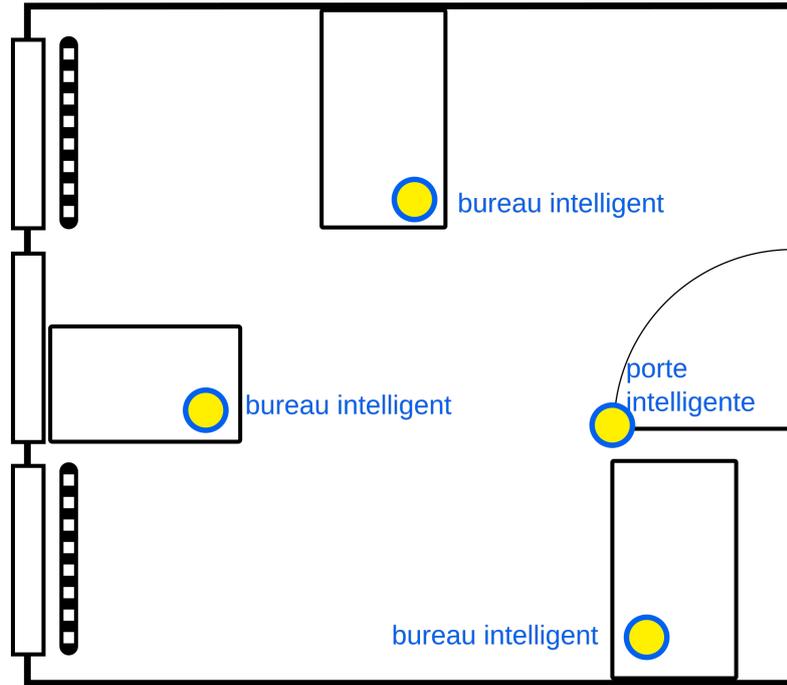


FIGURE 4.15 – Plan de la pièce dans le second scénario d’expérimentation montrant les objets inclus.

cela correspond à une règle triviale définissant le bureau comme obstacle (équation 4.8) ou à un triplet RDFs correspondant (équation 4.9), tous deux très simples.

$$\begin{aligned}
 & 2sees : obstacle(?x) \\
 & \wedge externe : pièce(?y) \\
 \wedge externe : objetDansPiece(?x, ?y) & \hspace{10em} (4.7) \\
 \rightarrow 2sees : obstacleASignaler(?x) &
 \end{aligned}$$

$$\begin{aligned}
 & externe : bureau(?x) \\
 \rightarrow 2sees : obstacle(?x) & \hspace{10em} (4.8)
 \end{aligned}$$

$$externe : bureau \text{ rdfs : subclassOf } 2sees : obstacle \hspace{10em} (4.9)$$

Comme les précédentes, les règles 4.7 et 4.8 sont « simplifiées » par rapport à une utilisation de WoT-TD en isolation, et supposent une ontologie externe. La règle 4.8 devient par exemple la règle 4.10. Il est également nécessaire que la base de connaissances prenne en compte les pièces et la présence d’objets à l’intérieur pour que les

atomes *externe* : *pièce(?y)* et *externe* : *objetDansPiece(?x, ?y)* fonctionnent.

$$\begin{aligned}
 & td : Thing(?objet) \\
 & \wedge td : title(?objet, ?titre) \\
 & \wedge swrlb : stringEqualIgnoreCase(?titre, "bureau") \\
 & \rightarrow 2sees : obstacle(?objet)
 \end{aligned} \tag{4.10}$$

## 4.4 Discussion

### 4.4.1 Analyse du contexte

Dans le premier scénario, deux aspects du contexte sont utilisés pour le filtrage des requêtes :

- Le lieu géographique, grâce au filtrage spatial implicite induit par la diffusion locale des IRIs ;
- L'analyse des descriptions d'objets pour déterminer leur usage.

Le mécanisme implémenté pour l'analyse des descriptions d'objets est très simple : elle consiste à extraire chaque nœud RDF de la description d'un objet. Chacun de ces nœuds représente ce que nous nommons un « indice de contexte » ou « indice contextuel », c'est-à-dire un signe porteur d'information sur le sens de l'objet. Par exemple, un capteur mentionné dans la description de l'objet constituera un tel indice. Nous lions donc les objets à des indices contextuels ; symétriquement, les requêtes sont elles aussi liées à des indices contextuels. Les requêtes qui sont proposées à la personne sont alors celles pour lesquelles les indices de contexte liés correspondent à ceux extraits des descriptions formelles des objets.

Avec ce type de relation entre objets et requêtes, la problématique est de construire les liens entre les requêtes et les indices de contexte, qui forment un graphe biparti (et, avec les objets, un graphe triparti – figure 4.16). L'idée qui nous apparaît comme étant la plus pertinente est celle d'une construction progressive du graphe, au fur et à mesure de l'utilisation du système 2SEES *dans son ensemble* – par tous ceux qui se servent de chaque canne intelligente 2SEES. L'alimentation en données concernant les objets serait alors faite de manière participative : quand un nouvel objet est rencontré par un système 2SEES, les indices de contexte sont extraits de sa description et manuellement liés aux requêtes déjà existantes. Cela nous pousse à envisager le cycle de vie des requêtes sous un angle similaire : elles peuvent être ajoutées lorsque le besoin s'en fait ressentir. Or, il est probable que le moment où de nouvelles requêtes deviennent nécessaires ou souhaitables coïncide avec la « découverte » de nouveaux objets, puisque ce sont eux qui apportent les possibilités d'interactions. Ainsi, le processus d'ajout de requêtes et celui de prise en compte de nouveaux objets intelligents sont fortement corrélés.

Afin d'effectuer ces deux tâches, une application web en Javascript/Python a été

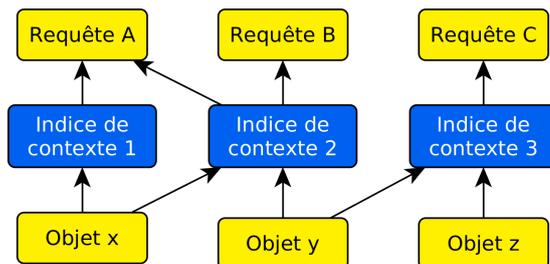


FIGURE 4.16 – Graphe des requêtes, indices de contexte et descriptions d’objets.

développée. Installée sur le *back-end*, elle offre la possibilité de manipuler visuellement la base de données d’indices contextuels et de requêtes (figure 4.17).

## 4.4.2 Bases de connaissances

Les raisonnements du système 2SEES ne sont pas faits sur une seule base de connaissances mais sur un ensemble de bases, distribuées entre les acteurs et son propre domaine. La première partie, la plus importante, est constituée par les descriptions des objets en WoT-TD, qui contiennent des liens vers les bases de données des objets, qui forment donc une autre partie des bases de connaissances. Les descriptions peuvent également être liées, par des relations RDF, avec des bases de connaissances soit internes aux gestionnaires d’objets, soit externes, formant encore un ensemble de bases de connaissances. Il est à noter que les données des objets peuvent être encodées en RDF ou non. La seconde partie importante est la base de connaissances du système 2SEES, construite pour définir et permettre le traitement des requêtes. On peut subdiviser cette base de connaissances en deux grandes parties : d’un côté, la logique particulière des requêtes, de l’autre, la logique partagée entre toutes les requêtes – une modélisation générale de l’environnement. Encore une fois, grâce au web sémantique et à la puissance du RDF et des couches supérieures, toutes les connaissances peuvent être liées à des ressources externes. Toutes ces bases de données sont liées et en quelque sorte assemblées par le 2SEES quand il répond à des requêtes (figure 4.18) ; cela se fait de manière transparente grâce à l’orientation vers le *web* de la pile de standards du W3C.

## 4.4.3 Requêtes SQWRL

L’encodage de la logique des requêtes en règles SWRL et en concepts et relations OWL génère la définition d’éléments qui sont seulement utiles à la réponse aux requêtes. Par exemple, la requête « Pourquoi fait-il froid ? » implique l’introduction d’un concept *cause possible* ; cela est dû à la structure du sujet de la requête, c’est-à-dire un modèle physique de la température de la pièce et de ses influences. La logique qui gouverne l’objet de la requête ne permet pas une réponse simple tenant à l’appartenance à une classe existante, et l’introduction d’un nouveau concept est nécessaire.

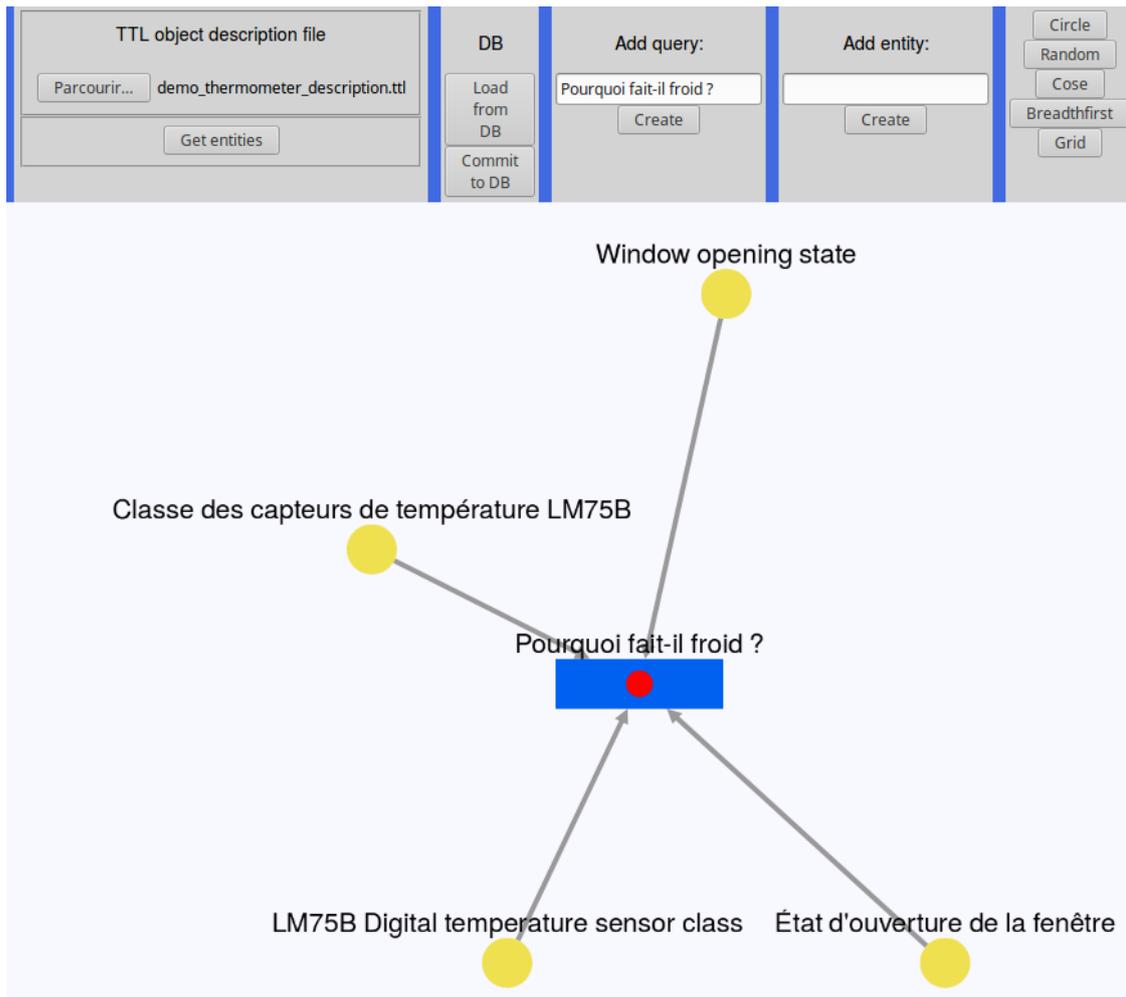


FIGURE 4.17 – Capture d’écran de l’application web de gestion des requêtes et des indices de contexte. La requête « Pourquoi fait-il froid ? » est ici liée à quatre indices contextuels (deux, dupliqués en français et anglais) : l’état d’ouverture de la fenêtre, et le capteur de température LM75B.

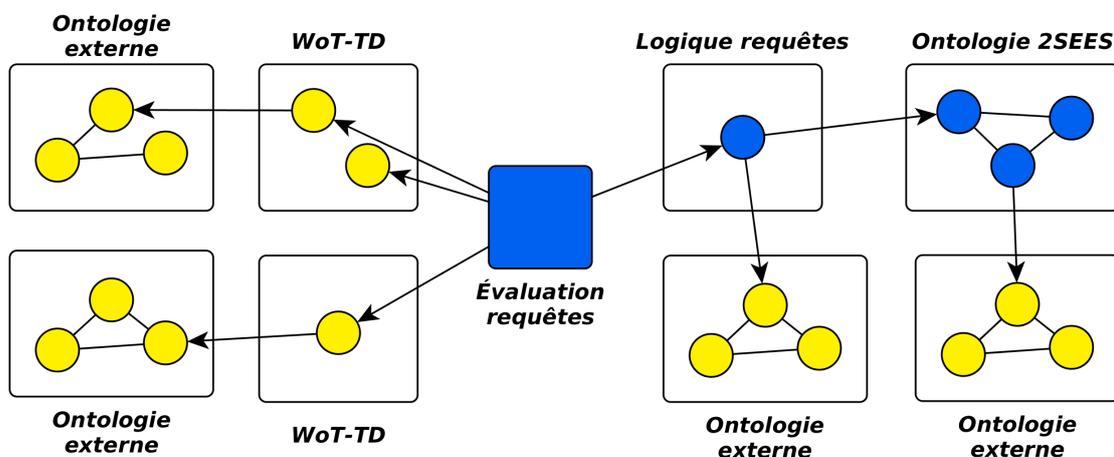


FIGURE 4.18 – Ensemble des bases de connaissances potentiellement mobilisées pour la réponse à une requête.

