



**HAL**  
open science

# Perception de scène par un système multi-capteurs, application à la navigation dans des environnements d'intérieur structuré

Marwa Chakroun

► **To cite this version:**

Marwa Chakroun. Perception de scène par un système multi-capteurs, application à la navigation dans des environnements d'intérieur structuré. Modélisation et simulation. Ecole nationale supérieure Mines-Télécom Atlantique; École nationale d'ingénieurs de Sfax (Tunisie), 2021. Français. NNT : 2021IMTA0278 . tel-03519333

**HAL Id: tel-03519333**

**<https://theses.hal.science/tel-03519333v1>**

Submitted on 10 Jan 2022

**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.

# THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE  
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE  
L'ÉCOLE NATIONALE D'INGENIEURS DE SFAX-UNIVERSITÉ DE  
SFAX

ECOLE DOCTORALE N° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Signal, Image, Vision*

ECOLE DOCTORALE SCIENCES ET TECHNOLOGIES  
Spécialité : *Génie électrique*

Par

**Marwa CHAKROUN**

## **Perception de scène par un système multi-capteurs, application à la navigation dans des environnements intérieurs structurés**

Thèse présentée et soutenue à ENIS, le 06 décembre  
Unités de recherche : iti et CEM\_Lab  
Thèse N° : 2021IMTA0278

### **Rapporteurs avant soutenance :**

Najoua BEN AMARA ESSOUKRI  
Olfa KANOUN

Professeur, ENISO  
Professeur, Technische Universität Chemnitz

### **Composition du Jury :**

Président : Ali KHENCHAF  
Examineurs : Abdennaceur KACHOURI  
Najoua BEN AMARA ESSOUKRI  
Olfa KANOUN  
Imene KHANFIR KALLEL  
Sonda AMMAR BOUHAMED

Professeur, ENSTA/Bretagne  
Professeur, ENIS  
Professeur, ENISO  
Professeur, Technische Universität Chemnitz  
Maitre de conférence, ISBS  
Maitre-assistant, ENET'Com

Dir. de thèse : Basel SOLAIMAN  
Co-encadrante : Houda DERBEL

Professeur, IMT Atlantique  
Maître de conférences, HDR, FSS

**Invité (s)**  
Jean-Paul HATON

Professeur, Université Henri-Poincaré



## Remerciements

*Les travaux présentés dans cette thèse sont le résultat de la recherche au sein du laboratoire de recherche "Control and Energy Management Laboratory (CEM\_Lab)" avec l'équipe 'Cybernetics Team' à l'Ecole Nationale d'Ingénieurs de Sfax en Tunisie et au sein du département ITI (Images et Traitement de l'Information) à l'IMT Atlantique en France.*

*Je remercie vivement Prof. **Ali KHENCHAF**, professeur à ENSTA Bretagne, pour l'honneur qu'il m'a accordé en acceptant de présider mon jury de thèse de doctorat.*

*Au terme de cette riche expérience, je voudrais exprimer ma haute reconnaissance à mon directeur de thèse Prof. **Basel SOLAIMAN**, chef du département ITI (Images et Traitement de l'Information) à Télécom Bretagne et Mme **Houda BEN JEMAA DERBEL**, maître de conférences à Faculté de Science de Sfax et , pour la qualité de leur encadrement, leur disponibilité, leurs compétences scientifiques et leurs hauts caractères personnels que j'ai énormément appréciés tout au long de l'élaboration de cette thèse.*

*Mes remerciements s'adressent aussi à Prof. **Najoua ESSOUKRI BEN AMARA**, professeur à ENISO, et Prof. **Olfa KANOUN**, professeur à Technische Universität Chemnitz, qui ont bien accepté l'examen de mes travaux de recherche et pour avoir accepté d'être rapporteurs de ce travail.*

*Mes remerciements s'adressent également à Prof. **Jean-Paul HATON**, professeur à l'Université Henri-Poincaré, pour l'honneur qu'il m'a fait en acceptant de faire partie du jury de cette thèse en tant qu'examineur de mes travaux de recherche.*

*J'adresse un remerciement particulier à Mme **Imene KHANFIR KALLEL** et Mme **Sonda AMMAR BOUHAMED** pour leurs conseils, leurs efforts fournis, leur confiance et pour les moments de discussion dont je ressortais toujours les idées plus claires.*

*Je remercie vivement Ing. **Zied LOUKIL** pour sa disponibilité et pour tous ses efforts fournis pour maîtriser le prototype d'acquisition utilisé au long des travaux de ma thèse.*

*Mes remerciements les plus chaleureuses à tous mes collègues et mes amies du l'équipe 'Cybernetics Team', du laboratoire CEM\_Lab et à toute l'équipe du département ITI pour leur soutien moral et leur solidarité dont elles ont toujours fait preuve.*

*Enfin, je voudrais exprimer ma haute reconnaissance à mes enseignants, qui ont contribué à ma formation et mes connaissances.*

### **Dédicaces**

*Toutes les lettres ne sauraient trouver les mots qu'il faut. . .*

*Tous les mots ne sauraient exprimer la gratitude, l'amour, le respect, la reconnaissance. . .*

**À mes chers parents Houria et Zouhair**, source de vie, d'affection et d'amour.

**À mon cher frère**, source de patience et de courage

**À toute ma famille**, source de motivation et d'espoir

**À tous mes amis**, source de bonheur et de joie

**À vous cher lecteur**

# Table des matières

<b>Table des figures</b>	<b>ix</b>
<b>Liste des tableaux</b>	<b>xv</b>
<b>I Cadre applicatif et théorique de l'étude</b>	<b>6</b>
I.1 Cadre applicatif . . . . .	7
I.1.1 Systèmes de localisation et de cartographie simultanées . . . . .	7
I.1.2 Les techniques de localisation . . . . .	9
I.1.2.1 Techniques basées sur les propriétés du signal . . . . .	9
I.1.2.2 Les Algorithmes de localisation . . . . .	10
I.1.3 Sources d'informations . . . . .	10
I.1.3.1 Capteurs proprioceptifs . . . . .	11
I.1.3.2 Capteurs extéroceptifs . . . . .	12
I.1.4 Modélisation d'un environnement intérieur . . . . .	17
I.1.4.1 Grille d'occupation . . . . .	17
I.1.4.2 Modélisation métrique . . . . .	18
I.1.4.3 Modélisation topologique . . . . .	18
I.1.4.4 Modélisation hybride . . . . .	18
I.1.4.5 Modèle de primitives . . . . .	19
I.2 Proposition d'un modèle générique d'interprétation de scène . . . . .	19
I.3 Cadre théorique pour le développement du système d'interprétation proposé . . . . .	23
I.3.1 Concepts fondamentaux . . . . .	23
I.3.1.1 Donnée, Information, Connaissance . . . . .	23
I.3.1.2 Les imperfections dans l'information . . . . .	25

I.3.2	Choix du cadre théorique . . . . .	25
I.3.3	La théorie de possibilités . . . . .	26
I.3.3.1	Distribution de possibilités . . . . .	27
I.3.3.2	Mesures de possibilité et de nécessité . . . . .	28
I.3.3.3	Mesures d'incertitude . . . . .	29
I.3.3.4	Mesures de similarité . . . . .	29
I.3.3.5	Fusion d'informations possibilistes . . . . .	31
I.3.3.6	Critères de décision dans un cadre possibiliste . . . . .	32
I.4	Conclusion . . . . .	34

## **II Développement d'outils de traitement possibilistes des données acquises pour la conception du système de navigation** **35**

II.1	Proposition d'approches pour la sélection d'attributs dans un cadre possibiliste .	37
II.1.1	App 1 : Estimation du pouvoir discriminant par la fusion de mesures statistiques et mesures d'incertitude (DP_SUMF : Discriminant power based on statistics and uncertainty measure fusion) . . . . .	37
II.1.2	App 2 : Estimation du pouvoir discriminant en se basant sur des mesures possibilistes de spécificité et de conflit (DP_PMSC : Discriminant power based on possibilistic measures of specificity and conflict) . . . . .	40
II.1.3	App 3 : Estimation du pouvoir discriminant en se basant sur des mesures possibilistes de consensus et de conflit (DP_PMCC : Discriminant power based on possibilistic measures of consensus and conflict) . . . . .	42
II.1.4	Résultats et discussions . . . . .	44
II.2	Proposition de classifieurs binaires hiérarchiques . . . . .	49
II.2.1	État de l'art . . . . .	49
II.2.2	Contexte de l'étude et motivation . . . . .	50
II.2.3	Proposition d'un classifieur binaire hiérarchique monocritère "CBH(1)" .	51
II.2.3.1	Principe . . . . .	51
II.2.3.2	Architecture du CBH(1) . . . . .	52
II.2.3.3	Phase de classification . . . . .	53
II.2.4	Proposition d'un classifieur binaire hiérarchique multicritère "CBH(2)" . .	55
II.2.4.1	Principe . . . . .	55

II.2.4.2	Architecture du CBH(2) . . . . .	56
II.2.4.3	Phase de classification . . . . .	57
II.3	Évaluation du CBH(1) et CBH(2) . . . . .	58
II.3.1	Classifieurs utilisés . . . . .	58
II.3.2	Bases de données utilisées . . . . .	59
II.3.3	Logiciel utilisé . . . . .	60
II.3.4	Mesures statistiques de pertinence . . . . .	60
II.3.5	Résultats et discussions . . . . .	61
II.4	Conclusion . . . . .	66
<b>III</b>	<b>Interprétation de scène basée sur un capteur ultrason</b>	<b>68</b>
III.1	Approche de détection et d'évitement d'obstacles . . . . .	69
III.2	Classification binaire : Catégorisation de la rigidité de l'objet détecté en Dur/Mou	71
III.2.1	Impact physique de la rigidité de matériaux des objets détectés sur l'écho ultrasonore . . . . .	72
III.2.2	Approche statistique pour la reconnaissance de la rigidité matérielle (SA_MRR : Statistical approach for material rigidity recognition) . . . .	73
III.2.2.1	Description . . . . .	73
III.2.2.2	Évaluation et discussions . . . . .	75
III.2.3	Modélisation possibiliste pour la reconnaissance de la rigidité matérielle (PM_MRR : Possibilistic modeling for material rigidity recognition) . . .	80
III.2.3.1	Description . . . . .	82
III.2.3.2	Évaluation et discussions . . . . .	82
III.3	Classification multi-classes : Reconnaissance de cinq matériaux . . . . .	86
III.3.1	Utilisation de CBH(1) pour la reconnaissance de cinq matériaux : Simu- lation numérique . . . . .	87
III.3.2	Utilisation de CBH(2) pour la reconnaissance de cinq matériaux : Simu- lation numérique . . . . .	91
III.3.3	Évaluation et discussions . . . . .	98
III.4	Conclusion . . . . .	99
<b>IV</b>	<b>Interprétation de scène basée sur une caméra monoculaire RVB</b>	<b>100</b>

IV.1	État de l'art . . . . .	102
IV.2	La saillance visuelle . . . . .	103
IV.2.1	Modèles de prédiction de la fixation oculaire . . . . .	104
IV.2.2	Modèles de détection des régions saillantes . . . . .	104
IV.3	Approche d'interprétation de scène par une caméra monoculaire RVB . . . . .	106
IV.3.1	Création d'une image panoramique . . . . .	106
IV.3.2	Génération d'une carte de saillance (1 <sup>ère</sup> simplification de scène) . . . . .	109
IV.3.3	Simplification de carte de saillance (2 <sup>ème</sup> simplification de scène) . . . . .	110
IV.3.3.1	Binarisation . . . . .	111
IV.3.3.2	Opérateurs morphologiques . . . . .	112
IV.3.4	Détection des objets saillants . . . . .	113
IV.3.5	Résultats et discussions . . . . .	113
IV.4	Conclusion . . . . .	115

## **V Perception de scène par un système multi-capteurs, application à l'aide de navigation des déficients visuels . . . . . 125**

V.1	État de l'art sur les systèmes d'aide à la navigation des déficients visuels . . . . .	126
V.2	Présentation du système d'aide à la navigation ("NA_System" : Navigation Aid System) . . . . .	128
V.2.1	Historique de NA_System . . . . .	129
V.2.1.1	Premier dispositif . . . . .	129
V.2.1.2	Deuxième dispositif . . . . .	130
V.2.1.3	Limitations des deux dispositifs . . . . .	130
V.2.2	Description du NA_System . . . . .	131
V.2.2.1	Capteurs utilisés . . . . .	132
V.2.2.2	Espace de travail . . . . .	133
V.3	Interprétation de scène par fusion de connaissances . . . . .	134
V.3.1	Adaptation du modèle générique d'interprétation de scène . . . . .	134
V.3.2	Recalage des informations . . . . .	135
V.3.2.1	Détermination du pixel pointé par le capteur ultrason dans l'image . . . . .	137
V.3.2.2	Recalage du signal de profil et l'image panoramique . . . . .	139

---

V.3.3 Résultats et discussions . . . . .	140
V.4 Conclusion . . . . .	141
<b>VI Conclusions et perspectives</b>	<b>148</b>
VI.1 Synthèse des travaux présentés dans ce mémoire . . . . .	148
VI.1.1 Motivations et problématique . . . . .	148
VI.1.2 Bilans et conclusions . . . . .	149
VI.2 Perspectives . . . . .	151
VI.2.1 Perspectives à court terme . . . . .	151
VI.2.2 Perspectives à long terme . . . . .	152
<b>Publications</b>	<b>154</b>
<b>Bibliographie</b>	<b>156</b>

# Table des figures

1	Les différents niveaux d'interprétation de scène selon l'échelle sémantique. . . . .	2
2	Structuration de la thèse. . . . .	5
I.1	Perception d'une boîte par un robot SLAM [25] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	8
I.2	Graphe de modélisation [25] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	8
I.3	Les techniques de localisation dans un environnement intérieur (illustration retirée afin de respecter les droits d'auteurs) . . . . .	9
I.4	Exemples de techniques basées sur les propriétés du signal [30] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	9
I.5	Techniques basées sur les algorithmes de localisation [30] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	10
I.6	Les différents capteurs dans un système d'interprétation (illustration retirée afin de respecter les droits d'auteurs) . . . . .	11
I.7	Exemple d'utilisation d'une caméra stéréoscopique [55] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	15
I.8	Exemple d'une image acquise par une caméra infrarouge [59] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	15
I.9	Exemple de grille d'occupation : Les zones sombres indiquent une forte probabilité de présence d'un obstacle [67] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	18

---

I.10	Exemple de carte métrique et de carte topologique [69] (illustration retirée afin de respecter les droits d’auteurs) . . . . .	18
I.11	Exemple de carte hybride [69] (illustration retirée afin de respecter les droits d’auteurs) . . . . .	19
I.12	Modélisation d’un environnement intérieur . . . . .	19
I.13	Exemples de scènes. . . . .	20
I.14	Système d’interprétation de scène par fusion de connaissance . . . . .	21
I.15	Effet de la profondeur sur l’interprétation de scène : (a) système de perception placé à une distance $d_o$ , (b) système de perception placé à une distance $d'_o < d_o$ .	22
I.16	Caractérisation d’un objet à partir du capteur de vision . . . . .	22
I.17	Positionnement des objets dans une scène . . . . .	24
I.18	Obtention d’informations distinctes à partir des mêmes données confrontées à différents contextes [7] (illustration retirée afin de respecter les droits d’auteurs)	25
I.19	Les connaissances véhiculées par une distribution de possibilités [7] (illustration retirée afin de respecter les droits d’auteurs) . . . . .	28
I.20	Fusion conjonctive par l’opérateur min (Fig.2.4 de [6]) (illustration retirée afin de respecter les droits d’auteurs) . . . . .	31
I.21	Fusion disjonctive par l’opérateur max (Fig.2.3 de [6])(illustration retirée afin de respecter les droits d’auteurs) . . . . .	32
II.1	Détermination de la spécificité intra-classe et inter-classe . . . . .	38
II.2	Différentes situations de représentation de deux classes $C_1$ et $C_2$ selon les formes et les dispositions de leurs modèles respectifs . . . . .	42
II.3	Différentes situations de représentation deux distributions de possibilités selon la forme de leur consensus . . . . .	44
II.4	Base de données synthétiques . . . . .	45
II.5	Pouvoirs discriminants déterminés selon les différentes approches de sélection d’attributs proposées. . . . .	46
II.6	Taux de correcte classifications obtenus par les différentes approches de sélection d’attributs sur les différentes bases de données . . . . .	47

II.7	Variation du TIC en fonction du TSA . . . . .	48
II.8	Moyenne des mesures TIC calculés pour les approches de sélection d'attributs évalués sur 4 bases de données . . . . .	48
II.9	Écart type des mesures TIC calculés pour les approches de sélection d'attributs évalués sur 4 bases de données . . . . .	48
II.10	Incertitude liée aux attributs dans un processus de classification . . . . .	51
II.11	Situation des mesures de dissimilarité utilisées sur une échelle sémantique . . . .	52
II.12	Architecture du CBH(1) . . . . .	53
II.13	Classification par maximum de possibilité . . . . .	54
II.14	Classification par maximum de similarité entre le modèle de l'objet X et les modèles des classes . . . . .	55
II.15	Architecture du CBH(2) . . . . .	57
II.16	Classification par maximum de similarité entre le modèle de l'objet X et les modèles des classes . . . . .	57
II.17	Performances des classifieurs sur la base de données colonne vertébrale . . . . .	62
II.18	Performances des classifieurs sur la base de données Diabètes . . . . .	63
II.19	Performances des classifieurs sur la base de données Graines . . . . .	63
II.20	Performances des classifieurs sur la base de données Ionosphère . . . . .	64
II.21	Performances des classifieurs sur la base de données Iris . . . . .	64
II.22	Performances des classifieurs sur la base de données Véhicule . . . . .	65
II.23	Performances des classifieurs sur la base de données Verre . . . . .	65
II.24	Performances des classifieurs sur la base de données Vin . . . . .	66
II.25	Analyse comparative basée sur le taux correct de classification . . . . .	66
III.1	Interprétation de scène observée par un capteur ultrason . . . . .	70
III.2	Différents scénarios de balayage d'une même scène (a,b,c,d), (A) leurs signaux de profil et (B) leurs détails d'ondelettes respectifs . . . . .	71
III.3	Approches proposées pour la reconnaissance de la rigidité de matériaux des objets	72