Ce type de concept représente une forme de « pollution » de la base de connaissances : il n'exprime pas un fait réel, ne décrit pas l'environnement, mais sert seulement à former la réponse à la requête. Une première manière de contenir cette pollution a été décrite plus haut : la séparation de la logique des requêtes d'avec la partie de la base de connaissances servant à décrire l'environnement. Le langage SQWRL (acronyme de l'anglais *Semantic Query-enhanced Web Rule Language*) représente une seconde manière de diminuer le bruit dans la base de connaissances. SQWRL permet de faire des requêtes directement sur une base de connaissances [O'Connor and Das, 2009] ; c'est, en quelque sorte, un hybride de SPARQL et de SWRL. Une requête SQWRL a la même forme qu'une règle SWRL. La différence se trouve dans le rôle que tient la *tête* de la formule : dans une règle, la tête est le *conséquent*, qui est considéré comme valide si le corps, l'*antécédent*, est valide ; en SQWRL, la tête est le *retour de la requête*, c'est-à-dire ce que la requête SQWRL retourne quand le conséquent est valide. Dans notre cas, il est alors très simple et avantageux de transformer une règle exprimant la logique d'une requête sémantique 2SEES en requête SQWRL. Par exemple, la requête SQWRL 4.11 renverra tous les radiateurs dont le point de fonctionnement est fixé à une température inférieure à 17 °C. Ainsi, SQWRL permet de s'affranchir des concepts créés dans le seul but de compiler les résultats d'une règle pour ensuite les interroger avec une requête SPARQL.

$$\begin{aligned}
 & \textit{externe} : \textit{radiateur} (?x) \\
 \wedge & \textit{externe} : \textit{parametrageRadiateur} (?x, ?temperature) \\
 \wedge & \textit{swrlb} : \textit{lessThanOrEqual} (?temperature, 18) \\
 \rightarrow & \textit{sqwrl} : \textit{select} (?x)
 \end{aligned} \tag{4.11}$$

#### 4.4.4 Gestion des erreurs

À côté de leur utilisation comme assistance à la perception spatiale et non-spatiale, les objets intelligents peuvent être utilisés pour la détection et la correction d'erreurs. En effet, les informations contextuelles extractibles à la fois de l'observation des occurrences des objets, de leurs descriptions formelles et des données qu'ils génèrent peuvent être utilisées pour vérifier les connaissances manipulées par d'autres composants du système 2SEES. La plupart de ces connaissances proviennent d'une manière ou d'une autre de données de capteurs dont on a eu un aperçu des limites en termes de précision et de validité. Toutes les informations présentes dans le système sont susceptibles d'être incorrectes.

Nous allons voir quelques cas où les informations relatives aux objets intelligents permettent de limiter les risques de défaillance du système, en permettant la détection et éventuellement la correction des erreurs à l'origine des défaillances.

### Localisation et navigation

Tout d'abord, prenons le cas des erreurs sur la position – sous-fonction essentielle de la navigation. Comme exposé dans la section 3.1.3, un certain degré d'imprécision dans la localisation est inévitable, à cause des capteurs et des méthodes de calcul. À partir d'un certain niveau, cette imprécision dégrade la bonne marche des fonctionnalités du système, par exemple en ayant un impact direct sur la navigation. Nous souhaitons idéalement une précision correspondant grossièrement à un petit pas, ce qui représente environ vingt à quarante centimètres. Si l'erreur sur la position est plus importante, d'environ un mètre, la personne ne peut plus se reposer sur la fonction de navigation pour réaliser des tâches simples relevant de la marche, telles le repérage des portes ou la connaissance du côté du couloir où elle se trouve. Elle peut encore largement profiter de la fonctionnalité pour se déplacer à une échelle plus large, ce qui relève de l'orientation. Plus l'erreur sur la position augmente, moins large est la gamme des tâches pour lesquelles le système est utile à la personne.

À un certain point, l'erreur est si grande que le système devient inutile à la personne. Ce peut être le cas si, par exemple, le système estime être dans un bâtiment alors qu'il est dans un autre, ou, à plus grande échelle, si la ville ou le pays estimés et réels diffèrent. Bien qu'ils aient l'air artificiellement exagérés, il est probable que ces cas puissent être rencontrés en pratique. Si le système est éteint avant un déplacement en voiture, en train, ou en avion, acte naturel, il se trouvera à son redémarrage dans un lieu radicalement éloigné de la dernière position qu'il connaissait, et qui devrait être celle qui est estimée comme étant la position courante. Le même type de défaillance au redémarrage peut être obtenu suite à une remise à zéro, volontaire ou non, effaçant l'état interne du système, y compris l'estimation de position. Une erreur dans une carte peut mener à une défaillance de cette ampleur de manière plus pernicieuse. L'échelon local du système ne manipule que des fragments de carte qui sont liés entre eux par certains points, par exemple une porte menant à un autre fragment de carte. Une erreur sur le fragment pointé par la porte générera une estimation de position complètement faussée lors du franchissement de cette porte.

Les informations données par les objets intelligents environnants peuvent permettre d'éviter ces défaillances. La mise en évidence des petites et moyennes imprécisions de localisation par l'utilisation des objets intelligents est difficile car elle implique que ceux-ci aient des données précises sur leur propre localisation. Les infrastructures de balises radio remplissent cette condition, mais étant destinées à la localisation, leur usage par le système 2SEES les placerait plus dans la catégorie *contribution à la fonctionnalité* que dans la catégorie *détection d'erreur sur la fonctionnalité*.

Les portes intelligentes représentent des objets intéressants pour la détection des imprécisions sur la localisation. Tout d'abord, leur positions exactes sont connues, ou bien faciles à connaître puisque, d'une part, elles sont parfaitement fixes<sup>15</sup>, et, d'autre part, elles font partie des plans des bâtiments faits et fournis par les

---

15. Contrairement au « mobilier ».

architectes. L'information concernant leur position est donc déjà existante et nécessite « seulement » un travail d'intégration dans une carte publique ou spécifique au système 2SEES. Une porte intelligente est également utile car elle offre une connaissance assez précise de l'endroit où se trouve la personne qui l'utilise. C'est le cas d'une porte qui observe simplement quand elle est manipulée, et c'est encore plus vrai pour une porte capable de détecter quand une personne la traverse (on peut aisément imaginer des applications pour de telles portes). L'exploitation des données de ce type de porte par le 2SEES est envisageable à condition que celui-ci puisse s'assurer avec une confiance raisonnable que c'est bien la personne qui l'utilise (et pas une autre) qui a interagi avec la porte en question.

La détection des erreurs de grande ampleur est plus simple. La simple détection d'un objet connecté offre déjà une information relativement précise sur sa propre position ; l'accès aux métadonnées et aux données encore plus. Il sera facile pour le système de détecter l'erreur devant l'évidence de l'écart entre l'estimation interne et les indices donnés par les objets. Si détecter de telles erreurs devrait être simple dans un environnement riche en objets intelligents, les corriger est plus difficile. Pour revenir à un niveau de précision à même de permettre le bon fonctionnement de la navigation, il n'est en effet pas suffisant de savoir dans quelle ville, quel bâtiment, ou même quel périmètre autour d'un objet dont la position est connue on se trouve ; il faut pouvoir redescendre à un niveau d'erreur de moins d'un mètre. On tombe alors dans les mêmes problèmes que pour les erreurs plus légères.

Les erreurs de localisation sont globalement difficiles à corriger : elles nécessitent des objets dont la position est connue précisément, ce qui est loin d'être évident quand cela ne fait pas partie de la fonction de l'objet.

### Détection d'obstacles

Des défaillances peuvent également avoir lieu dans la fonctionnalité de détection d'obstacles ; une défaillance est constituée par l'absence de détection d'un obstacle réel. Des objets peuvent ne pas être détectés pour trois raisons : à cause de leur forme et de leur position (p. ex. : un câble suspendu), à cause du matériau qui les constitue (par exemple une porte en verre), ou bien à cause de leur mobilité (cas d'une voiture).

Le cas des objets non-détectés à cause de leur forme est difficile. D'une part parce que les capteurs sont conçus pour assurer la détection de tous les objets, intelligents ou pas, et d'autre part parce qu'on retombe encore une fois dans les mêmes problèmes qu'on a avec la correction d'erreur de positionnement :

1. La position relative de l'objet par rapport à la personne doit être connue précisément, ce qui implique un des points qui suivent ;
2. Soit la position de la personne *et* celle de l'objet sont connues avec une très grande précision et on peut alors vérifier si leurs trajectoires vont se rencontrer ;
3. Soit la position relative est accessible directement par un autre moyen tel que la détection visuelle avec une caméra, par le 2SEES ou bien par l'objet.

Il paraît peu probable que des objets intelligents localisent avec précision la personne. Le contraire est plus envisageable : on a décrit dans la section 4.2.3 un mécanisme dans lequel un objet est détecté avec une caméra une fois son apparence connue, ce qui permet d’avoir une information directe sur la position relative de l’objet.

Pour les objets dont ce sont les caractéristiques du matériau qui posent problème aux capteurs, comme, par exemple, les portes en verre, il est possible que le système utilise l’information de leur présence pour activer ou augmenter la priorité des capteurs qui sont à même de les détecter – c’est-à-dire des capteurs ultrasoniques dans le cas d’une porte en verre. Ici, l’identification par le réseau joue un rôle de facilitation et d’augmentation de la fiabilité de l’algorithme de détection d’obstacles. On peut également informer la personne de l’état de verrouillage d’une porte en avance pour éviter toute surprise.

Le cas le plus intéressant est probablement le troisième, celui des obstacles difficiles à détecter à cause de leur mobilité. En effet, on assiste aujourd’hui à un fort développement des systèmes de transports intelligents ou STI (*intelligent transportation systems* ou *ITS* en anglais), avec par exemple des véhicules autonomes, et des communications entre les véhicules et entre les véhicules et les infrastructures. Or, un composant essentiel des systèmes de transports intelligents est la localisation des véhicules, qu’elle soit mesurée par les véhicules eux-mêmes (GPS, odométrie) ou bien par les infrastructures (caméras, radars, etc.) ; ces données sont partagées le long de la route afin d’augmenter l’efficacité et la sécurité du réseau routier (par exemple, pour éviter les embouteillages).

Leur utilisation par le système 2SEES s’inscrit parfaitement dans le cadre de l’augmentation de la sécurité. Les véhicules suspectés de ne pas s’arrêter à un feu parce qu’en excès de vitesse, ou parce qu’ils ont déjà brûlé un feu précédent, pourraient être signalés à la personne utilisatrice. Symétriquement, la présence de la personne pourrait être signalée aux véhicules, leur permettant soit d’alerter leur conductrice ou conducteur, ou bien, en cas de danger, de freiner directement. La connectivité pourrait être assurée grâce à des balises spécifiques à proximité des passages piétons, par une adaptation du 2SEES permettant la connexion aux infrastructures, ou bien, comme précédemment, par un accès aux données par Internet.

#### 4.4.5 Problématiques d’interopérabilité

Toutes les interactions décrites jusqu’ici impliquent un niveau élevé d’interopérabilité entre le système 2SEES et les différents acteurs liés aux objets intelligents. Nous sommes proches de la définition même de l’interopérabilité : la capacité pour des systèmes à fonctionner ensemble.

Les métadonnées décrivant les objets et les interactions possibles avec eux ainsi que les données qu’ils produisent doivent être encodées dans un format que le système 2SEES peut comprendre au niveau sémantique.

On l’a vu, la fonctionnalité SO2SEES repose sur plusieurs hypothèses :

- Les objets intelligents sont identifiés par une IRI ;

- Un mécanisme de découverte locale des IRI est implémenté ;
- Cette IRI mène à une description formelle des objets ;
- Cette IRI mène également aux données des objets, de manière directe ou indirecte.

L'utilité de la fonctionnalité SO2SEES est directement liée à sa capacité à interagir avec un grand nombre d'objets intelligents. Plus large est l'éventail d'objets interopérables, meilleure sera la fonctionnalité : d'une part, parce que la couverture géographique augmente avec le nombre d'objets, et, d'autre part, parce que les possibilités d'interactions s'enrichissent avec la densité des objets et surtout leur diversité, qui permet l'exploitation de synergies entre leurs fonctionnalités individuelles pour proposer des requêtes plus synthétiques et concises.

Cependant, de telles interactions avec des entités externes sont possibles uniquement si les parties en présence arrivent à s'accorder sur l'utilisation de standards communs – de préférence ouverts – qui définissent les mécanismes intervenant à chaque niveau d'interaction, depuis la connectivité physique jusqu'à l'échange d'informations sémantiques. Malheureusement, il est peu probable que la prise en compte des problématiques liées au handicap soit un sujet central dans la conception et l'adoption de tels standards par l'industrie de l'Internet des objets : cela ajoute de la complexité, et ces problématiques pèsent peu dans la balance économique en comparaison avec les applications touchant une plus grande population. Pour s'en convaincre, on peut observer la difficulté à faire appliquer de tels standards dans le cadre pourtant théoriquement contraint par la législation qu'est l'urbanisme. Puisqu'il est peu probable que les acteurs industriels tiers s'alignent sur le système 2SEES, c'est au système 2SEES de s'adapter à eux : il faut suivre autant que possible les standards et les concepts adoptés ou – plus difficile – en passe d'être adoptés.

L'architecture du système 2SEES doit elle-même se soumettre aux contraintes venant de la nécessité d'interopérabilité car l'architecture a une influence sur la compatibilité avec les différentes formes que prend l'environnement IoT. Le système doit être conçu de manière à minimiser les contraintes sur les acteurs externes, ou plutôt, plus précisément, à minimiser la friction générée par les contraintes : une contrainte ne pose pas de problème si elle est satisfaite par défaut, ce n'est plus une contrainte qui pèse sur la réussite, elle est en quelque sorte « gratuite ». Nos choix dans la conception du système ont donc été dirigés en grande partie par ces contraintes.

### Accès aux objets intelligents

Le premier de ces choix concerne l'accès aux objets intelligents. Ce choix est qualifié de premier à la fois par sa position temporelle dans la conceptualisation, et pour le rôle déterminant qu'il exerce sur d'autres choix ultérieurs comme les communications bas niveau et l'adressage des objets. Deux alternatives sont disponibles pour les communications entre le système 2SEES et les objets intelligents :

- Des communications directes par les couches basses du modèle OSI ou de la pile TCP/IP, c'est-à-dire sans fil sur un même médium physique ou segment

- réseau ;
- Des communications indirectes, passant par les couches supérieures du modèle OSI ou de la pile TCP/IP (couche réseau/IP ou plus hautes).

La connectivité sans fil directe permet des interactions rapides et dispense l'échelon local du 2SEES (LMSC et smartphone) de la nécessité d'être en permanence connecté à Internet, mais impose logiquement une compatibilité au niveau des couches basses du réseau. L'accès indirect ne pose pas de contraintes sur les objets eux-mêmes mais implique une infrastructure lourde : accès à Internet partout et surtout plates-formes *back-end* accessibles derrière les objets. Nous pensons que les contraintes de l'accès indirect sont moins lourdes ; il déplace essentiellement les contraintes depuis les objets intelligents vers les éléments placés plus haut dans la hiérarchie : serveurs locaux et surtout *back-ends*. Dans beaucoup des applications qui constituent l'Internet des objets, les objets en question sont déjà soumis à de très fortes contraintes, particulièrement en termes de consommation d'énergie et de puissance, toutes deux très limitées. Dans ces conditions, il paraît peu raisonnable de leur imposer des contraintes supplémentaires : diffusion périodique de messages contenant leur identifiant pour la découverte, sortie de mode veille pour répondre à des requêtes, etc. Au niveau matériel et des protocoles de bas niveau, les objets sont techniquement adaptés à leur application (consommation, portée, débit), et il est peu probable que le LMSC et le smartphone embarquent à eux deux suffisamment de matériel et les logiciels nécessaires à une connectivité directe avec chacun d'entre eux. Un exemple extrême mais pertinent de ce problème est la connectivité filaire : elle est impossible pour le 2SEES, alors que des objets intéressants comme les systèmes de contrôle de la température et de l'éclairage l'utilisent. On peut utiliser des passerelles sans fil comme des points d'accès WiFi, mais il faudrait alors idéalement assurer une couverture dans l'intégralité d'un bâtiment.

L'accès indirect résout ce problème. Les contraintes qui lui sont associées sont plus légères : un accès à Internet mobile (3G, 4G, 5G ou WiFi) est très probable dans les environnements riches en objets intelligents. La seconde contrainte, l'accès possible aux données des objets intelligents par une IRI, est plus délicate. Nous croyons néanmoins qu'elle est susceptible d'être satisfaite : le transport des données vers des plates-formes sur Internet simplifie grandement le traitement des données agrégées. La généralisation de l'utilisation d'IRI est comparativement moins probable mais c'est en fait une contrainte assez souple : il est aisé d'adapter le système à plusieurs méthodes d'accès. Le point crucial est l'ouverture de l'accès aux données des objets ; celle-ci soulève des problèmes de sécurité potentiellement difficiles à gérer que nous ne traiterons pas ici, et n'est probablement pas dans l'intérêt le plus immédiat des gestionnaires d'objets. Les données sont une ressource précieuse qu'il est tentant de garder pour soi. Néanmoins, la généralisation des objets intelligents devrait mener les fabricants et gestionnaires à partager une partie non-négligeable des données, afin d'offrir des services plus avancés que ce que permettent des architectures en silos non interopérables – des services comme celui que vise à offrir la fonctionnalité SO2SES.

La présence du serveur local comme nœud-passerelle dans un réseau d'objets intelligents est chose commune. WoT-TD mentionne deux types d'équipements de ce type dans la liste des architectures dont la prise en charge est visée : les passe-

relles simples et les *edge routers*, soit des passerelles dotées de capacités de calcul et jouant un rôle actif dans le processus [Kovatsch et al., 2019]. Le second cas correspond aux capacités du serveur local.

### Découverte des objets

Un autre choix important est celui de la méthode de découverte des objets. Le premier mécanisme, la transmission locale par les objets, est le plus direct, voire le plus naturel, mais présente un désavantage majeur : il pèse sur les fabricants d'objets, acteurs externes qui se verraient imposer l'implémentation d'un mécanisme de diffusion de messages sur médium local dans leurs objets intelligents pour la compatibilité avec la fonctionnalité SO2SEES. Malheureusement, leur marge de manœuvre sur cette contrainte est négativement liée à la facilité d'implémentation côté système 2SEES : plus les fabricants ont de liberté sur le mécanisme précis de diffusion locale des IRIs<sup>16</sup>, plus il sera difficile pour les développeurs du système 2SEES d'assurer la compatibilité avec les différents fabricants d'objets intelligents car il sera nécessaire de s'adapter à l'implémentation mise en place par chaque fabricant. L'adaptation requiert l'achat des équipements en question, le développement d'un code spécifique, et surtout, de tests ; de plus, cela veut potentiellement dire une prise en charge de nombreux protocoles sans fil par le 2SEES avec la complexité que cela implique : plusieurs émetteurs-récepteurs, éventuellement plusieurs antennes, des pilotes de périphériques et des piles réseaux adaptés, et surtout la gestion de la commutation entre les différents protocoles pour la gestion de l'énergie.

Reporter la charge sur les fabricants d'objets en contraignant le mécanisme de diffusion peut alors paraître plus simple mais ce serait probablement une erreur. D'une part, c'est techniquement réducteur : une seule méthode n'est pas forcément adaptée à tous les équipements, qui doivent avant tout satisfaire les contraintes découlant de leurs fonctionnalités. D'autre part, ce serait une compromission de l'interopérabilité et donc de la capacité du système à être adopté, car toute contrainte mise sur les acteurs externes augmente le risque qu'ils fassent le choix de la simplicité en n'adoptant pas la norme.

La seconde possibilité requiert des bases de données géographiques à jour et accessibles, idéalement maintenues par les gestionnaires d'objets intelligents ; elle requiert également que la fonction d'estimation de position du système 2SEES soit fonctionnelle et suffisamment précise. Le second aspect de l'accès aux objets est l'adressage. Dans le modèle OSI, ou dans la pile standard TCP/IP, un adressage à la couche 3 (réseau/IP<sup>17</sup>) pourrait être envisagé. Mais nombre d'objets ne possèdent pas les ressources nécessaires à une connexion directe à un réseau IP. Un adressage indirect, de plus haut niveau, permet de régler ce problème : la machine pointée par l'adresse peut être l'objet lui-même, mais également une machine tierce, représentante intermédiaire de l'objet. Nous pouvons découpler l'adresse de l'objet lui-même car nous n'avons pas besoin de lui, mais de ses données. Nous avons choisi

---

16. Ce qui signifie par exemple qu'ils peuvent utiliser le protocole de diffusion et le protocole réseau (802.15.4, WiFi, Bluetooth, etc.) de leur choix.

17. En IPv6 en tout cas, le nombre d'adresses étant en principe suffisant pour tous les objets.

d'utiliser des IRI déréférencables à la fois comme identifiants et comme méthode d'accès aux objets – c'est-à-dire à leurs données et métadonnées hébergées sur leurs plates-formes *back-ends*. Les IRI présentent l'avantage d'être standardisées et largement utilisées [Duerst and Suignard, 2005]; en particulier, elles sont le mécanisme d'identification des concepts en RDF [Wood et al., 2014].

Dans l'implémentation concrète du scénario, c'est une troisième voie qui a été choisie, dans laquelle les IRI des objets sont diffusées par le serveur local. Cette approche semble n'apporter que des avantages par rapport à la première solution envisagée :

- Elle implique une implémentation beaucoup moins large du mécanisme spécifique car il y a normalement plus d'objets intelligents que de passerelles. Du côté du 2SEES, cela signifie que moins d'implémentations concurrentes et potentiellement différente du standard sont à gérer (tests en particuliers). De leur côté, les fabricants de matériel sont moins nombreux à être impactés, et sur moins de produits ;
- Les problèmes d'incompatibilité entre la diffusion périodique de messages et les fortes contraintes énergétiques de certains objets intelligents sont éliminées : on peut imaginer que les serveurs locaux disposent d'une alimentation par le secteur.

Tout en évitant les problèmes de la seconde solution :

- Il n'y a pas de contrainte de cartographie des objets connectés, qui était nécessaire pour réaliser le registre géographique (qui est une forme de carte) ;
- Les contraintes sont posées sur les fabricants de matériel plutôt que sur les gestionnaires.

Enfin, ce changement permet de paramétrer la diffusion des messages. Nous avons justement profité de ce dernier avantage pour affiner la précision du « filtrage » contextuel géographique. Nous souhaitons que la présence des objets soit portée à la connaissance de la fonctionnalité SO2SEES quand celle-ci se trouve dans le même lieu qu'eux. Si les objets sont chargés de la diffusion des messages (premier mécanisme envisagé), l'échelon local du 2SEES les détectera une fois « à portée » ; la proximité spatiale joue un rôle essentiel dans la portée des émetteurs et récepteurs mais elle n'est pas seule ; les obstacles, murs, caractéristiques et orientations des antennes y ont par exemple une influence. L'aspect géographique du filtrage n'est alors pas vraiment contrôlé, il est même, en quelque sorte, accidentel.

En revanche, si la diffusion est à la charge du serveur local, on peut la soumettre à des conditions précises. Dans la configuration choisie pour notre scénario, on peut en particulier subordonner la diffusion à l'état de la porte : *si* la porte est ouverte (ou bien ouverte puis fermée), *alors* le serveur local diffuse les IRIs. Ainsi on a une très grande probabilité que l'échelon local du 2SEES ne soit informé des IRIs que lorsqu'il en a réellement besoin, c'est-à-dire une fois que la personne qui l'utilise entre dans une pièce<sup>18</sup>. Par extension, le même serveur local peut gérer plusieurs pièces, chacune équipée d'une porte intelligente ; selon la porte poussée, le serveur local pourra alors diffuser la liste précise des IRIs se trouvant dans la pièce, plutôt que la liste de toutes les IRIs des objets qu'il gère.

18. La possibilité alternative étant que quelqu'un d'autre utilise la porte quand la personne déficiente visuelle utilisant le système 2SEES se trouve à portée du serveur local.

## Adoptabilité du WoT-TD

Nous avons choisi le standard WoT-TD à la fois parce qu'il correspond bien à notre application et parce qu'il semble avoir de bonnes chances d'être adopté. En effet, c'est un standard du W3C, le principal organisme de standardisation du web, et du web sémantique. Les travaux au sein du W3C sont également menés en très étroite collaboration avec des industriels qui fabriquent des objets connectés. Ainsi, des implémentations de WoT-TD dans des objets existants sont testées. Dans une liste sur la page [Kaebisch et al., 2019b], on trouve des objets connectés testés : des hubs pour ampoules connectées Philips, Panasonic et Ikea, des aspirateurs et des climatiseurs Panasonic, des composants industriels Siemens et Festo, etc. Tous les aspects du standard sont évalués. L'approche de WoT-TD, basée sur des descriptions simples, constitue un autre facteur favorisant son adoptabilité future.

### 4.4.6 Apports perceptifs

Dans un second temps, on peut envisager de combiner WoT-TD avec SSN et SOSA. Les deux ontologies ont l'immense intérêt d'apporter, en plus de la modélisation des objets eux-même, la modélisation de leur lien avec l'environnement, voire une modélisation basique de l'environnement. Ces connaissances pourraient être renforcées et utilisées dans le cadre d'une nouvelle ontologie spécifique aux personnes malvoyantes et aveugles.