---

III.4 Différents phénomènes d'interaction de l'onde ultrasonore . . . . .	72
III.5 Impact de la rigidité de matériaux des objets détectés sur l'écho ultrasonore . . .	73
III.6 Approche statistique pour la reconnaissance de la rigidité des matériaux des objets détectés . . . . .	74
III.7 Exemples de représentation temps-fréquence de l'écho ultrasonore réfléchi par (a) un objet dur et (b) un objet mou . . . . .	74
III.8 Largeurs des faisceaux des capteurs à ultrasons utilisés . . . . .	75
III.9 Objets utilisés pour la construction de la base de données . . . . .	76
III.10 Exemples d'échos ultrasonores réfléchis par un objet dur et un objet mou . . .	76
III.11 Structure de la base de données . . . . .	77
III.12 Impact de la distance sur l'écho ultrasonore . . . . .	78
III.13 Comparaison de précision de EZ2 et EZ4 à une distance 50 cm . . . . .	78
III.14 Taux des différentes métriques pour chaque classifieur . . . . .	80
III.15 Les valeurs de mesures d'entropie . . . . .	81
III.16 Approche possibiliste pour la reconnaissance de la rigidité de matériaux des objets détectés . . . . .	83
III.17 Structure de la base de données utilisée dans PM_MRR . . . . .	84
III.18 Distributions de possibilités correspondantes aux attributs sélectionnés . . . . .	85
III.19 Processus de prise de décision basé sur la mesure de similarité . . . . .	85
III.20 Taux de classification achevés avant et après sélection d'attributs en se basant sur une prise de décision vectorielle versus une prise de décision basée sur des modèles possibilistes . . . . .	86
III.21 Distributions de possibilités de l'ensemble $\Omega^0$ selon $f_{12}$ . . . . .	87
III.22 Distributions de possibilités de l'ensemble $\Omega^1$ selon $f_{11}$ . . . . .	88
III.23 Distributions de possibilités de l'ensemble $\Omega^2$ selon $f_{13}$ . . . . .	89
III.24 Distributions de possibilités correspondantes aux classes $C_2$ et $C_4$ selon $f_9$ . . .	90
III.25 CBH(1) résultant . . . . .	91
III.26 Distributions de possibilités correspondantes à $C_{op}^1$ et $\Omega^1$ . . . . .	92

---

III.27 Distributions de possibilités correspondantes à $C_{op}^2$ et $\Omega^2$ . . . . .	93
III.28 Distributions de possibilités correspondantes à $C_{op}^3$ et $\Omega^3$ . . . . .	94
III.29 Distributions de possibilités correspondantes à $C_4$ et $C_5$ . . . . .	97
III.30 CBH(2) résultant . . . . .	97
III.31 Performances des classifieurs sur des données ultrasonores . . . . .	98
IV.1 Le fossé sensoriel et le fossé sémantique (modification de la Fig.1.1 de [6]) (illustration retirée afin de respecter les droits d'auteurs) . . . . .	102
IV.2 Génération de cartes de saillance en appliquant (a) un modèle de prédiction de fixation oculaire et (b) un modèle de prédiction des régions saillantes . . . . .	106
IV.3 Approche d'interprétation de scène à partir d'une caméra monoculaire RVB . . . . .	107
IV.4 Présentation de ICE . . . . .	108
IV.5 Exemples de panoramas (a) avant et (b) après l'extrapolation . . . . .	108
IV.6 Exemple illustratif de connexions aux limite (illustration retirée afin de respecter les droits d'auteurs) . . . . .	109
IV.7 Cartes de saillance générées appliquant différents (a) modèles de détection d'objets saillants et (b) modèles de prédiction de fixation oculaire . . . . .	114
IV.8 Application de la stratégie proposée sur un exemple de scénarii . . . . .	124
V.1 Quelques systèmes d'aides à la navigation des aveugles [210] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	127
V.2 Premier dispositif . . . . .	130
V.3 Deuxième dispositif . . . . .	130
V.4 Influence de la direction de balayage sur la construction du signal de profil . . . . .	131
V.5 Vue en 3D du dispositif du le système d'aide à la navigation conçu [7] . . . . .	132
V.6 Schéma fonctionnel du "NA_System" conçu . . . . .	132
V.7 Carte Raspberry Pi 3 Model B (illustration retirée afin de respecter les droits d'auteurs) . . . . .	133
V.8 Connexion à SmartCane . . . . .	134

---

V.9 Interface du django Web "SmartCane" . . . . .	134
V.10 Module de capteurs considéré pour la perception de scène . . . . .	134
V.12 Positionnement des capteurs dans NA_System . . . . .	135
V.11 Système d'interprétation de scène perçue par des capteurs ultrasons et une caméra monoculaire RVB . . . . .	136
V.13 Différents champs de vision pour la caméra intégrée dans NA_System . . . . .	137
V.14 Détermination du nombre de pixel x entre le centre du capteur ultrason et le centre de la caméra en fonction de la distance séparant le système de l'objet . . . . .	138
V.15 Détermination du pixel pointé par le capteur ultrason dans l'image par rapport au centre de l'image . . . . .	139
V.16 Principe de recalage le signal de profil avec l'image panoramique . . . . .	140
V.17 Application du modèle d'interprétation proposé sur un exemple de scénarii . . . . .	146

# Liste des tableaux

II.1	Pouvoirs discriminants déterminés selon les différentes approches de sélection d'attributs proposées. . . . .	45
II.2	Résultats de la sélection d'attributs obtenus pour différentes approches. . . . .	47
II.3	Détails des bases de données utilisées. . . . .	59
II.4	Évaluation de la base de données Colonne vertébrale. . . . .	62
II.5	Évaluation de la base de données Diabètes. . . . .	63
II.6	Évaluation de la base de données Graines. . . . .	63
II.7	Évaluation de la base de données Ionosphère. . . . .	64
II.8	Évaluation de la base de données Iris. . . . .	64
II.9	Évaluation de la base de données Véhicule. . . . .	65
II.10	Évaluation de la base de données Verre. . . . .	65
II.11	Évaluation de la base de données Vin. . . . .	66
III.1	Matrice de confusion de chaque classifieur . . . . .	79
III.2	TPD et TPM pour chaque classifieur . . . . .	80
III.3	Pouvoirs discriminants d'attributs correspondants à la discrimination de " $\pi_{C_{Dur}}$ " et " $\pi_{C_{Mou}}$ ". . . . .	85
III.4	Contexte. . . . .	87
III.5	Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 1 <sup>ère</sup> itération du CBH(1). . . . .	88
III.6	Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 2 <sup>ème</sup> itération du CBH(1). . . . .	89

---

III.7 Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 3 <sup>ème</sup> itération du CBH(1). . . . .	90
III.8 Pouvoirs discriminants d'attributs correspondants à la discrimination de $C_2$ et $C_4$ . . . . .	90
III.9 Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 1 <sup>ère</sup> itération du CBH(2). . . . .	95
III.10 Score du chaque couple possible pour la 1 <sup>ère</sup> itération du CBH(2). . . . .	95
III.11 Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 2 <sup>ème</sup> itération du CBH(2). . . . .	95
III.12 Score du chaque couple possible pour la 2 <sup>ème</sup> itération du CBH(2). . . . .	96
III.13 Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 3 <sup>ème</sup> itération du CBH(2). . . . .	96
III.14 Score du chaque couple possible pour la 3 <sup>ème</sup> itération du CBH(2). . . . .	96
III.15 Pouvoirs discriminants d'attributs correspondants à la discrimination de $C_4$ et $C_5$ . . . . .	96
III.16 Évaluation de CBH(1) et CBH(2) sur une base de données des signaux ultrasonores. . . . .	98
IV.1 Comparaison du temps de traitement requis pour chaque modèle. . . . .	114
V.1 Systèmes d'aide à la navigation des personnes à déficience visuelle [210] (illustration retirée afin de respecter les droits d'auteurs) . . . . .	128
V.2 Comparaison entre Raspberry Pi 2 et Raspberry Pi 3. (illustration retirée afin de respecter les droits d'auteurs) . . . . .	133
V.3 Spécification de la carte Arduino Nano V3.0 [252]. (illustration retirée afin de respecter les droits d'auteurs) . . . . .	133

# Introduction générale

## Description du sujet de thèse

À l'origine, les robots ont été installés dans les industries, afin d'accomplir des tâches répétitives avec une précision constante. La tendance actuelle est de les employer partout. Dans ce contexte la navigation perceptuelle d'un robot de façon autonome et robuste, dans des environnements intérieurs structurés et fortement encombré, est un cadre de recherche d'actualité. L'étude dans le présent projet est focalisée sur le développement d'un système d'interprétation de scènes d'environnements intérieurs, observée par un système multi-capteurs pour assurer une navigation autonome. Dans ce cadre, trois concepts sont à définir, à savoir "la navigation", "une scène" et "un système d'interprétation".

**Navigation** : Le mot navigation est dérivé du latin, du mot "navis" qui signifie navire, et le mot "agere" qui signifie déplacement, ou manipulation [1, 2]. La définition de la navigation varie en fonction de l'activité. Par exemple, certains chercheurs ont défini la navigation par rapport aux objets ou aux véhicules plutôt qu'aux humains[1, 2], alors qu'autres chercheurs pensent que les humains sont engagés dans les systèmes de navigation [3, 4]. Néanmoins, toutes ces définitions ont trois points en commun : (i) déterminer le point de départ d'une personne ou d'un objet, (ii) connaître la destination de la personne ou de l'objet, et (iii) déterminer les informations indispensables pour un meilleur itinéraire du point de départ à la destination [5].

**Scène** : Le terme "scène" fait référence à un sous-ensemble d'un environnement dans lequel sont considérés plusieurs objets ayant une signification sémantique et des relations spatiales entre eux [6].

**Système d'interprétation** : Un système d'interprétation a plusieurs aspects : d'une part, un système d'interprétation utilise des capteurs capables de collecter des informations du monde extérieur. D'autre part, un système d'interprétation intègre des connaissances a priori liées à l'expert et ses capacités cognitives appropriés. Par exemple, dans le cadre de traitement

d'images médicales, les médecins sont considérés comme des experts qui fournissent la vérité terrain des images afin d'évaluer la stratégie d'interprétation considérée.

Il est particulièrement intéressant de noter que l'interprétation chez les humains est une tâche simple dans la vie quotidienne. Cependant, les humains ne sont pas tout à fait passifs dans leurs interprétations car ils prennent en considération leurs connaissances a priori d'une part et leurs capacités cognitive et d'observation d'autre part. De même, un système d'interprétation de scène peut intégrer multiples capteurs pour extraire autant d'informations que possible.