Les questions soulevées par les deux versants de la fonctionnalité sont différentes. Pour ce qui est de l'assistance à la perception spatiale, le problème principal est de savoir si une telle fonctionnalité, malgré sa simplicité – la description verbale simple de l'environnement – est à même d'enrichir la représentation spatiale d'une personne. Nous le pensons ; la perception spatiale chez les humains adultes semble se baser prioritairement sur les objets et utiliser le langage [Wang and Spelke, 2002]. La question principale soulevée par l'aspect non-spatial de l'assistance est celle du niveau d'assistance, déjà évoquée dans le chapitre 2. Plus les interactions avec le système 2SEES seront limitées à des requêtes de bas niveau, plus les utilisateurs devront passer du temps à spécifier quels attributs ils veulent connaître, rentrer dans le détail du fonctionnement des objets intelligents, passer du temps à combiner les informations reçues en réponse, et à poser sur elles des jugements. D'un côté, cela signifie que les requêtes de haut niveau augmentent l'efficacité du système ; mais de l'autre côté, cela correspond à une augmentation du rôle du 2SEES dans le partage des tâches du système { personne + dispositif d'assistance }. Une partie du travail de réflexion est déplacée de la personne vers le système d'assistance. Avec les défauts que cela implique : dirigisme ou perte d'initiative, et uniformisation des expériences.

Il est intéressant de voir l'utilisation du terme *affordance* dans le vocabulaire de WoT-TD : on y trouve la définition de concepts tels que *PropertyAffordance*, *ActionAffordance*, etc. Dans le cadre de WoT-TD, ce mot signifie la possibilité d'action offerte par un objet. Ce mot provient des travaux de Gibson [Gibson, 2014] que nous avons évoqués dans la section 1.3.1. Chez Gibson, l'environnement offre aux

êtres des possibilités d'interactions : on peut marcher sur le sol, voir les choses, respirer l'air, etc. La perception est un élément essentiel de ces interactions. On a dans notre situation une transposition de ces interactions dans le contexte des objets intelligents : un bouton offre la possibilité d'être actionné, les propriétés peuvent être exposées, etc. On a donc un parallèle entre la perception de l'environnement naturel et de l'environnement artificiel qui se trouve formalisé dans le vocabulaire même du système servant à décrire ce dernier. Gibson a écrit « *The world of physical reality does not consist of meaningful things. The world of ecological reality, as I have been trying to describe it, does.* »<sup>19</sup> [Gibson, 2014] ; en ouvrant des possibilités d'interactions avec les objets intelligents, nous espérons créer un nouvel environnement écologique.

---

19. Ce qu'on pourrait traduire par : « Le monde de la réalité physique n'est pas formé de choses dotés d'un sens. Le monde de la réalité écologique, que je me suis appliqué à décrire, l'est. »



# Conclusion

## Bilan

Dans cette thèse, nous nous sommes intéressés à la création d'une aide électronique à la mobilité des personnes déficientes visuelles. Cette problématique générale est large, et nous avons d'abord dû circonscrire le cadre dans lequel nous avons réalisé nos travaux : choix d'un modèle de mobilité, explicitation du type de déficience visuelle ciblée, et aperçu des différentes approches de l'étude la perception.

Ce cadre nous a permis d'analyser une assez grande diversité de dispositifs existants, sous plusieurs angles : par rapport à leurs fonctions, leurs technologies, leurs interfaces avec la personne, mais aussi leurs positionnements perceptifs. À notre connaissance, ce dernier critère d'analyse constitue une approche novatrice d'analyse des dispositifs d'assistance. Il présente notamment l'intérêt de s'appliquer indifféremment des considérations techniques des dispositifs, étant valable pour une canne classique comme pour le service d'aide fourni par le personnel d'un aéroport.

Au cours de cette thèse, le système 2SEES a été conceptualisé, conçu et réalisé. Continuation du projet SEES et de son prototype, le 2SEES comprend des avancées à la fois conceptuelles et techniques par rapport au SEES et aux dispositifs existants. La prise en compte explicite de la tolérance aux fautes et de la robustesse du système constitue une de ces avancées. Afin de présenter une interface unique et cohérente aux utilisateurs, le système 2SEES réunit trois familles de fonctionnalités dans un même système : assistance à la détection d'obstacles, assistance à la navigation, et assistance à la perception de l'environnement. La conception mécanique du prototype de canne intelligente, son architecture électronique et son système d'exploitation ont été orientés vers la création d'un dispositif tolérant aux fautes.

La détection d'obstacles est une sous-fonction indispensable de la mobilité, en raison des risques de heurt et de chute que posent en permanence les obstacles et les singularités de l'environnement. Elle est donc logiquement au cœur d'un grand nombre d'aides à la mobilité, au premier rang desquelles se trouve la canne blanche classique. Pour la fonction de détection du 2SEES, nous nous sommes focalisés sur la robustesse plutôt que sur la performance spatiale de la fonction de détection, sa rapidité, ou sur sa précision, domaines largement explorés dans la littérature. Elle repose sur des groupes sensoriels intégrant à la fois des capteurs redondants (identiques) pour la tolérance aux pannes, et des capteurs hétérogènes pour la robustesse informationnelle, c'est-à-dire la capacité du système à faire face à des

variations importantes de l'environnement.

La conception de la fonctionnalité d'assistance à la navigation a été guidée par les contraintes induites par l'idéal d'universalité choisi. Cela impliquait en particulier la possibilité de se passer d'infrastructures de localisation et d'abaisser la consommation d'énergie du système à un niveau compatible avec l'application. La solution retenue est une version simplifiée de filtre particulière, utilisant les données de capteurs inertiels et d'une roue munie d'un encodeur rotatif, et une cartographie du bâtiment. Ce filtre est léger et s'exécute en temps réel sur un smartphone.

La maîtrise de la consommation d'énergie est un aspect crucial du système. Elle est obtenue grâce à l'application du principe d'*active multi-sensor context awareness*, qui signifie l'activation et la désactivation des capteurs et des composants en fonction de leur utilité dans le contexte courant, ainsi que les choix que cela implique : présence de capteurs hétérogènes, analyse du contexte, etc. Les fonctionnalités de détection d'obstacles et de navigation sont ainsi deux applications partielles de ce principe.

Nous avons conceptualisé et implémenté une fonctionnalité permettant aux personnes déficientes visuelles d'interagir avec les objets intelligents. Nommée SO2SEES, celle-ci fonctionne par le biais de requêtes : la personne peut interroger le système 2SEES, qui agglomère les données des objets intelligents présents autour d'elle pour répondre à ses requêtes. Cette fonctionnalité s'inscrit dans le cadre de l'assistance à la perception de l'environnement ; elle est novatrice, car l'environnement avec lequel elle interagit n'est pas constitué par le monde des objets physiques simples mais par celui des objets intelligents. Les objets intelligents constituent un substrat très intéressant, car il est possible d'exploiter les informations qu'ils génèrent pour fournir des réponses concises et de haut niveau aux utilisateurs. Il est à noter qu'une croissance de la densité et de la diversité des objets intelligents dispersés est probable dans les prochaines années.

En résumé, les contributions de cette thèse sont au nombre de cinq :

1. Fonctionnalité SO2SEES pour la perception des objets intelligents ;
2. Analyse des dispositifs d'assistance selon le positionnement perceptif ;
3. Prise en compte de la tolérance aux fautes et de la robustesse, y compris informationnelle ;
4. Réalisation d'un prototype de plate-forme multi-fonctionnelle et multi-capteurs ;
5. Conception et implémentation d'un algorithme léger de suivi de la trajectoire.

## Perspectives

Cette thèse ouvre de nombreuses perspectives. La plus urgente est probablement la mise en place d'expérimentations complètes pour valider plus concrètement les aspects développés, particulièrement la fonctionnalité SO2SEES. En effet, la nouveauté et l'intérêt de cette fonctionnalité résident dans les informations qu'elle peut apporter à des utilisateurs et dans la manière dont elle effectue cette tâche.

De manière générale, les fonctionnalités ont été développées de façon relativement indépendante les unes des autres, et gagneraient à être mieux intégrées ensemble. Cela passerait particulièrement par une utilisation universelle de la localisation estimée et des données géographiques. Les questions de cartographie soulevées par cet objectif sont intéressantes mais complexes, car les éléments avec lesquels les différentes fonctionnalités du système 2SEES interagissent sont hétérogènes et absentes des données cartographiques standards.

Le mécanisme d'*active multi-sensor context awareness* vise à contrôler l'intégralité du fonctionnement matériel du système. En l'état, il n'est implémenté qu'à un niveau limité, et son amélioration passerait, d'un côté, par une intégration poussée dans les systèmes d'exploitations et l'électronique, et, de l'autre, par une meilleure acquisition, estimation, et représentation des informations contextuelles.

Plus prosaïquement, un certain nombre de sous-fonctionnalités envisagées dans la conception du 2SEES n'ont été que partiellement implémentées et mériteraient d'être finalisées. C'est notamment le cas d'une partie des fonctionnalités de tolérance aux fautes du système d'exploitation.

Enfin, le pré-prototype du 2SEES conçu et réalisé au cours de cette thèse doit encore être amélioré : il faut poursuivre sa miniaturisation et son intégration dans un format plus pratique. Ces perfectionnements sont particulièrement importants en raison de l'influence de la praticité du système sur les expérimentations, et surtout sur sa future utilisation par les personnes malvoyantes et aveugles, qui devraient être les premières bénéficiaires.



# Bibliographie

- Adenike Abiose, August Colenbrander, Christian Corbé, Terry A. Cox, Leon B. Ellwein, Astrid Fletcher, Y. Hiratsuka, Brien Holden, Lea Hyvärinen, Ahmed Trabelsi, Rohit Varma, Karla Zadnik, and Jialang Zhao. Consultation on development of standards for characterization of vision loss and visual functioning. Recommendations, World Health Organisation (WHO), Geneva, Switzerland, September 2003.
- Académie Française. Mobilité. In Florence Monier, Jean-Mathieu Pasqualini, and Yves Pouliquen, editors, *Dictionnaire de l'Académie Française*, page 349. Académie Française, Paris, 9e edition, 2019.
- Muhammad Haris Afzal, Valérie Renaudin, and Gérard Lachapelle. Assessment of indoor magnetic field anomalies using multiple magnetometers. In *Proceedings of the 23rd International Technical Meeting of the Satellite Division of The Institute of Navigation*, pages 525–533, Portland, Oregon, USA, September 2010.
- Ehud Ahissar and Eldad Assa. Perception as a closed-loop convergence process. *eLife*, 5 :e12830, May 2016. ISSN 2050-084X. doi:10.7554/eLife.12830. URL <https://doi.org/10.7554/eLife.12830>.
- Dragan Ahmetovic, Cristian Bernareggi, Andrea Gerino, and Sergio Mascetti. Zebra-recognizer : Efficient and precise localization of pedestrian crossings. In *22nd International Conference on Pattern Recognition*, pages 2566–2571, August 2014. doi:10.1109/ICPR.2014.443.
- I. F. Akyildiz, T. Melodia, and K. R. Chowdhury. Wireless Multimedia Sensor Networks : Applications and Testbeds. *Proceedings of the IEEE*, 96(10) :1588–1605, October 2008. doi:10.1109/JPROC.2008.928756.
- I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks : a survey. *Computer Networks*, 38(4) :393–422, 2002. ISSN 1389-1286. doi:10.1016/S1389-1286(01)00302-4. URL <http://www.sciencedirect.com/science/article/pii/S1389128601003024>.
- Éric Alexandre, Hughes de Chaumont, Hogo Hatzig, David Jamoneau, Philippe Lemaire, and Martin Rochon. *La balise sonore en questions*. Délégation ministérielle à l'égalité, ministère de la Transition écologique et solidaire et ministère de la Cohésion des territoires, Arche paroi sud, 92055 La Défense cedex, March 2018.
- Saleh Alghamdi, Ron van Schyndel, and Ibrahim Khalil. Accurate positioning using long range active RFID technology to assist visually impaired people. *Journal of Network and Computer Applications*, 41 :135–147, 2014. ISSN 1084-8045.