Sur l'échelle sémantique d'interprétation, quatre niveaux de représentation des connaissances sont placés selon la sémantique contenue dans les connaissances [7], comme le montre la Figure 1.

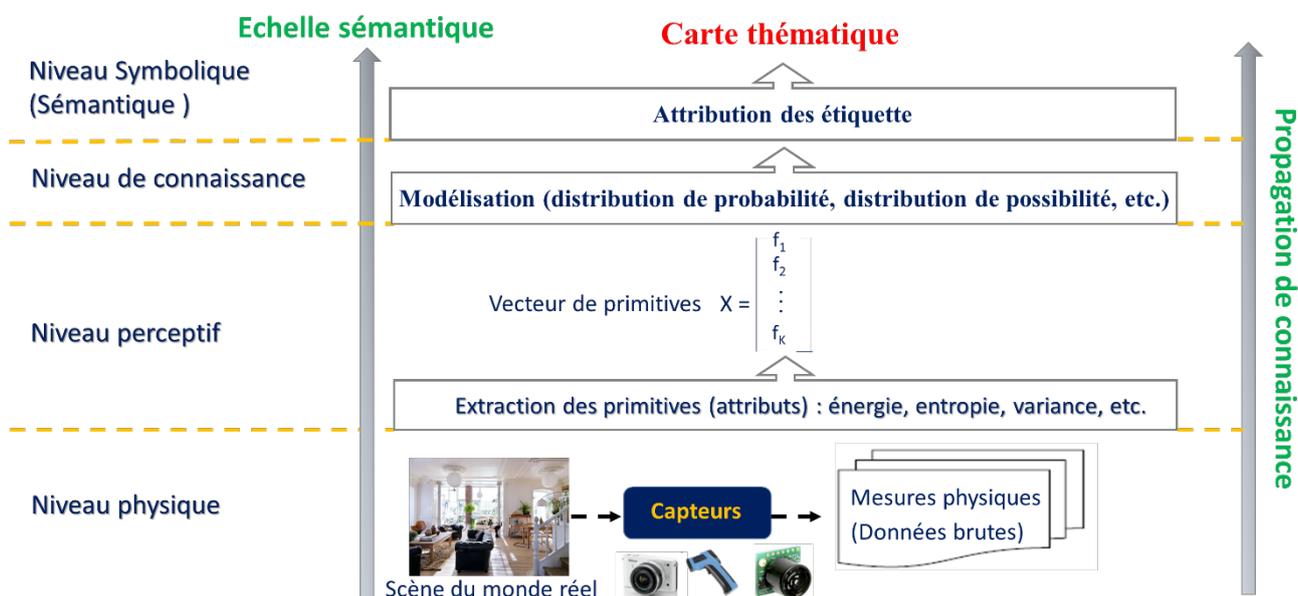


FIGURE 1 – Les différents niveaux d'interprétation de scène selon l'échelle sémantique.

- *Niveau physique* : Le niveau physique est le niveau le plus bas de l'échelle. Il consiste en représentation des éléments de la scène par les données brutes issues des capteurs.
- *Niveau perceptif* : À ce niveau, les éléments de scène sont représentés par des paramètres (attributs) à partir des transformations des données brutes ou par des connaissances injectées de l'extérieur. En d'autres termes, chaque élément de la scène est relié à un vecteur de primitives qui le représente.
- *Niveau de connaissance* : Ce niveau comprend l'injection de connaissances a priori de l'expert liées aux classes thématiques ("objet", "arrière-plan") présentées dans la scène à interpréter afin de construire un modèle destiné au système d'interprétation. À ce niveau,

un élément de la scène est projeté dans un autre espace de connaissances, où il est caractérisé par un vecteur qui représente son score de proximité comparé à chaque modèle des classes thématiques. Comme exemples de modèles, nous trouvons la distribution de masse, la distribution de probabilité, la distribution de possibilité, etc.

- *Niveau symbolique ou sémantique* : Il s'agit du niveau d'échelle le plus élevé, qui représente un élément de scène par le biais de connaissances symboliques (par exemple, des étiquettes ou encore appelées des labels), qui spécifient son appartenance à une classe thématique. À ce niveau, tous les éléments de la scène sont transformés en une carte thématique qui représente le résultat final de l'interprétation. Cette interprétation implique l'attribution d'une étiquette à chaque élément de la scène. Ainsi, une étiquette est considérée comme le représentant le plus élevé sur l'échelle sémantique de l'élément. La définition des étiquettes (classes) est prédéterminée par des experts après analyse des structures de la scène.

## Problématiques d'un système d'interprétation de scène

Dans un système multi-capteurs d'interprétation de scène, les informations, extraites des données fournies par les capteurs, présentent des restrictions du monde réel et sont généralement affectées par des imperfections de nature épistémique. Ces imperfections doivent être pris en compte lors de la modélisation. L'incertitude épistémique provient soit de la description linguistiques des connaissances, soit de l'utilisation d'un petit nombre d'échantillons.

Par conséquent, nous remarquons que dans le contexte des systèmes d'interprétation de scène, la théorie des possibilités est un cadre approprié pour représenter et traiter ces connaissances. Cette théorie présente de nombreux avantages qui la rendent très utile lors de l'interprétation de scène, à savoir :

- La capacité à représenter des connaissances ambiguës
- La simplicité des outils utilisés par cette pour traiter et fusionner les informations ambiguës, et pour quantifier l'incertitude associée à l'occurrence de tous les événements du contenu informationnel (par l'utilisation des mesures de possibilité et de nécessité)
- La flexibilité et la simplicité pour traiter et fusionner les informations provenant de différentes sources de connaissances

Outre la gestion des informations imparfaites lors du développement d'un système d'interprétation multi-capteurs, il est également nécessaire d'adopter un modèle simple qui doit être conforme à la tâche que nous souhaitons réaliser. C'est dans ce contexte que s'inscrit notre

---

sujet de thèse, qui consiste au développement d'un système d'interprétation de scènes d'environnements intérieurs, observées par un système multi-capteurs combinant un capteur ultrason et une caméra RVB.

De manière générale, nos recherches visent à intégrer le système proposé dans un dispositif utilisé par les personnes à déficience visuelle, ou encore dans un robot opérant dans des espaces meublés.

## Contributions

Le travail présenté dans cette thèse englobe plusieurs contributions. En effet, comme première étape, nous avons établi un état de l'art sur les systèmes de navigation dans un environnement d'intérieur. L'objectif était de connaître les capteurs utilisés dans tels systèmes, les techniques adoptées, et les modèles générés pour l'interprétation de scène.

Pour une meilleur fiabilité du système d'interprétation de scène, nous avons opté à fusionner deux stratégies d'interprétation de scène. La première stratégie consiste à utiliser des capteurs ultrasons pour détecter, déterminer la profondeur des obstacles et identifier leur rigidité matérielle. La deuxième stratégie est basée sur une caméra RVB, qui consiste à simplifier la scène autant que possible. En effet, nous avons opté à détecter que les objets d'intérêt.

Également, dans la chaîne de traitement de données adoptées pour l'analyse et l'interprétation de la scène, de nouvelles approches ont été proposées, notamment pour la sélection d'attributs, la classification, la fusion et le recalage de données issues des deux capteurs utilisés.



FIGURE 2 – Structuration de la thèse.

# Chapitre I

## Cadre applicatif et théorique de l'étude

### Sommaire

---

I.1	Cadre applicatif . . . . .	7
I.2	Proposition d'un modèle générique d'interprétation de scène . . .	19
I.3	Cadre théorique pour le développement du système d'interprétation proposé . . . . .	23
I.4	Conclusion . . . . .	34

---

## Introduction

Au cours de ces dernières années, l'utilisation des systèmes de navigation intérieure ont connu une croissance exponentielle. Plusieurs systèmes pour le suivi et la localisation d'objets meublant des environnements intérieurs, ont été proposés. L'utilisation de multiples capteurs et processus de traitement ont eu tendance à augmenter la complexité de ces systèmes notamment au niveau des modèles de scène qu'ils adoptent pour l'interprétation. Dans une optique d'alléger le traitement des données issues de multiples sources dans un système d'aide à la navigation, nous proposons une solution pour l'interprétation d'une scène d'un environnement intérieur perçue par un système multi-capteurs, basée sur un modèle simple qui, malgré sa simplicité, assure une aide à la navigation.

Il est à noter que l'analyse d'une scène est une étape fondamentale dans le processus d'interprétation de cette scène. Cette étape comprend l'extraction d'informations à partir des données provenant de plusieurs sources de connaissances. Dans la plupart des cas, ces informations présentent des imperfections, tels que l'ambiguïté, l'incomplétude, l'inexactitude, l'incertitude, le conflit, etc. Ces imperfections persistent depuis l'acquisition des données via des sources de connaissances jusqu'au traitement et l'interprétation [8]. Par conséquent, afin de rendre l'interprétation de scène fiable, cohérente et riche, ces imperfections doivent être prises en compte à chaque étape du traitement et d'analyse.

Dans cette optique, nous commençons par présenter dans ce chapitre, notre étude menée pour l'exploration des systèmes de navigation les plus connus dans la littérature, ainsi que les techniques utilisées par ces systèmes et leurs capteurs intégrés, puis nous enchaînons par la présentation de notre proposition pour un modèle d'interprétation de scène. Enfin, nous introduisons les concepts de base liés à l'information au son sens habituel, aux imperfections qui l'affectent et les théories qui les traitent. Une attention particulière a été menée à la théorie de possibilités qui est considérée comme le cadre général du chapitre suivant.

## I.1 Cadre applicatif

### I.1.1 Systèmes de localisation et de cartographie simultanées

Les systèmes de localisation et de cartographie simultanées sont connus en anglais sous le nom de SLAM (Simultaneous Localization and Mapping). Ces systèmes sont utilisés pour établir une carte d'environnement et de s'y localiser [9]. Cadena *et al.*[10] estime que l'architecture des robots SLAM se compose de deux parties : le front-end et le back-end, comme le montre la

Figure 2 de leur article.

Un robot SLAM est muni d'un ensemble de capteurs et a pour objectifs d'assurer une topologie cartographique efficace, une estimation de localisation fiable, etc. [11, 12]. Le problème de la cartographie est l'un des problèmes les plus importants et les plus complexes des robots SLAM [13]. En effet, le robot doit se situer précisément [14], et pour cela, il a besoin d'une carte précise. La corrélation entre la carte et la position globale rend le problème difficile [15]. La seule façon de résoudre ce problème est d'estimer continûment et simultanément la position du robot et de la carte [16].

Les robots SLAM sont coûteux. Dans ce cadre, de nombreux travaux ont été proposés. Par exemple, dans [17], un microcontrôleur est utilisé comme un processeur principal. Néanmoins, le système est loin d'être en temps réel car il nécessite une demi-minute pour chaque étape de prédiction. Dans [18], les auteurs ont envisagé d'utiliser 6 capteurs infrarouges au lieu d'utiliser un seul télémètre coûteux, tandis que d'autres chercheurs suggèrent d'équiper leur système par des capteurs à ultrasons vu que ces capteurs sont de bon marché [19, 20].

Outre le prix élevé de ces robots, les algorithmes SLAM nécessitent un ordinateur de capacité de traitement élevée, c'est-à-dire un ordinateur de bureau ou un ordinateur portable. Dans [21], les auteurs ont proposé une optimisation d'architecture en utilisant plusieurs processeurs (DSP, processeur RISC, GPU) et un coprocesseur pour le prétraitement des données. D'autres travaux [22, 23] ont proposé d'implémenter la partie de l'algorithme qui nécessite une puissance de calcul élevée sur une FPGA (Field-Programmable Gate Array).

La complexité du graphe de modélisation présente un autre problème pour les robots SLAM [24]. Par exemple, la Figure I.1 présente une observation réelle d'une boîte prise par un robot SLAM équipé d'une caméra RVB-D et d'un capteur laser, ainsi que les différents descripteurs utilisés. La Figure I.2 présente le graphe de modélisation correspondant à cette boîte, qui paraît très compliqué.

FIGURE I.1 – Perception d'une boîte par un robot SLAM [25] (illustration retirée afin de respecter les droits d'auteurs)

Il est à noter que malgré toutes les méthodes d'optimisation mentionnées ci-dessus, il n'est pas souhaitable d'utiliser un robot SLAM sans tirer avantage de son matériel et de ses logiciels à haute puissance.

FIGURE I.2 – Graphe de modélisation [25] (illustration retirée afin de respecter les droits d'auteurs)

## I.1.2 Les techniques de localisation

Afin de déterminer l'emplacement d'un objet, deux techniques de localisation ont été proposées [26, 27] comme le montre la Figure I.3.

FIGURE I.3 – Les techniques de localisation dans un environnement intérieur (illustration retirée afin de respecter les droits d'auteurs)

### I.1.2.1 Techniques basées sur les propriétés du signal

La détermination de la position d'un objet par ces techniques est basée sur des paramètres géométriques du signal, à savoir l'angle, la distance et la puissance du signal. Dans ce qui suit, nous présentons les techniques de localisation basées sur les propriétés de signal les plus courantes [27, 28].

**Angle d'arrivée (Angle of Arrival : AoA) :** est déterminé en fonction de l'intersection des lignes directrices entre des points de référence [28]. Sachant l'AoA, la position de l'objet peut alors être estimée et utilisée à des fins de suivi ou de navigation. Dans [29], les auteurs ont souligné qu'en pratique, les systèmes basés sur cette mesure sont souvent coûteux et complexes.

**Temps d'arrivée (Time of Arrival : ToA) :** Bien que l'AoA soit basée sur la mesure de l'angle, le ToA est essentiellement basé sur le temps nécessaire pour qu'un signal atteigne un récepteur d'un émetteur fixe. Le ToA est également reconnu en anglais par ToF (Time of Flight) [29].

**Différence du temps d'arrivée (Time Difference of Arrival : TDoA) :** est considérée comme la différence du ToA entre deux capteurs différents, éliminant ainsi la nécessité de connaître quand le signal a été envoyé [27, 28].

**Indication de force du signal reçu (Received Signal Strength Indication : RSSI) :** est une mesure du niveau de puissance lors de la réception d'un signal radio. Son objectif est d'indiquer l'intensité du signal pour une utilisation ultérieure dans l'estimation de la distance [31]. Cependant, cette technique peut être affectée par les trajets multiples dans les environnements intérieurs, ce qui peut conduire à une précision réduite [32, 33].

FIGURE I.4 – Exemples de techniques basées sur les propriétés du signal [30] (illustration retirée afin de respecter les droits d'auteurs)

### I.1.2.2 Les Algorithmes de localisation

Les propriétés du signal et les algorithmes de localisation fonctionnent ensemble pour estimer la position d'un objet. En fait, l'algorithme de localisation traite la propriété du signal pour générer ultérieurement l'emplacement d'un objet cible [26]. Ainsi, la précision des informations de localisation dépend à la fois de la précision de la propriété du signal et de l'exactitude de l'algorithme de localisation utilisé.

**Triangulation** : La triangulation repose sur les caractéristiques géométriques des triangles pour déterminer la position des objets en se basant sur l'AoA entre deux points de référence connus [33, 34].

**Trilatération** : La trilatération est souvent considérée l'équivalent de la triangulation. En fait, ces deux algorithmes s'appuient sur les caractéristiques géométriques des triangles pour déterminer l'emplacement des objets [31, 32]. Cependant, la triangulation est basée sur l'AoA, tandis que la trilatération se base sur ToA ou TDoA [27].

FIGURE I.5 – Techniques basées sur les algorithmes de localisation [30] (illustration retirée afin de respecter les droits d'auteurs)

**Proximité** : L'algorithme "proximité" repose sur la propriété du signal RSSI pour déterminer l'emplacement des objets mobiles à partir d'une grille d'antenne [26, 32]. Lorsqu'un objet en mouvement est détecté, l'antenne la plus proche de lui intervient pour calculer sa position. Dans le cas où un objet est détecté par plusieurs antennes, l'antenne dont le signal le plus puissant intervient pour calculer sa position [26, 27].

**Analyse de scène** : L'algorithme "Analyse de scène" collecte d'abord des informations sur la scène, puis il estime l'emplacement des objets en faisant correspondre ces informations avec celles d'une base de données de valeurs RSSI [27, 32, 35]. Les informations recueillies représentent l'empreinte digitale, qui caractérise une scène d'une autre. Cependant, lorsque l'emplacement d'un objet a été modifié, cet algorithme nécessite beaucoup de temps et d'efforts pour mettre à jour la base de données [26]. Par conséquent, les principaux défis de cet algorithme sont le long temps de traitement et le coût de calcul élevé [32].

### I.1.3 Sources d'informations

Chez l'être humain, une scène est interprétée à travers leurs organes sensoriels. Dans les applications de l'intelligence artificielle, les organes sensoriels sont équivalents à des capteurs (Par exemple, une caméra peut être utilisée pour la vision, un télémètre peut être utilisé pour le

positionnement, etc.). Ainsi, une scène fait référence à un sous-ensemble des résultats de mesure capturés par un ou plusieurs capteurs et représentés sur différents types de supports (description linguistique, son, image, vidéo, etc.).

Comme tous les êtres vivants, tout le monde n'a pas la même vision (la vision des humains est complètement différente de celle des insectes). À l'instar du monde biologique, différentes perceptions peuvent être générées en utilisant différents capteurs correspondant à la même catégorie (par exemple, pour les capteurs de vision, on peut trouver des caméras monoculaires, des caméras omnidirectionnelles, etc.).

Pour assurer une navigation autonome, un tel système a besoin d'interpréter soigneusement son environnement. Ce défi est fortement lié aux capteurs dont le système était équipé. De ce fait, on a tenté de présenter dans ce qui suit un bref revue sur les capteurs utilisés dans les systèmes de navigation, tout en les évaluant selon différents critères tels que leur coût, leur portée de détection, etc. Comme le montre la Figure I.6, les capteurs équipant les systèmes de navigation appartiennent à l'une des deux catégories : capteurs proprioceptifs et capteurs extéroceptifs [36].

FIGURE I.6 – Les différents capteurs dans un système d'interprétation (illustration retirée afin de respecter les droits d'auteurs)

### I.1.3.1 Capteurs proprioceptifs

Les capteurs proprioceptifs fournissent des informations internes du système de navigation, y comprennent la vitesse, l'accélération, les angles de rotation et d'attitude. Comme exemple du capteur proprioceptif le plus répandu, on peut citer l'encodeur [37], qui est généralement fixé sur la roue motrice ou directement sur l'arbre moteur associé à la roue. Les capteurs proprioceptifs peuvent être divisés en deux familles [38] :

1. **Les capteurs de déplacement** : Ces capteurs permettent de mesurer le déplacement élémentaire, le changement d'accélération ou de vitesse sur des pistes rectilignes ou curvilignes. Parmi ces capteurs, on peut trouver l'encodeur, l'odomètre, l'accéléromètre, le RADAR Doppler, et les mesureurs optiques.
2. **Les capteurs d'attitude** : Ces capteurs sont principalement de type inertiel, qui peuvent mesurer deux types de données, à savoir les angles de capteur et celles de roulis et de tangage. Cette famille y comprend les gyromètres, les gyroscopes, les inclinomètres, les capteurs inertiels composites, et les magnétomètres.