- doi:10.1016/j.jnca.2013.10.015. URL <http://www.sciencedirect.com/science/article/pii/S1084804513002580>.
- K. Amarasinghe, K. Kenney, and M. Manic. Toward Explainable Deep Neural Network Based Anomaly Detection. In *11th International Conference on Human System Interaction (HSI)*, pages 311–317, July 2018. doi:10.1109/HSI.2018.8430788.
- AS5040 : 10 Bit 360 °Programmable Magnetic Rotary Encoder*. ams, November 2015. v2-11.
- B. Ando. Electronic sensory systems for the visually impaired. *IEEE Instrumentation Measurement Magazine*, 6(2) :62–67, June 2003. ISSN 1094-6969. doi:10.1109/MIM.2003.1200287.
- Jean Arlat, Yves Crouzet, Yves Deswarte, Jean-Charles Fabre, Jean-Claude Laprie, and David Powell. Tolérance aux fautes. In J. Akoka and I. Comyn-Wattiau, editors, *Encyclopédie de l'informatique et des systèmes d'information*, pages 240–270. Vuibert, Paris, France, 2006. ISBN 27117-4846-4.
- ATmega128RFA1 : 8-bit AVR Microcontroller with Low Power 2.4GHz Transceiver for Zigbee and IEEE 802.15.4*. Atmel, 2014. Rev. 8266F-MCU Wireless-09/14.
- Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things : A survey. *Computer Networks*, 54(15) :2787 – 2805, 2010. ISSN 1389-1286. doi:<https://doi.org/10.1016/j.comnet.2010.05.010>. URL <http://www.sciencedirect.com/science/article/pii/S1389128610001568>.
- Malika Auvray. Remplacer un sens par un autre : la suppléance perceptive. In P. Fuchs, editor, *Le Traité de la Réalité Virtuelle*, volume 1, L'homme et l'environnement virtuel, pages 173–188. Les Presses de l'École des Mines, 3 edition, 2006.
- Malika Auvray, Sylvain Hannequin, Charles Lenay, and Kevin O'Regan. There is something out there : distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, 4(4) :505–21, December 2005. URL <https://hal.archives-ouvertes.fr/hal-00173591>.
- Malika Auvray, Sylvain Hannequin, and J Kevin O'Regan. Learning to perceive with a visuo – auditory substitution system : Localisation and object recognition with 'the voice'. *Perception*, 36(3) :416–430, 2007. doi:10.1068/p5631. PMID : 17455756.
- AEDR-8300 Series Encoders*. Avago Technologies, 2007.
- Algirdas Avizienis, Jean-Claude Laprie, and Brian Randell. Fundamental concepts of dependability. Research report 1145, LAAS-CNRS, April 2001.
- P. Bach-y Rita, K. A. Kaczmarek, M. E. Tyler, and J. Garcia-Lara. Form perception with a 49-point electro-tactile stimulus array on the tongue : a technical note. *Journal of rehabilitation research and development*, 35(4) :427–430, October 1998. ISSN 0748-7711. URL <http://europepmc.org/abstract/MED/10220221>.
- Paul Bach-y Rita. *Brain mechanisms in sensory substitution*. Academic Press Inc, 1972.

- Paul Bach-y-Rita and Stephen W. Kercel. Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7(12) :541–546, 2003. ISSN 1364-6613. doi:10.1016/j.tics.2003.10.013. URL <http://www.sciencedirect.com/science/article/pii/S1364661303002900>.
- Paul Bach-y Rita, Carter C Collins, Frank A Saunders, Benjamin White, and Lawrence Scadden. Vision substitution by tactile image projection. *Nature*, 221(5184) :963, 1969. doi:10.1038/221963a0.
- Bryan Bashin. Truncated-dome threat still looming. *The Braille Monitor*, March 2001. URL <https://nfb.org/images/nfb/publications/bm/bm01/bm0103/bm010303.htm>.
- Russel M. Bauer. The agnosias. In P.J. Snyser, P.D. Nussbaum, and D.L. Robins, editors, *Clinical Neuropsychology : A Pocket Handbook for Assessment*, pages 508–533. American Psychological Association, Washington, DC, 2nd edition, 2006.
- T. Berners-Lee, R. Fielding, and L. Masinter. Uniform Resource Identifier (URI) : Generic Syntax. RFC 3986, IETF, January 2005. URL <https://www.ietf.org/rfc/rfc3987.txt>.
- Shah Bhatti, James Carlson, Hui Dai, Jing Deng, Jeff Rose, Anmol Sheth, Brian Shucker, Charles Gruenwald, Adam Torgerson, and Richard Han. MANTIS OS : An Embedded Multithreaded Operating System for Wireless Micro Sensor Platforms. *Mobile Networks and Applications*, 10(4) :563–579, August 2005. ISSN 1383-469X. URL <http://dl.acm.org/citation.cfm?id=1160162.1160178>.
- David Bolt. Beneficial blindness : literary representation and the so-called positive stereotyping of people with impaired vision. *New Zealand Journal of Disabilities Studies*, 12 :80–100, 2006.
- J. Borenstein and I. Ulrich. The GuideCane—a computerized travel aid for the active guidance of blind pedestrians. In *Proceedings IEEE International Conference on Robotics and Automation*, volume 2, pages 1283–1288 vol.2, Albuquerque, New Mexico, USA, April 1997. doi:10.1109/ROBOT.1997.614314.
- Johann Borenstein and Yoram Koren. Obstacle avoidance with ultrasonic sensors. *IEEE Journal on Robotics and Automation*, 4(2) :21–3–218, April 1988. ISSN 0882-4967. doi:10.1109/56.2085.
- BMP180 : Digital pressure sensor*. Bosch Sensortec, May 2015. Rev. 2.8.
- J. A. Brabyn. New developments in mobility and orientation aids for the blind. *IEEE Transactions on Biomedical Engineering*, BME-29(4) :285–289, April 1982. ISSN 0018-9294. doi:10.1109/TBME.1982.324945.
- Lesley A. Brabyn and John A. Brabyn. An Evaluation of "Talking Signs" for the Blind. *Human Factors*, 25(1) :49–53, 1983. doi:10.1177/001872088302500105. URL <https://doi.org/10.1177/001872088302500105>. PMID : 6840772.
- Thomas Butkiewicz. Low-cost coastal mapping using kinect v2 time-of-flight cameras. In *2014 Oceans - St. John's*, pages 1–9, September 2014. doi:10.1109/OCEANS.2014.7003084.

- S. Caraiman, A. Morar, M. Owczarek, A. Burlacu, D. Rzeszotarski, N. Botezatu, P. Hergelegiu, F. Moldoveanu, P. Strumillo, and A. Moldoveanu. Computer vision for the visually impaired : the sound of vision system. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 1480–1489, October 2017. doi:10.1109/ICCVW.2017.175.
- Matteo Carandini. Charting the Structure of Neuroscience. *Neuron*, 102(4) :732 – 734, 2019. ISSN 0896-6273. doi:<https://doi.org/10.1016/j.neuron.2019.05.012>. URL <http://www.sciencedirect.com/science/article/pii/S0896627319304374>.
- Tsung-Hsiang Chang, Chien-Ju Ho, David Chawei Hsu, Yuan-Hsiang Lee, Min-Shieh Tsai, Mu-Chun Wang, and Jane Hsu. icane – a partner for the visually impaired. In Tomoya Enokido, Lu Yan, Bin Xiao, Daeyoung Kim, Yuanshun Dai, and Lawrence T. Yang, editors, *Embedded and Ubiquitous Computing – EUC 2005 Workshops*, number 3823 in Lecture Notes in Computer Science, pages 393–402. Springer, Berlin, Heidelberg, Germany, December 2005. ISBN 978-3-540-30803-4 978-3-540-32296-2. doi:10.1007/11596042\_41. URL [http://link.springer.com/chapter/10.1007/11596042\\_41](http://link.springer.com/chapter/10.1007/11596042_41).
- J. Choo and S. Liu. Visual Analytics for Explainable Deep Learning. *IEEE Computer Graphics and Applications*, 38(4) :84–92, July 2018. doi:10.1109/MCG.2018.042731661.
- Marion Chottin. Mobility Technologies for Visually Impaired People Through the Prism of Classic Theories of Perception. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People : Fundamentals and ICT Assistive Technologies*, pages 77–108. Springer International Publishing, Cham, Switzerland, 2018. ISBN 978-3-319-54446-5. doi:10.1007/978-3-319-54446-5\_3.
- Alice T. Chuang, Curtis E. Margo, and Paul B. Greenberg. Retinal implants : a systematic review. *British Journal of Ophthalmology*, 98(7) :852–856, 2014. doi:10.1136/bjophthalmol-2013-303708.
- Sakmongkon Chumkamon, Peranitti Tuvaphanthaphiphat, and Phongsak Keeratiwintakorn. A blind navigation system using rfid for indoor environments. In *5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, volume 2, pages 765–768, May 2008. doi:10.1109/ECTICON.2008.4600543.
- James M. Coughlan and Huiying Shen. The Crosswatch Traffic Intersection Analyzer : A Roadmap for the Future. In Klaus Miesenberger, Arthur Karshmer, Petr Penaz, and Wolfgang Zagler, editors, *Computers Helping People with Special Needs*, number 7383 in Lecture Notes in Computer Science, pages 25–28. Springer, Berlin, Heidelberg, Germany, July 2012. ISBN 978-3-642-31533-6 978-3-642-31534-3. doi:10.1007/978-3-642-31534-3\_4. URL [http://link.springer.com/chapter/10.1007/978-3-642-31534-3\\_4](http://link.springer.com/chapter/10.1007/978-3-642-31534-3_4).
- James M. Coughlan and Huiying Shen. Crosswatch : a system for providing guidance to visually impaired travelers at traffic intersections. *Journal of assistive technologies*, 7(2), April 2013. ISSN 1754-9450. doi:10.1108/17549451311328808. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3864896/>.
- Luigi F. Cuturi, Elena Aggius-Vella, Claudio Campus, Alberto Parmiggiani, and Monica Gori. From science to technology : Orientation and mobility in blind children

- and adults. *Neuroscience & Biobehavioral Reviews*, 71 :240–251, 2016. ISSN 0149-7634. doi:10.1016/j.neubiorev.2016.08.019. URL <http://www.sciencedirect.com/science/article/pii/S014976341630210X>.
- Christophe de Vault. *Étude et Développement d'un micro-noyau réparti, temps-reel et tolérant aux fautes : DREAM*. PhD thesis, Université Blaise Pascal, 2003.
- Christophe De Vault and Kun Mean HOU. DREAM : un micro noyau temps réel orienté pour la tolérance aux fautes. In *Télé médecine et e-santé*, volume 13 of *Informatique et Santé*, pages 63–69. Springer-Verlag France, 2002. URL <https://hal-clermont-univ.archives-ouvertes.fr/hal-01788080>.
- Gunnar Declerck and Charles Lenay. Living in space. a phenomenological account. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People*, chapter 1, pages 3–51. Springer International Publishing, Cham, Switzerland, 2018. doi:10.1007/978-3-319-54446-5.
- René Descartes. *Discours de la Méthode*. 1637.
- Diderot. *Lettre sur les aveugles*. Le Livre de Poche Libretti – Librairie Générale Française, 31, rue de Fleurus – 75278 Paris Cedex 06, 4e édition, 2011. (Original publié en 1749 ; présentation et notes par Catherine Bouttier-Couqueberg).
- M. Duerst and M. Suignard. Internationalized Resource Identifiers (IRIs). RFC 3987, IETF, January 2005. URL <https://www.ietf.org/rfc/rfc3987.txt>.
- Cormac Duffy, Utz Roedig, John Herbert, and Cormac J. Sreenan. A Comprehensive Experimental Comparison of Event Driven and Multi-Threaded Sensor Node Operating Systems. *Journal of Networks*, 3(3) :57–70, 2008. URL <https://eprints.lancs.ac.uk/id/eprint/41393>.
- A. Dunkels, B. Gronvall, and T. Voigt. Contiki – a lightweight and flexible operating system for tiny networked sensors. In *29th Annual IEEE International Conference on Local Computer Networks*, pages 455–462, November 2004. doi:10.1109/LCN.2004.38.
- Robert Efron. Biology without consciousness—and its consequences. *Perspectives in Biology and Medicine*, 11(1) :9–36, 1967. doi:10.1353/pbm.1967.0022.
- Robert Efron. What is perception? In Robert S. Cohen and Marx W. Wartofsky, editors, *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*, pages 137–173. Springer Netherlands, Dordrecht, Netherlands, 1969. ISBN 978-94-010-3378-7. doi:10.1007/978-94-010-3378-7\_4. URL [https://doi.org/10.1007/978-94-010-3378-7\\_4](https://doi.org/10.1007/978-94-010-3378-7_4).
- Gaute T. Einevoll, Alain Destexhe, Markus Diesmann, Sonja Grün, Viktor Jirsa, Marc de Kamps, Michele Migliore, Torbjörn V. Ness, Hans E. Plesser, and Felix Schürmann. The Scientific Case for Brain Simulations. *Neuron*, 102(4) :735 – 744, 2019. ISSN 0896-6273. doi:<https://doi.org/10.1016/j.neuron.2019.03.027>. URL <http://www.sciencedirect.com/science/article/pii/S0896627319302909>.
- Riyad A. El-laithy, Jidong Huang, and Michael Yeh. Study on the use of microsoft kinect for robotics applications. In *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, pages 1280–1288, April 2012. doi:10.1109/PLANS.2012.6236985.