### I.1.3.2 Capteurs extéroceptifs

Les capteurs extéroceptifs établissent une connexion forte entre le système de navigation et son environnement [39]. Comme le montre la Figure I.6, les capteurs extéroceptifs peuvent être des capteurs télémétriques, des capteurs de vision, ou bien des capteurs hybrides.

#### 1. Capteurs télémétriques

Les capteurs télémétriques sont principalement utilisés pour détecter les obstacles dans une scène. La distance maximale mesurée par ces capteurs dépend fortement de leur champ de vision. Dans ce qui suit, nous présentons les différents types de capteurs télémétriques les plus couramment utilisés dans les systèmes de navigation.

- **Télémètre infrarouge** : Il utilise un rayonnement électromagnétique pour mesurer la distance entre eux et les obstacles. Ce rayonnement est caractérisé par une longueur d'onde plus courte que celle des micro-ondes mais plus longue que la lumière visible. Le calcul de la distance est effectué en utilisant une des techniques de localisation, à savoir le ToA, le TDoA ou bien la trilatération. En général, les systèmes infrarouges ne consomment pas beaucoup d'énergie, sont très directionnels et prennent très peu de place.

Même s'il ne fait aucun doute que le coût de mise en oeuvre d'un télémètre infrarouge au niveau d'un système est très faible, sa précision et sa portée sont limitées [40]. Pour être mis en oeuvre dans de grandes surfaces et avoir une bonne précision, un système infrarouge nécessite l'intégration de nombreux récepteurs [35] ce qui augmente la complexité et également le coût du système. Cependant, le coût n'est pas la seule limite. Il existe encore des limitations techniques entre l'émetteur et le récepteur, notamment les problèmes d'interférence entre les rayons infrarouges avec la lumière du soleil ou la lumière fluorescente.

- **Télémètre ultrasons** : L'émetteur ultrason émet périodiquement de courtes impulsions sonores hautes fréquences (au-dessus du champ d'audibilité pour l'homme). Ces impulsions se propagent dans l'air à la vitesse du son qui est égale à 340m/s. Lorsqu'elles font face à un objet, elles se réfléchissent et reviennent sous forme d'écho au capteur [41]. Le principe de mesure de distance pour ces capteurs est basé sur les techniques de localisation ToA/TDoA. Les systèmes de navigation intérieure à ultrasons les plus connus incluent le système Dolphin [42, 43], le système Cricket [44], et le système Active Bat [45]. De nouveaux systèmes [46, 47, 48] ont été pro-

posés comme alternatives aux systèmes susmentionnés. L'utilisation d'un système à ultrasons peut être suffisant pour certaines applications et présente l'avantage d'être bon marché et moins encombrant. De plus, tous les matériaux réfléchissant le son, même les objets transparents, peuvent être détectés.

La portée des capteurs à ultrasons varie en fonction de leur fréquence (Par exemple, pour une fréquence de 40 kHz, la portée maximale peut atteindre 10 m, et pour une fréquence de 200 kHz, la portée maximale peut atteindre 1,5 m). Ces capteurs n'occupent qu'un petit espace et sont également rentables. En raison de leur compacité, ils sont faciles à remplacer. Les capteurs à ultrasons ne sont pas sensibles à la lumière et à la couleur, ils conviennent donc aux environnements sombres. De plus, un autre avantage de ces capteurs est que leurs ondes ne sont pas affectées par la poussière et les environnements brumeux. Cependant, il faut éviter les courants d'air qui dévient le signal de leurs destinations.

- **Télémètre laser** : Il peut faire la distinction entre les longues distances et les courtes distances. En effet, selon la plage de mesure, le laser utilise trois méthodes :
  - i) Dans le cas où la plage de mesure est inférieure ou égale à un mètre, le laser utilise la triangulation pour déterminer la distance.
  - ii) Dans le cas où la plage de mesure peut aller jusqu'à plusieurs kilomètres, le laser utilise une méthode de corrélation de phase. Les principaux inconvénients de cette méthode sont la faible résolution et le temps de réponse long.
  - iii) Pour les très longues portées, la distance est basée sur la technique ToA.

Le télémètre laser est très directionnel et possède une excellente précision lors de la mesure des distances. Néanmoins, il est très sensible aux conditions environnementales (neige, pluie, température, etc.). Il faut également noter que l'effet thermique du laser peut endommager les yeux.

- **Télémètre RADAR (Radio Detection And Ranging)** : C'est un système qui utilise des ondes électromagnétiques, qui ne sont pas affectées par les conditions météorologiques. Un tel système est utilisé pour détecter la direction d'un objet cible, déterminer sa distance, ainsi que sa vitesse. La direction de la cible est obtenue en captant la position angulaire de l'antenne. La distance est mesurée en fonction de la technique ToA, alors que la vitesse est déterminée en fonction du décalage de fréquence de l'écho- RADAR généré par l'effet Doppler.

Le RADAR est caractérisé par sa rapidité et sa portée extrêmement élevée, qui

lui permettant de détecter les objets stationnaires ou en mouvement à plus de 200 mètres. Cependant, il est difficile pour le RADAR de détecter les petits obstacles. De plus, le RADAR ne peut pas distinguer les objets ni les localiser. Par conséquent, les systèmes basés sur des télémètres RADAR conviennent mieux aux applications de navigation extérieure qu'aux applications de navigation intérieure.

- **Télémètre LiDAR (Light Detection and Ranging)** : Il utilise des ondes électromagnétiques (laser). Il peut collecter des informations sur la distance, l'humidité, la température, etc. Le calcul de la distance est basé sur la technique ToA. Le LiDAR est généralement un capteur encombrant fourni dans un boîtier robuste (Par exemple, le HDL-64E, qui mesure 8 pouces x 8 pouces x 11 pouces et pèse 13 Kg), mais une version compacte et légère a récemment été introduite (HDL-32). Selon le fabricant et le produit choisi, la plage de profondeur varie d'environ 50 m à plusieurs centaines de mètres, avec une précision de quelques centimètres.

Le LiDAR est l'un des capteurs utilisés les plus chers (Par exemple, le prix du HDL-64E est de 75 000 \$ tandis que le prix du VLP-16 est passé de 8 000 \$). En raison de la nature rotative de certains composants, le LiDAR peut subir une usure au cours du temps. Ainsi, un diagnostic fréquent est donc nécessaire. Étant donné que le LiDAR utilise des télémètres lasers, une analyse approfondie des risques d'interférences lasers multiples doit être effectuée.

## 2. Capteurs de vision

Dans cette section, nous présentons les capteurs de vision les plus couramment utilisés.

- **Caméra standard** : La caméra standard est une caméra RVB. Sa durée de vie dépend fortement de la durée de vie de l'obturateur. En d'autres termes, cela dépend du nombre de déclencheurs de ce dernier, avant que la caméra ne devienne aveugle. En règle générale, la durée de vie de la plupart des caméras standards est de 30 000 à 150 000 clics. La taille physique de ces caméras monoculaires est très compacte (généralement autour 5 *pouces*<sup>3</sup> et 20 *pouces*<sup>3</sup>). En raison de son matériau simple, de son poids léger, de sa compacité, de son assemblage simple et de son interface série, la caméra standard est facile à remplacer et est largement utilisée dans les systèmes de localisation et de cartographie simultanés SLAM [49, 50, 51].
- **Caméra stéréoscopique** : Inspirée de l'oeil humain, la stéréo vision utilise deux caméras standards installées côte à côte pour observer la même scène sous deux angles différents. Parmi les caméras stéréoscopiques les plus populaires, on trouve :

Ensenso, Bumblebee 2, Capella, Surveyor Stereo Vision System, etc. Ces caméras ont mené d'un certain nombre d'études dans le domaine de la navigation [52, 53, 54].

FIGURE I.7 – Exemple d'utilisation d'une caméra stéréoscopique [55] (illustration retirée afin de respecter les droits d'auteurs)

Un problème courant avec les systèmes de vision stéréo est le temps de traitement des images. Plusieurs solutions ont été proposées pour réduire le temps de calcul, notamment en utilisant des systèmes reconfigurables (comme le FPGA [56]). Après avoir installé le banc stéréo sur un robot, la nécessité de le calibrer régulièrement est considérée comme l'un des principaux problèmes de la vision stéréo. En fait, les performances d'un banc stéréo dépendent entièrement de l'étalonnage qui présente un problème difficile, surtout lorsqu'il doit être résolu en temps réel.

- **Caméra omnidirectionnelle** : La caméra omnidirectionnelle est semblable à une caméra standard mais avec un champ de vision de 360 degrés. Parmi les caméras omnidirectionnelles, on peut citer : GoPro Fusion, Samsung Gear 360 Ricoh Theta V, Yi Technology 360VR et Detu Twin. En raison de leur compacité, de leur interface série, elles sont utilisées dans différentes applications, telles que la navigation [57], la création de l'art panoramique [58], etc.
- **Caméra thermique/infrarouge** : La caméra infrarouge et la caméra thermique ont le même principe de fonctionnement qui réside dans l'enregistrement des rayonnements infrarouges émis par les objets et les corps. Cependant, chacune a son champ d'application. En effet, une caméra thermique est sensible aux ondes émises avec des longueurs d'onde de l'ordre de 10 à 5 microns, tandis qu'une caméra infrarouge est sensible aux ondes émises avec des longueurs d'onde de l'ordre de 10 à 6 microns. Ces deux types de caméras conviennent aux applications qui ne nécessitent pas beaucoup de détails.