- Giulia V. Elli, Stefania Benetti, and Olivier Collignon. Is there a future for sensory substitution outside academic laboratories? *Multisensory Research*, 27(5-6) : 271–291, 2014. doi:10.1163/22134808-00002460. URL <http://booksandjournals.brillonline.com/content/journals/10.1163/22134808-00002460>.
- Michael W. Eysenck. *Principles of Cognitive Psychology*. Principles of Psychology. Psychology Press, 27 Church Road, Hove, East Sussex, BN3 2FA, UK, 2nd edition, 2001.
- René Farcy, Roger Leroux, Alain Jucha, Roland Damaschini, Colette Grégoire, and Aziz Zogaghi. Electronic travel aids and electronic orientation aids for blind people : technical, rehabilitation and everyday life points of view. In M.A. Hersh, editor, *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments*, 2006.
- Fastrax UP501*. Fastrax Ltd, Rev. 1.2 edition, October 2010.
- Gustav Theodor Fechner. *Elemente der Psychophysik*. 1860.
- Vítor Filipe, Filipe Fernandes, Hugo Fernandes, António Sousa, Hugo Paredes, and João Barroso. Blind navigation support system based on microsoft kinect. *Procedia Computer Science*, 14 :94 – 101, 2012. ISSN 1877-0509. doi:<http://dx.doi.org/10.1016/j.procs.2012.10.011>. URL <http://www.sciencedirect.com/science/article/pii/S1877050912007739>. Proceedings of the 4th International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2012).
- Chaz Firestone and Brian J. Scholl. "top-down" effects where none should be found : The el greco fallacy in perception research. *Psychological Science*, 25 (1) :38–46, 2014. doi:10.1177/0956797613485092. URL <https://doi.org/10.1177/0956797613485092>. PMID : 24297777.
- Chaz Firestone and Brian J. Scholl. Seeing and Thinking in Studies of Embodied "Perception" : How (Not) to Integrate Vision Science and Social Psychology. *Perspectives on Psychological Science*, 12(2) :341–343, 2017. doi:10.1177/1745691616679944. URL <https://doi.org/10.1177/1745691616679944>. PMID : 28346111.
- Gerald A. Fishman. When your eyes have a wet nose : the evolution of the use of guide dogs and establishing the seeing eye. *Survey of Ophthalmology*, 48(4) :452–458, 2003. ISSN 0039-6257. doi:[https://doi.org/10.1016/S0039-6257\(03\)00052-3](https://doi.org/10.1016/S0039-6257(03)00052-3). URL <http://www.sciencedirect.com/science/article/pii/S0039625703000523>.
- Seth R. Flaxman, Rupert R. A. Bourne, Serge Resnikoff, Peter Ackland, Tasanee Braithwaite, Maria V. Cicinelli, Aditi Das, Jost B. Jonas, Jill Keeffe, John H. Kempen, Janet Leasher, Hans Limburg, Kovin Naidoo, Konrad Pesudovs, Alex Silvester, Gretchen A. Stevens, Nina Tahhan, Tien Y. Wong, Hugh R. Taylor, Rupert Bourne, Aries Arditi, Yaniv Barkana, Banu Bozkurt, Alain Bron, Donald Budenz, Feng Cai, Robert Casson, Usha Chakravarthy, Jaewan Choi, Maria Vittoria Cicinelli, Nathan Congdon, Reza Dana, Rakhi Dandona, Lalit Dandona, Iva Dekaris, Monte Del Monte, Jenny deya, Laura Dreer, Leon Ellwein, Marcela Frazier, Kevin Frick, David Friedman, Joao Furtado, Hua Gao, Gus Gazzard, Ronnie George, Stephen Gichuhi, Victor Gonzalez, Billy Hammond, Mary Elizabeth Hartnett, Minguang He, James Hejtmancik, Flavio Hirai, John Huang, April Ingram, Jonathan Javitt, Jost Jonas, Charlotte Joslin, John Kempen, Moncef Khairallah, Rohit Khanna, Judy Kim, George Lambrou, Van Charles

- Lansingh, Paolo Lanzetta, Jennifer Lim, Kaweh Mansouri, Anu Mathew, Alan Morse, Beatriz Munoz, David Musch, Vinay Nangia, Maria Palaïou, Maurizio Battaglia Parodi, Fernando Yaacov Pena, Tunde Peto, Harry Quigley, Murugesan Raju, Pradeep Ramulu, Zane Rankin, Dana Reza, Alan Robin, Luca Rossetti, Jinan Saaddine, Mya Sandar, Janet Serle, Tueng Shen, Rajesh Shetty, Pamela Sieving, Juan Carlos Silva, Rita S. Sitorus, Dwight Stambolian, Gretchen Stevens, Hugh Taylor, Jaime Tejedor, James Tielsch, Miltiadis Tsilimbaris, Jan van Meurs, Rohit Varma, Gianni Virgili, Ya Xing Wang, Ning-Li Wang, Sheila West, Peter Wiedemann, Tien Wong, Richard Wormald, and Yingfeng Zheng. Global causes of blindness and distance vision impairment 1990&#x2013;2020 : a systematic review and meta-analysis. *The Lancet Global Health*, 5(12) :e1221–e1234, Dec 2017. ISSN 2214-109X. doi:10.1016/S2214-109X(17)30393-5. URL [https://doi.org/10.1016/S2214-109X\(17\)30393-5](https://doi.org/10.1016/S2214-109X(17)30393-5).
- G. Fortino, A. Guerrieri, and W. Russo. Agent-oriented smart objects development. In *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 907–912, May 2012. doi:10.1109/CSCWD.2012.6221929.
- Allen Foster and Clare Gilbert. Epidemiology of childhood blindness. *Eye*, 6(2) :173–176, 1992. ISSN 1476-5454. doi:10.1038/eye.1992.34. URL <https://doi.org/10.1038/eye.1992.34>.
- Marcel Foucault. La psychophysique. *L'Année psychologique*, 8, 1901.
- Lidar Lite v3 Operation Manual and Technical Specifications*. Garmin, September 2016. Rev. 0A.
- Marcelo C. Ghilardi, Rafael C.O. Macedo, and Isabel H. Manssour. A new approach for automatic detection of tactile paving surfaces in sidewalks. *Procedia Computer Science*, 80 :662–672, 2016. ISSN 1877-0509. doi:10.1016/j.procs.2016.05.356. URL <http://www.sciencedirect.com/science/article/pii/S1877050916307748>. International Conference on Computational Science 2016, ICCS 2016, 6-8 June 2016, San Diego, California, USA.
- James J. Gibson. *The Ecological Approach to Visual Perception*. Psychology Press, New York, Nov. 2014. doi:10.4324/9781315740218.
- Nicholas A. Giudice and Gordon E. Legge. *Blind Navigation and the Role of Technology*, chapter 25, pages 479–500. John Wiley & Sons, Inc., 2008. ISBN 9780470379424. doi:10.1002/9780470379424.ch25. URL <http://dx.doi.org/10.1002/9780470379424.ch25>.
- Reginald G. Golledge, James R. Marston, Jack M. Loomis, and Roberta L. Klatzky. Stated preferences for components of a personal guidance system for nonvisual navigation. *Journal of Visual Impairment & Blindness*, 98(3) :135–147, March 2004. ISSN 0145-482X. URL <http://www.editlib.org/p/97814>.
- Android 9 Compatibility Definition*. Google Inc., February 2019. URL <https://source.android.com/compatibility/cdd>.
- GoSense SAS. Rango, la canne blanche intelligente – GoSense. URL <http://www.gosense.com/fr/rango/>. Accédé le 5 septembre 2019.

- Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. Internet of things (iot) : A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7) :1645–1660, 2013. ISSN 0167-739X. doi:10.1016/j.future.2013.01.010. URL <http://www.sciencedirect.com/science/article/pii/S0167739X13000241>.
- Frederik Gustafsson. Particle filter theory and practice with positioning applications. *IEEE Aerospace and Electronic Systems Magazine*, 25(7) :53–82, July 2010. ISSN 0885-8985. doi:10.1109/MAES.2010.5546308.
- Armin Haller, Krzysztof Janowicz, Simon Cox, Danh Le Phoc, Kerry Taylor, Maxime Lefrançois, Rob Atkinson, Raúl García-Castro, Joshua Lieberman, and Claus Stadler. Semantic Sensor Network Ontology. W3C Recommendation OGC 16-079, W3C, December 2017. URL <https://www.w3.org/TR/vocab-ssn/>.
- M. Hamdi, N. Boudriga, and M. S. Obaidat. Bandwidth-Effective Design of a Satellite-Based Hybrid Wireless Sensor Network for Mobile Target Detection and Tracking. *IEEE Systems Journal*, 2(1) :74–82, March 2008. doi:10.1109/JSYST.2007.916049.
- Chih-Chieh Han, Ram Kumar, Roy Shea, Eddie Kohler, and Mani Srivastava. A Dynamic Operating System for Sensor Nodes. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, pages 163–176, New York, NY, USA, 2005. ACM. ISBN 1-931971-31-5. doi:10.1145/1067170.1067188. URL <http://doi.acm.org/10.1145/1067170.1067188>.
- Handisco. Sherpa par handisco. <https://handisco.com>, 2019. URL <https://handisco.com>. (Accessed on the 25th of July 2019).
- Vanessa Harrar, Sébrina Aubin, Daniel-Robert Chebat, Ron Kupers, and Maurice Ptito. The multisensory blind brain. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People : Fundamentals and ICT Assistive Technologies*, pages 111–136. Springer International Publishing, Cham, Switzerland, 2018. ISBN 978-3-319-54446-5. doi:10.1007/978-3-319-54446-5\_4.
- Faiz M. Hasanuzzaman, Xiaodong Yang, and Yingli Tian. Robust and effective component-based banknote recognition for the blind. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6) :1021–1030, Nov 2012. ISSN 1094-6977. doi:10.1109/TSMCC.2011.2178120.
- L. Hermer-Vazquez, A. Moffet, and P. Munkholm. Language, space, and the development of cognitive flexibility in humans : the case of two spatial memory tasks. *Cognition*, 79(3) :263–299, May 2001.
- Marion A. Hersh. Mobility, inclusion and exclusion. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People : Fundamentals and ICT Assistive Technologies*, pages 631–648. Springer International Publishing, Cham, Switzerland, 2018. doi:10.1007/978-3-319-54446-5\_21.
- Jason Hill, Robert Szewczyk, Alec Woo, Seth Hollar, David Culler, and Kristofer Pister. System Architecture Directions for Networked Sensors. In *Proceedings of the Ninth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS IX, pages 93–104, New York, NY, USA, 2000. ACM. ISBN 1-58113-317-0. doi:10.1145/378993.379006. URL <http://doi.acm.org/10.1145/378993.379006>.

- Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosf, and Mike Dean. SWRL : A Semantic Web Rule Language Combining OWL and RuleML. Technical report, W3C, May 2004. URL <https://www.w3.org/Submission/SWRL/>.
- Brian Hsu, Cheng-Han Hsieh, Sung-Nien Yu, Ehud Ahissar, Amos Arieli, and Yael Zilbershtain-Kra. A tactile vision substitution system for the study of active sensing. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3206–3209, July 2013. doi:10.1109/EMBC.2013.6610223.
- Human Brain Project. Overview. URL <https://www.humanbrainproject.eu/en/about/overview/>. Accessed : 05 july 2019.
- Jorge E. Ibarra-Esquer, Félix F. González-Navarro, Brenda L. Flores-Rios, Larysa Burtseva, and María A. Astorga-Vargas. Tracking the evolution of the internet of things concept across different application domains. *Sensors*, 17(6), 2017. ISSN 1424-8220. doi:10.3390/s17061379. URL <http://www.mdpi.com/1424-8220/17/6/1379>.
- Volodymyr Ivanchenko, James Coughlan, and Huiying Shen. Detecting and locating crosswalks using a camera phone. *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008 :4563143, 2008. ISSN 1063-6919. doi:10.1109/CVPRW.2008.4563143. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2874980/>.
- Volodymyr Ivanchenko, James Coughlan, and Huiying Shen. Staying in the crosswalk : A system for guiding visually impaired pedestrians at traffic intersections. *Assistive technology research series*, 25(2009) :69–73, 2009. ISSN 1383-813X. doi:10.3233/978-1-60750-042-1-69. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2964893/>.
- G. Jansson. The effect of the range of laser cane on the detection of objects by the blind. Technical report, Department of Psychology, University of Uppsala, 1975.
- Gabriele Janzen and Miranda Van Turenout. Selective neural representation of objects relevant for navigation. *Nature neuroscience*, 7(6) :673–677, 2004. doi:10.1038/nn1257.
- Javier Jiménez, Jesús Olea, Jesús Torres, Inmaculada Alonso, Dirk Harder, and Konstanze Fischer. Biography of Louis Braille and Invention of the Braille Alphabet. *Survey of Ophthalmology*, 54(1) :142–149, 2009. ISSN 0039-6257. doi:10.1016/j.survophthal.2008.10.006. URL <http://www.sciencedirect.com/science/article/pii/S0039625708001860>.
- Sebastian Kaebisch, Takuki Kamiya, Michael McCool, and Victor Charpenay. Web of Things (WoT) Thing Description. W3C Candidate Recommendation 16 May 2019, W3C, May 2019a. URL <https://www.w3.org/TR/2019/CR-wot-thing-description-20190516/>.
- Sebastian Kaebisch, Takuki Kamiya, Michael McCool, and Matthias Kovatsch. Web of Things (WoT) Thing Description : Implementation Report . W3c implementation report, W3C, October 2019b. URL <https://w3c.github.io/wot-thing-description/testing/report.html>.

- Hiroiyuki Kajimoto, Masaki Suzuki, and Yonezo Kanno. Hamsatouch : Tactile vision substitution with smartphone and electro-tactile display. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, pages 1273–1278, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2474-8. doi:10.1145/2559206.2581164. URL <http://doi.acm.org/10.1145/2559206.2581164>.
- A. M. Kassim, H. I. Jaafar, M. A. Azam, N. Abas, and T. Yasuno. Design and development of navigation system by using rfid technology. In *IEEE 3rd International Conference on System Engineering and Technology*, pages 258–262, August 2013. doi:10.1109/ICSEngT.2013.6650181.
- Leslie Kay. An ultrasonic sensing probe as a mobility aid for the blind. *Ultrasonics*, 2(2) :53–59, 1964.
- Leslie Kay. A sonar aid to enhance spatial perception of the blind : engineering design and evaluation. *Radio and Electronic Engineer*, 44 :605–627, November 1974. ISSN 0033-7722. doi:10.1049/ree.1974.0148.
- Winthrop N. Kellogg. Sonar system of the blind. *Science*, 137(3528) :399–404, 1962. ISSN 0036-8075. doi:10.1126/science.137.3528.399. URL <https://science.sciencemag.org/content/137/3528/399>.
- R. Kleihorst, B. Schueler, A. Danilin, and M. Heijligers. Smart camera mote with high performance vision system. In *Workshop on distributed smart cameras (DSC)*, Boulder, Colorado, USA, October 2006. URL [https://pervasive.aau.at/BR/dsc06/CR/dsc06\\_p11\\_cr.pdf](https://pervasive.aau.at/BR/dsc06/CR/dsc06_p11_cr.pdf).
- Kevin Klues, Chieh-Jan Mike Liang, Jeongyeup Paek, Răzvan Musăloiu-E, Philip Levis, Andreas Terzis, and Ramesh Govindan. TOSThreads : Thread-safe and Non-invasive Preemption in TinyOS. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, SenSys '09, pages 127–140, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-519-2. doi:10.1145/1644038.1644052. URL <http://doi.acm.org/10.1145/1644038.1644052>.
- Kurt Konolige, Joseph Augenbraun, Nick Donaldson, Charles Fiebig, and Pankaj Shah. A low-cost laser distance sensor. In *IEEE International Conference on Robotics and Automation*, pages 3002–3008, May 2008. doi:10.1109/ROBOT.2008.4543666.
- Matthias Kovatsch, Ryuichi Matsukura, Michael Lagally, Toru Kawaguchi, Kuni-hiko Toumura, and Kazuo Kajimoto. Web of Things (WoT) Architecture. W3c candidate recommendation, W3C, May 2019. URL <https://www.w3.org/TR/wot-architecture/>.
- C. Ronald Kube. A minimal infrared obstacle detection scheme. *The Robotics Practitioner*, 2(2) :15–20, 1996. URL [http://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/motionplanning/papers/sbp\\_papers/integrated1/kube\\_infrared\\_obs\\_detect.pdf](http://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/motionplanning/papers/sbp_papers/integrated1/kube_infrared_obs_detect.pdf).
- V. Kulyukin, C. Gharpure, J. Nicholson, and S. Pavithran. Rfid in robot-assisted indoor navigation for the visually impaired. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1979–1984 vol.2, September 2004. doi:10.1109/IROS.2004.1389688.

- Vladimir Kulyukin, Chaitanya Gharpure, John Nicholson, and Grayson Osborne. Robot-assisted wayfinding for the visually impaired in structured indoor environments. *Autonomous Robots*, 21(1) :29–41, June 2006. ISSN 0929-5593, 1573-7527. doi:10.1007/s10514-006-7223-8. URL <http://link.springer.com/article/10.1007/s10514-006-7223-8>.
- Jeremy Landt. The history of RFID. *IEEE Potentials*, 24(4) :8–11, October 2005. ISSN 0278-6648. doi:10.1109/MP.2005.1549751.
- Charles Lenay and Gunnar Declerck. Technologies to access space without vision. some empirical facts and guiding theoretical principles. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People : Fundamentals and ICT Assistive Technologies*, chapter 2, pages 53–75. Springer International Publishing, Cham, Switzerland, 2018. ISBN 978-3-319-54446-5. doi:10.1007/978-3-319-54446-5\_2.
- Charles Lenay, Olivier Gapenne, Sylvain Hanneton, Catherine Marque, and Christelle Genou elle. Sensory substitution : Limits and perspectives. In *Advances in consciousness research*, pages 276–292. John Benjamins Publishing Company, Amsterdam, Netherlands, 2003. doi:10.1075/aicr.53.22len.
- Susan Lewallen and Paul Courtright. Blindness in Africa : present situation and future needs. *British Journal of Ophthalmology*, 85(8) :897–903, 2001. ISSN 0007-1161. doi:10.1136/bjo.85.8.897. URL <https://bjo.bmj.com/content/85/8/897>.
- Xiongmin Lin, Jean Connier, Philippe Vaslin, Chengcheng Guo, Tianyi Zang, Jian-Jin Li, Christophe de Vault, and Kun Mean Hou. Indoor navigation with the smart environment explorer stick (SEES). In *New and Smart Information and Communication Science and Technology to Support Sustainable Development 2015 (NICST 2015)*. Institut du Mat eriau dans le Syst eme (IMS), Bordeaux, France, 2015.
- Xing Liu, Kun Mean Hou, Christophe De Vault, Hongling Shi, and Khalid El Gholami. MIROS : A Hybrid Real-Time Energy-Efficient Operating System for the Resource-Constrained Wireless Sensor Nodes. *Sensors*, 14(9) :17621–17654, 2014. ISSN 1424-8220. doi:10.3390/s140917621. URL <https://www.mdpi.com/1424-8220/14/9/17621>.
- J. M. Loomis, R. G. Golledge, R. L. Klatzky, J. M. Speigle, and J. Tietz. Personal guidance system for the visually impaired. In *Proceedings of the First Annual ACM Conference on Assistive Technologies*, Assets '94, pages 85–91, New York, NY, USA, 1994. ACM. ISBN 0-89791-649-2. doi:10.1145/191028.191051.
- Jack M Loomis, Roberta L Klatzky, and Nicholas A Giudice. Sensory substitution of vision : Importance of perceptual and cognitive processing. In Roberto Manduchi and Sri Kurniawan, editors, *Assistive technology for blindness and low vision*. CRC Press, Boca Raton, FL, USA, 2012.
- W. Loughborough. Talking lights. *Journal of Visual Impairment and Blindness*, 73 :243, 1979.
- Jiangyan Lu, Kin Wai Michael Siu, and Ping Xu. A comparative study of tactile paving design standards in different countries. In *9th International Conference on Computer-Aided Industrial Design and Conceptual Design*, pages 753–758, Nov 2008. doi:10.1109/CAIDCD.2008.4730674.

- Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. A Realistic Evaluation and Comparison of Indoor Location Technologies : Experiences and Lessons Learned. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, IPSN '15, pages 178–189, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3475-4. doi:10.1145/2737095.2737726. URL <http://doi.acm.org/10.1145/2737095.2737726>.
- Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, BGalit Buchs, Daniel-Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. The "eyecane", a new electronic travel aid for the blind : Technology, behaviour & swift learning. *Restorative Neurology and Neuroscience*, 32(6) :813–824, 2014. doi:10.3233/RNN-130351.
- Roberto Manduchi and Sri Kurniawan. Mobility-Related Accidents Experienced by People with Visual Impairment. *Insight : Research and Practice in Visual Impairment and Blindness*, 4(2), 2011. URL <https://users.soe.ucsc.edu/~manduchi/papers/MobilityAccidents.pdf>.
- David Marr. *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman and Company, New York, NY, USA, 1982.
- LV-MaxSonar<sup>®</sup>-EZ<sup>™</sup>Series*. MaxBotix Inc., 2015. Rev. 3.
- Peter B. L. Meijer. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39(2) :112–121, Feb 1992. ISSN 0018-9294. doi:10.1109/10.121642.
- Peter B.L. Meijer. The voice - new frontiers in sensory substitution. Web, 2019. URL <https://www.seeingwithsound.com/>.
- Maurice Merleau-Ponty. *Le primat de la perception*. Editions Verdier, 11220 Lagrasse, 2014.
- Microsoft. Seeing ai. Website, 2017. URL <https://www.microsoft.com/en-us/seeing-ai/>. Accessed : 29 July 2019.
- MIPSoft. Blindsquare. Website, 2019. URL <https://www.blindsquare.com/>.
- Kei Mizuno, Masaaki Tanaka, Kouzi Yamaguti, Osami Kajimoto, Hirohiko Kuratsune, and Yasuyoshi Watanabe. Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. *Behavioral and Brain Functions*, 7(1) :17, 2011. ISSN 1744-9081. doi:10.1186/1744-9081-7-17. URL <https://doi.org/10.1186/1744-9081-7-17>.
- Hideo Mori and Shinji Kotani. Robotic travel aid for the blind : HARUNOBU-6. In P. Sharkey, F. D. Rose, and J.-I. Lindström, editors, *Proceedings of the 2nd European conference on disability, virtual reality and associated technologies (ECDVRAT)*, pages 193–202. The University of Reading, 1998. ISBN 0704911418. URL <http://centaur.reading.ac.uk/27455/>.
- A. Munir, A. Gordon-Ross, and S. Ranka. Multi-core embedded wireless sensor networks : Architecture and applications. *IEEE Transactions on Parallel and Distributed Systems*, 25(6) :1553–1562, June 2014. doi:10.1109/TPDS.2013.219.

- Arslan Munir, Ann Gordon-Ross, and Sanjay Ranka. *Modeling and optimization of parallel and distributed embedded systems*. Wiley – IEEE Press, Hoboken, NJ, December 2015. ISBN 978-1-119-08639-0.
- Serge Nicolas. La fondation de la psychophysique de Fechner : des présupposés métaphysiques aux écrits scientifiques de weber. *L'Année psychologique*, 102(2) :255–298, 2002. doi:10.3406/psy.2002.29592.
- LM75B : Digital temperature sensor and thermal watchdog*. NXP B.V., February 2015. Rev. 6.1.
- Martin O'Connor and Amar Das. Sqwrl : A query language for owl. In *Proceedings of the 6th International Conference on OWL : Experiences and Directions - Volume 529*, OWLED'09, pages 208–215, Aachen, Germany, Germany, 2009. CEUR-WS.org. URL <http://dl.acm.org/citation.cfm?id=2890046.2890072>.
- Adrian M. Owen. The Search for Consciousness. *Neuron*, 102(3) :526 – 528, 2019. ISSN 0896-6273. doi:<https://doi.org/10.1016/j.neuron.2019.03.024>. URL <http://www.sciencedirect.com/science/article/pii/S0896627319302879>.
- Fred Paas, Alexander Renkl, and John Sweller. Cognitive Load Theory and Instructional Design : Recent Developments. *Educational Psychologist*, 38(1) :1–4, 2003. doi:10.1207/S15326985EP3801\_1. URL [https://doi.org/10.1207/S15326985EP3801\\_1](https://doi.org/10.1207/S15326985EP3801_1).
- E. Pissaloux, R. Velazquez, M. Hersh, and G. Uzan. Towards a cognitive model of human mobility : an investigation of tactile perception for use in mobility device. *Journal of Navigation*, 70(1) :1–17, January 2017. doi:10.1017/S0373463316000461. URL <http://eprints.gla.ac.uk/132599/>.
- Shelly Powers. *Practical RDF*. O'Reilly Media, 1005 Gravenstein Highway North, Sebastopol, CA, USA, 1 edition, November 2003. ISBN 0-596-00263-7.
- N. Pressey. Mowat sensor. *Focus*, 11(3) :35–39, 1977.
- Zenon Pylyshyn. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioural and Brain Sciences*, 22(3) :341–365, Jun 1999.
- CL Reed, RJ Caselli, and MJ Farah. Tactile agnosia. underlying impairment and implications for normal tactile object recognition. *Brain : a journal of neurology*, 119 (Pt 3) :875–888, June 1996. ISSN 0006-8950. doi:10.1093/brain/119.3.875. URL <https://doi.org/10.1093/brain/119.3.875>.
- GDP Research. The miniguide mobility aid. Website, 2005. URL [http://www.gdp-research.com.au/minig\\_1.htm](http://www.gdp-research.com.au/minig_1.htm).
- Irvin Rock and Stephen Palmer. The Legacy of Gestalt Psychology. *Scientific American*, 263(6) :84–91, 1990. ISSN 00368733, 19467087. URL <http://www.jstor.org/stable/24997014>.
- Uta R. Roentgen, Gert Jan Gelderblom, Mathijs Soede, and Luc P. de Witte. Inventory of electronic mobility aids for persons with visual impairments : A