FIGURE I.8 – Exemple d'une image acquise par une caméra infrarouge [59] (illustration retirée afin de respecter les droits d'auteurs)

Les caméras infrarouges/thermiques sont sensibles à la température. Ces caméras sont généralement encombrées à cause de leur grande dimension qui est d'environ 7 pouces x 2,76 pouces x 2,76 pouces. Le poids de ces caméras peut atteindre 0,7 Kg. En fonction de sa résolution et de ses accessoires, le prix d'une caméra infrarouge / thermique varie de 300 \$ et 12K \$

- **Caméra d'évènement** : Les caméras d'évènements ont des caractéristiques intrinsèques différentes des caméras standards. Ce type de caméra génère une trame d'évènements asynchrones à sa sortie au lieu d'une image d'intensité standard, ce qui rend les algorithmes de vision conventionnels moins adaptés pour eux. Par conséquent, certaines recherches ont été menées pour exploiter la nature asynchrone et la résolution temporelle élevée des évènements [60, 61].

La durée de vie des caméras d'évènements nouvellement lancées ne dépasse pas dix ans. En fonction de leur résolution et de leur latence, le prix de ces caméras varie de 3000 \$ à 5000 \$. La caméra d'évènement est très économe en énergie vu qu'elle ne capture pas les informations redondantes dans une scène par défaut. Généralement, ces caméras sont faciles à remplacer car elles sont très compactes (Par exemple, la caméra d'évènement mini-eDVS mesure 0,7 pouces x 0,7 pouces x 0,3 pouces).

### 3. Capteurs hybrides

On désigne par un capteur hybride, tout capteur intégrant à la fois un capteur télémétrique et un capteur de vision.

- **Caméra à temps de vol optique 3D** : La caméra à temps de vol optique 3D, est souvent utilisée pour détecter les obstacles à proximité. La manière d'obtenir la 3D est la même que pour la vision stéréo, mais une des deux caméras est remplacée par un laser. L'idée derrière de ces capteurs est d'illuminer l'environnement avec une lumière régulière puis de l'observer avec une caméra. Les caméras temps de vol optiques sont utilisées pour avoir des images de profondeur, pour éviter les obstacles, et pour la numérisation d'objet 3D. Connaissant la position du laser et celle de la caméra, ces informations peuvent être obtenues par la triangulation.

Vu que chaque pixel de la caméra fournit une mesure de distance, la caméra à temps de vol peut obtenir des informations en peu de temps sans scanner l'ensemble de la scène. Les caméras à temps de vol peuvent acquérir jusqu'à 50 images 3D par seconde. Contrairement à une caméra stéréo nécessitant des algorithmes puissants et une structure de référence sur l'objet pour extraire ses données 3D, une caméra à temps de vol ne nécessite aucun algorithme spécifique et ses mesures sont tout à fait indépendantes de la structure de référence d'objet. Il est à noter que ce type de caméra est encore très cher (Par exemple, les Swiss Rangers coûtent plus de 5000\$). Cependant, le nouveau capteur Kinect RVB-D de Microsoft qui coûte 150 \$ peut changer les règles du jeu.

- **Capteur RVB-D** : Le capteur RVB-D se compose d'une caméra standard, d'un émetteur et d'un récepteur infrarouge (IR). Il fournit des informations sur les couleurs de la scène et la profondeur estimée pour chaque pixel. Comme le capteur RVB-D présente les avantages à la fois d'un capteur de vision et d'un capteur de profondeur, il a suscité un grand intérêt pour la robotique [50]. Le capteur RVB-D le plus populaire est le Kinect RVB-D développé par Microsoft en novembre 2010.

En termes de durée de vie, les capteurs RVB-D ont montré des performances encourageantes au fil des ans. Selon le fabricant et les caractéristiques du capteur, le prix de RVB-D varie de 150 \$ à 400 \$. La portée maximale de ce capteur ne dépasse pas 5 m (par exemple, la portée du Kinect V2 est de 4,5 m, et celle de Kinect V1 et Asus Xtion ne dépasse pas 3,5 m).

#### I.1.4 Modélisation d'un environnement intérieur

Ces dernières années ont vu une demande accrue de modélisation de scènes d'intérieur [62] en raison de ses applications potentielles dans la planification, la construction, la cartographie et la navigation [63, 64, 65, 66]. Après une étude approfondie de la modélisation intérieure, nous pouvons constater qu'il existe en fait cinq modèles.

##### I.1.4.1 Grille d'occupation

Une grille d'occupation est une simple représentation de surface de l'environnement. Ce modèle discrétise l'environnement en fonction d'une grille régulière en cellules rectangulaires ou carrées de même taille. Chacune de ces cellules possède un indice désignant si elle représente un espace libre ou un espace occupé. Quant aux méthodes de construction des grilles d'occupation, elles sont assez économiques en termes de ressources informatiques et elles sont rapides et faciles à mettre à jour.

Le principal avantage des grilles d'occupation est qu'elles peuvent représenter l'espace de façon très dense, selon le pas de discrétisation de la grille. Elles conviennent à tout type d'environnement et fournissent des estimations probabilistes ou statistiques des données, ainsi que des informations sur l'occupation et donc sur la localisation des obstacles. Cependant, le principal inconvénient des grilles d'occupation est le manque de compacité. Une telle modélisation est plus adaptée pour représenter des environnements encombrés et semble moins efficace dans les grandes espaces vides.

FIGURE I.9 – Exemple de grille d’occupation : Les zones sombres indiquent une forte probabilité de présence d’un obstacle [67] (illustration retirée afin de respecter les droits d’auteurs)

#### I.1.4.2 Modélisation métrique

Le but de la modélisation métrique est de construire une carte (graphe) géométrique de l’environnement depuis les données acquises. Les informations telles que la distance, la longueur, la position, etc., sont clairement définies dans les cartes métriques et sont souvent définies dans un référentiel unique. Ces cartes sont généralement faciles à interpréter par les humains car elles fournissent une relation clairement définie avec le monde réel [68].

#### I.1.4.3 Modélisation topologique

Le but de la modélisation topologique est d’établir une carte abstraite montrant la relation entre l’ensemble d’éléments environnementaux, sans recourir à un référentiel absolu. Une carte topologique est caractérisée par des sommets qui correspondent à des positions généralement associées aux informations de perception que le robot peut reconnaître, et par des arêtes renseignant sur l’existence d’une piste, entre deux sommets, qui peut être parcourue par le robot. Le principal avantage de la modélisation topologique est qu’il n’y a pas d’incertitude liée au mouvement du robot.

FIGURE I.10 – Exemple de carte métrique et de carte topologique [69] (illustration retirée afin de respecter les droits d’auteurs)

La Figure I.10 illustre un exemple de modélisation métrique et de modélisation topologique correspondantes au même environnement intérieur. De cette figure, on peut constater que les pièces et les couloirs sont bien définis dans la carte topologique, alors que la carte métrique permet de présenter les murs.

#### I.1.4.4 Modélisation hybride

La modélisation hybride inclue l’utilisation simultanée de deux ou plusieurs types de carte afin de tirer parti d’avantages de chaque représentation. Par exemple, certains auteurs ont proposé une modélisation hybride métrique-topologique, dont le but est de combiner la précision de la carte métrique avec l’extensibilité de la carte topologique [70]. Comme Bailey a souligné [71], la topologie permet à l’environnement d’être subdivisé naturellement, réduisant ainsi l’espace de stockage et le temps de calcul, et assurant la connectivité et la cohérence à grande échelle. En contrepartie, la carte métrique conduit à une précision locale plus élevée, une quantification de l’incertitude et à une certaine généralité. Ceci explique le grand nombre de modèles

hybrides, dans la littérature, qui se basent sur la combinaison de la modélisation métrique et de la modélisation topologique.

FIGURE I.11 – Exemple de carte hybride [69] (illustration retirée afin de respecter les droits d'auteurs)

#### I.1.4.5 Modèle de primitives

Le modèle de primitives est défini par Guivant *et al.* [72] en tant que des parties spécifiques de l'environnement. Afin d'assurer la cohérence du graphe de l'environnement, le modèle de primitives doit prendre en considération l'incertitude liée aux primitives, qui sont souvent représentées par des distributions gaussiennes. En raison du nombre limité de primitives, les mises en correspondance sont généralement moins chères que pour une grille d'occupation. D'autre part, contrairement à la grille d'occupation, la représentation par primitives n'est généralement pas trop dense, et ne permet que la cartographie et le positionnement sans planification, et sans représentation des obstacles.

## I.2 Proposition d'un modèle générique d'interprétation de scène

Dans ce travail, le système de perception utilisé pour observer une scène réelle intérieure, est un système multi-capteurs équipé de capteurs télémétriques et de visions. Actuellement les modèles de perception de scènes sont plus ou moins complexes ce qui nous motive de proposer un modèle simple qui, malgré sa simplicité, permet une navigation autonome. Comme le montre la Figure I.12, le modèle d'interprétation de scène proposé consiste à partager le champ intercepté par les différents capteurs en trois zones majeures, à savoir : gauche, face et droite. La Figure I.13 présente quelques exemples de modélisation de scène.