- literature review. *Journal of Visual Impairment & Blindness*, 102(11) :702–724, 2008. doi:10.1177/0145482X0810201105. URL <https://doi.org/10.1177/0145482X0810201105>.
- Uta R. Roentgen, Gert Jan Gelderblom, Mathijs Soede, and Luc P. De Witte. The impact of electronic mobility devices for persons who are visually impaired : A systematic review of effects and effectiveness. *Journal of Visual Impairment & Blindness*, 103(11) :743–753, 2009. doi:10.1177/0145482X0910301104.
- David A. Ross and Bruce B. Blasch. Development of a wearable computer orientation system. *Personal and Ubiquitous Computing*, 6(1) :49–63, January 2002. ISSN 1617-4909. doi:10.1007/s007790200005.
- L Russel. Travel path sounder. In *Proceedings of the Rotterdam Mobility Research Conference*, New York, 1965. American Foundation for the Blind.
- Rikuo Sakaguchi, Shino Takasu, and Tetsuo Akiyama. Study concerning the colours of tactile blocks for the visually handicapped – visibility for the visually handicapped and scenic congruence for those with ordinary sight and vision. In *JIPEA World Congress*, pages 453–462, 2000. URL <http://www.sept.org/techpapers/1148.pdf>.
- Dona Sauerburger. O&M Living History – Where Did Our O&M Techniques Come From? *Newsletter of the Metropolitan Washington Orientation and Mobility Association*, May 1996. URL <http://www.sauerburger.org/dona/omhistory.htm>.
- Daniel L. Schacter, Daniel T. Gilbert, Daniel M. Wegner, and Matthew K. Nock. *Psychology*. Worth Publishers, New York, NY, USA, 3rd edition, 2014.
- Jasper Schipperijn, Jacqueline Kerr, Scott Duncan, Thomas Madsen, Charlotte Klinker, and Jens Troelsen. Dynamic accuracy of gps receivers for use in health research : A novel method to assess gps accuracy in real-world settings. *Frontiers in Public Health*, 2 :21, March 2014. ISSN 2296-2565. doi:10.3389/fpubh.2014.00021.
- Albrecht Schmidt, Michael Beigl, and Hans-W Gellersen. There is more to context than location. *Computers & Graphics*, 23(6) :893–901, 1999. ISSN 0097-8493. doi:[https://doi.org/10.1016/S0097-8493\(99\)00120-X](https://doi.org/10.1016/S0097-8493(99)00120-X). URL <http://www.sciencedirect.com/science/article/pii/S009784939900120X>.
- Simone Schnall. No magic bullet in sight : A reply to firestone and scholl (2017) and durgin (2017). *Perspectives on Psychological Science*, 12(2) :347–349, 2017. doi:10.1177/1745691617691948. URL <https://doi.org/10.1177/1745691617691948>. PMID : 28346116.
- Charles E Schroeder, Donald A Wilson, Thomas Radman, Helen Scharfman, and Peter Lakatos. Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, 20(2) :172–176, 2010. ISSN 0959-4388. doi:<https://doi.org/10.1016/j.conb.2010.02.010>. URL <http://www.sciencedirect.com/science/article/pii/S0959438810000322>.
- Sendero Group. Sendero group : The seeing eye gps app for cell-enabled ios devices. Web, 2019. URL <https://www.senderogroup.com/products/shopseeingeyegps.html>.

- M. Serrão, J.M.F. Rodrigues, J.I. Rodrigues, and J.M.H. du Buf. Indoor Localization and Navigation for Blind Persons using Visual Landmarks and a GIS. *Procedia Computer Science*, 14 :65–73, 2012. ISSN 1877-0509. doi:10.1016/j.procs.2012.10.008. URL <http://www.sciencedirect.com/science/article/pii/S1877050912007703>. Proceedings of the 4th International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2012).
- GP2Y0A02YK : Long Distance Measuring Sensor*. Sharp.
- Huiying Shen and James M. Coughlan. Towards a real-time system for finding and reading signs for visually impaired users. In Klaus Miesenberger, Arthur Karshmer, Petr Penaz, and Wolfgang Zagler, editors, *Computers Helping People with Special Needs*, number 7383 in Lecture Notes in Computer Science, pages 41–47. Springer, Berlin, Heidelberg, Germany, July 2012. ISBN 978-3-642-31533-6 978-3-642-31534-3. doi:10.1007/978-3-642-31534-3\_7. URL [http://link.springer.com/chapter/10.1007/978-3-642-31534-3\\_7](http://link.springer.com/chapter/10.1007/978-3-642-31534-3_7).
- Hongling Shi, Kun Mean Hou, Haiying Zhou, and Xing Liu. Energy efficient and fault tolerant multicore wireless sensor network : E<sup>2</sup>MWSN. In *7th International conference on wireless communications, networking and mobile computing (WiCOM)*, pages 1–4, 2011.
- Hongling Shi, Kun Mean Hou, Xunxing Diao, Xing Liu, and Christophe de Vault. A robust multi-core multi-support and modular wireless multimedia sensor network : MiLive. In *ECOTECHS*, Montoldre, France, October 2013.
- Y. Shiizu, Y. Hirahara, K. Yanashima, and K. Magatani. The development of a white cane which navigates the visually impaired. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5005–5008, August 2007. doi:10.1109/IEMBS.2007.4353464.
- GP Soong, JE Lovie-Kitchin, and B Brown. Does mobility performance of visually impaired adults improve immediately after orientation and mobility training? *Optometry and vision science : official publication of the American Academy of Optometry*, 78(9) : 657–666, September 2001. ISSN 1040-5488. doi:10.1097/00006324-200109000-00011. URL <https://doi.org/10.1097/00006324-200109000-00011>.
- Sound of Vision Consortium. Final report. Technical report, Sound of Vision Consortium, 2018. URL <https://soundofvision.net/download/780/>.
- LSM330DLC*. ST Microelectronics, 2012. Rev. 2.
- LSM303DLHC*. ST Microelectronics, 2013. Rev. 2.
- VL53L0X : World smallest Time-of-Flight ranging and gesture detection sensor*. ST Microelectronics, 2016. Rev. 1.
- Graham Stokes. *On Being Old : The Psychology Of Later Life*. Routledge, Taylor & Francis, London, 1992.
- P. Strumillo, M. Bujacz, P. Baranski, P. Skulimowski, P. Korbek, M. Owczarek, K. Tomalczyk, A. Moldoveanu, and R. Unnthorsson. Different approaches to aiding blind persons in mobility and navigation in the “naviton” and “sound of vision” projects. In Edwige Pissaloux and Ramiro Velazquez, editors, *Mobility of Visually Impaired People :*

- Fundamentals and ICT Assistive Technologies*, pages 435–468. Springer International Publishing, Cham, 2018. ISBN 978-3-319-54446-5. doi:10.1007/978-3-319-54446-5\_15.
- Ruxandra Tapu, Bogdan Mocanu, Andrei Bursuc, and Zaharia Titus. A smartphone-based obstacle detection and classification system for assisting visually impaired people. In *2013 IEEE International conference on computer vision workshop*, pages 444–451, 2013. doi:10.1109/ICCVW.2013.65.
- Christoph Teufel and Bence Nanay. How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, 47 :17 – 25, 2017. ISSN 1053-8100. doi:https://doi.org/10.1016/j.concog.2016.05.008. URL <http://www.sciencedirect.com/science/article/pii/S1053810016300927>. Cognitive Penetration and Predictive Coding.
- Yingli Tian and Shuai Yuan. Clothes Matching for Blind and Color Blind People. In Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer, editors, *Computers Helping People with Special Needs*, pages 324–331, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-14100-3. doi:10.1007/978-3-642-14100-3\_48.
- Yingli Tian, Xiaodong Yang, and Aries Arditi. Computer vision-based door detection for accessibility of unfamiliar environments to blind persons. In Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer, editors, *Computers Helping People with Special Needs*, number 6180 in Lecture Notes in Computer Science, pages 263–270. Springer, Berlin, Heidelberg, Germany, July 2010. ISBN 978-3-642-14099-0 978-3-642-14100-3. doi:10.1007/978-3-642-14100-3\_39. DOI : 10.1007/978-3-642-14100-3\_39.
- E. C. Tolman. Cognitive maps in rats and men. *The Psychological Review*, 1948.
- Barbara Tversky. Distortions in cognitive maps. *Geoforum*, 23 :131–138, 1992.
- Barbara Tversky. Functional significance of visuospatial representations. In Priti Shah and Akira Miyake, editors, *The Cambridge Handbook of Visuospatial Thinking*, Cambridge Handbooks in Psychology, pages 1–34. Cambridge University Press, 2005. doi:10.1017/CBO9780511610448.002.
- Claude Vincent, François Routhier, Valérie Martel, Marie-Ève Mottard, Frédéric Dumont, Lise Côté, and Danielle Cloutier. Field testing of two electronic mobility aid devices for persons who are deaf-blind. *Disability and Rehabilitation : Assistive Technology*, 9(5) :414–420, 2014. doi:10.3109/17483107.2013.825929.
- TEPT4400 : Ambient Light Sensor*. Vishay, August 2014. Rev. 1.7.
- W3C OWL Working Group. OWL 2 web ontology language document overview (second edition). Technical report, W3C, December 2012. URL <http://www.w3.org/TR/2012/REC-owl2-overview-20121211/>.
- Ranxiao Frances Wang and Elizabeth S Spelke. Human spatial representation : insights from animals. *Trends in Cognitive Sciences*, 6(9) :376–382, 2002. ISSN 1364-6613. doi:https://doi.org/10.1016/S1364-6613(02)01961-7. URL <http://www.sciencedirect.com/science/article/pii/S1364661302019617>.

- Shuihua Wang and YingLi Tian. Indoor signage detection based on saliency map and bipartite graph matching. In *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, pages 518–525, November 2011. doi:10.1109/BIBMW.2011.6112422.
- Scooter Willis and Sumi Helal. Rfid information grid for blind navigation and wayfinding. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pages 34–37, October 2005. doi:10.1109/ISWC.2005.46.
- Kathleen E. Wirth and David B. Rein. The economic costs and benefits of dog guides for the blind. *Ophthalmic Epidemiology*, 15(2) :92–98, 2008. doi:10.1080/09286580801939353.
- David Wood, Markus Lanthaler, and Richard Cyganiak. RDF 1.1 concepts and abstract syntax. W3C recommendation, W3C, February 2014. URL <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>.
- World Health Organisation. *International statistical classification of diseases and related health problems (11th Revision)*. World Health Organisation, Geneva, Switzerland, 2018. [Online : <https://icd.who.int/browse11/l-m/en>].
- World Health Organisation. Priority eye diseases. Webpage, World Health Organisation, Geneva, Switzerland, 2019. [Online, accessed on the 21st of June 2019 : <https://www.who.int/blindness/causes/priority/en/index1.html>].
- Huadong Wu. *Sensor Data Fusion for Context-aware Computing Using Dempster-shafer Theory*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2004. AAI3126933.
- Chucui Yi and Yingli Tian. Assistive Text Reading from Natural Scene for Blind Persons. In Gang Hua and Xian-Sheng Hua, editors, *Mobile Cloud Visual Media Computing*, pages 219–241. Springer International Publishing, 2015. ISBN 978-3-319-24700-7. doi:10.1007/978-3-319-24702-1\_9.
- Chucui Yi, Roberto W. Flores, Ricardo Chinchá, and YingLi Tian. Finding objects for assisting blind people. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 2(2) :71–79, February 2013. ISSN 2192-6662, 2192-6670. doi:10.1007/s13721-013-0026-x.
- Moustafa Youssef. Towards Truly Ubiquitous Indoor Localization on a Worldwide Scale. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '15, pages 12 :1–12 :4, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3967-4. doi:10.1145/2820783.2820883.
- M. Yusro, K.M. Hou, E. Pissaloux, H.L. Shi, K. Ramli, and D. Sudiana. Sees : Concept and design of a smart environment explorer stick. In *2013 The 6th International Conference on Human System Interaction (HSI)*, pages 70–77, June 2013a. doi:10.1109/HSI.2013.6577804.
- Muhammad Yusro. *Development of a new algorithm for improving accuracy of pole detection to the supporting system of mobility aid for visually impaired person*. PhD thesis, Universitas Indonesia et Université Clermont Auvergne, June 2017. URL <https://hal.archives-ouvertes.fr/tel-02004329>. HAL Id : tel-02004329, version 1.

- Muhammad Yusro, Kun Mean Hou, Edwige Pissaloux, Kalamullah Ramli, Dodi Sudiana, Zhang Lizhong, and Hong Ling Shi. Design and Implementation of SEE-Phone in SEES (Smart Environment Explorer Stick). In *Proc. Inter. Workshop of New Information Communication, Science and Technology for Sustainable Development (NICST 2013)*, 2013b.
- Xianchao Zheng, Hui Yang, Weiwei Tang, Shuxiang Pu, Lingxiang Zheng, Huiru Zheng, Bruce Liao, and Jolly Wang. Indoor Pedestrian Navigation with Shoe-Mounted Inertial Sensors. In James J. (Jong Hyuk) Park, Shu-Ching Chen, Joon-Min Gil, and Neil Y. Yen, editors, *Multimedia and Ubiquitous Engineering*, number 308 in Lecture Notes in Electrical Engineering, pages 67–73. Springer, Berlin, Heidelberg, Germany, 2014. ISBN 978-3-642-54899-4 978-3-642-54900-7. doi:10.1007/978-3-642-54900-7\_10. DOI : 10.1007/978-3-642-54900-7\_10.
- Haiying Zhou, Kun Mean Hou, and Christophe de Vault. SDREAM : A Super-Small Distributed Real-time Microkernel Dedicated to Wireless Sensors. *International Journal of Pervasive Computing and Communications*, 2(4) :398–410, 2006.
- Julia Zucker and Eric Alexandre. *Le chien guide d'aveugle ou le chien d'assistance*. Délégation ministérielle à l'égalité, ministère de la Transition écologique et solidaire et ministère de la Cohésion des territoires, Arche Sud, 92055 La Défense cedex, February 2018. URL <https://www.ecologique-solidaire.gouv.fr/sites/default/files/DMA%20-%20Le%20chien%20guide%20d%27aveugle%20ou%20le%20chien%20d%27assistance.pdf>.