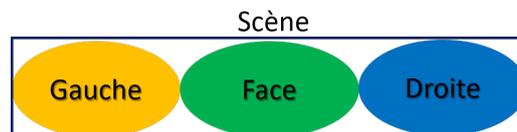


FIGURE I.12 – Modélisation d'un environnement intérieur

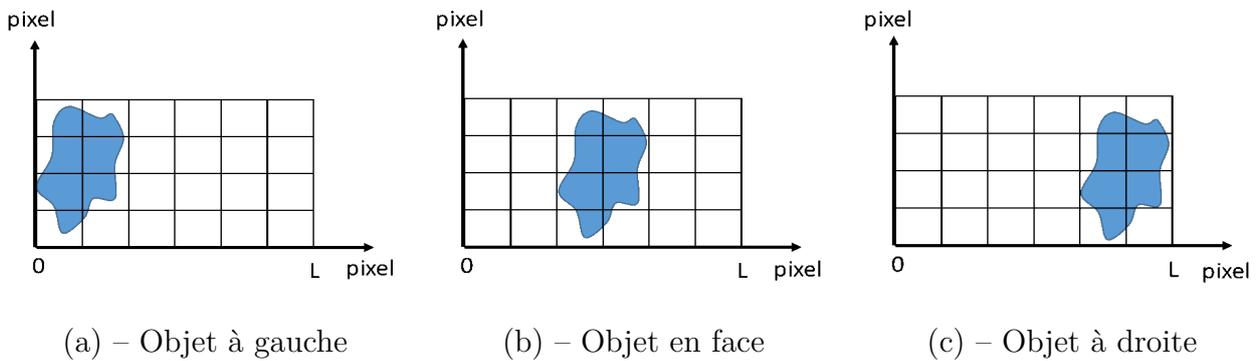


FIGURE I.13 – Exemples de scènes.

Pour une interprétation efficace de scène, toutes les connaissances fournies par les différents capteurs du système d'acquisition doivent être prises en compte. À cet effet, une étape de recalage matériel est nécessaire entre les images obtenues par les capteurs de vision et les signaux des capteurs télémétriques. Lors de recalage, plusieurs connaissances a priori doivent être prises en compte, à savoir, le positionnement des capteurs dans le système d'acquisition, la vitesse du balayage de l'utilisateur, la direction du balayage, les champs de vision des capteurs, etc.

Pour le système d'interprétation de scène proposé, les données de chaque capteur sont d'abord traitées et analysées séparément. Par conséquent, on obtient des informations spécifiques de chaque type de capteur conduisant à une interprétation conséquente de la scène, et l'attribution d'une ou plusieurs étiquettes décrivant chaque objet. Ensuite, les informations relatives aux capteurs télémétriques sont recalées avec les informations relatives aux capteurs de vision. Le recalage de ces différentes informations génère de nouvelles connaissances. Par exemple, certaines informations sont spécifiques à un capteur, telles que la distance qui est spécifique au capteur télémétrique et la forme qui est spécifique au capteur de vision. Suite au recalage, ces informations sont attribués conjointement aux objets de la scène. La Figure I.14 résume le scénario d'interprétation de scène proposé basé sur le recalage d'informations et la fusion de connaissances.

Il est à noter que la modélisation de la scène dépend à la fois du champ de vision du capteur d'imagerie définissant le largeur  $L$  de l'image correspondante à la scène, ainsi que la distance séparant la scène du capteur télémétrique. La Figure I.15 présente l'effet de la profondeur sur l'interprétation de la scène. En fait, pour une grande distance " $d_o$ " (Cf. Figure I.16(a)), le modèle de scène se compose de trois objets (objet à gauche, objet en face, et objet à droite), alors que pour une faible distance " $d'_o$ " (Cf. Figure I.16(b)), le modèle de scène se compose seulement d'un objet en face.

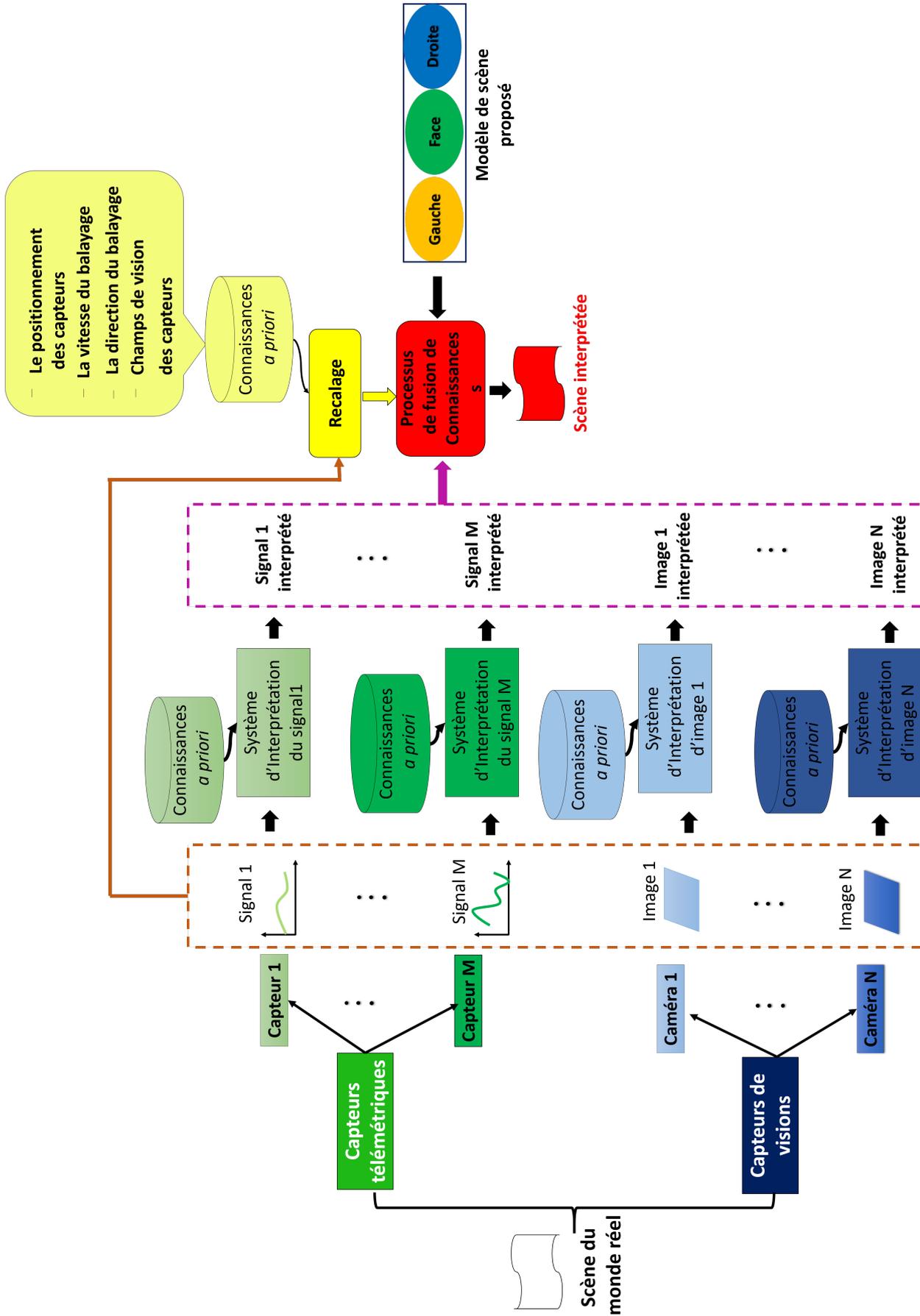


FIGURE I.14 – Système d'interprétation de scène par fusion de connaissance

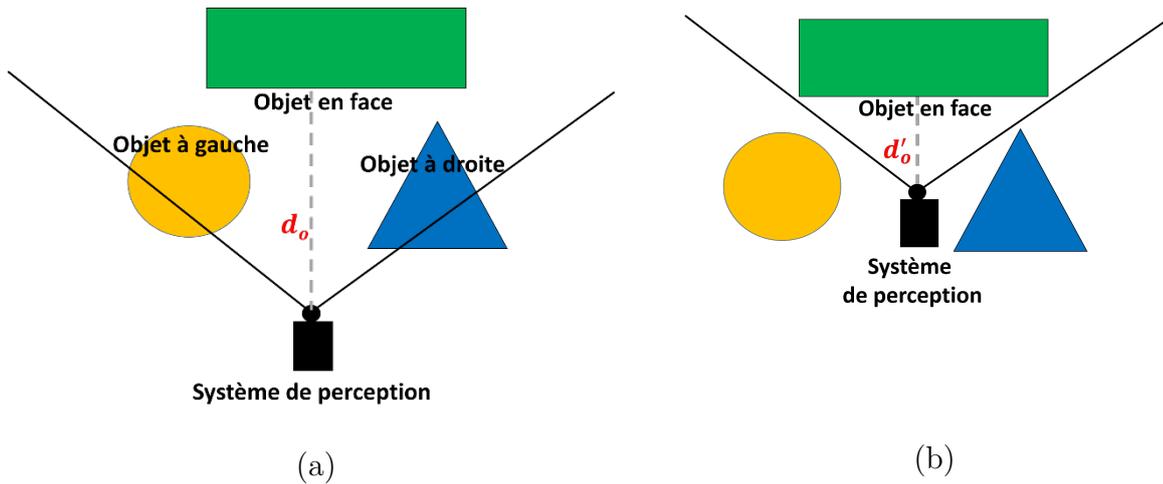


FIGURE I.15 – Effet de la profondeur sur l'interprétation de scène : (a) système de perception placé à une distance  $d_o$ , (b) système de perception placé à une distance  $d'_o < d_o$

L'algorithme I.1 présente le processus de positionnement des objets dans une scène. Il est à noter que ce processus repose sur 3 informations, à savoir :

- La distance " $d_o$ " séparant l'objet du système d'acquisition (information fournie par le capteur télémétrique).
- Le barycentre " $x_o$ " de l'objet détecté
- La largeur " $L_o$ " de l'objet détecté (information fournie par le capteur de vision).

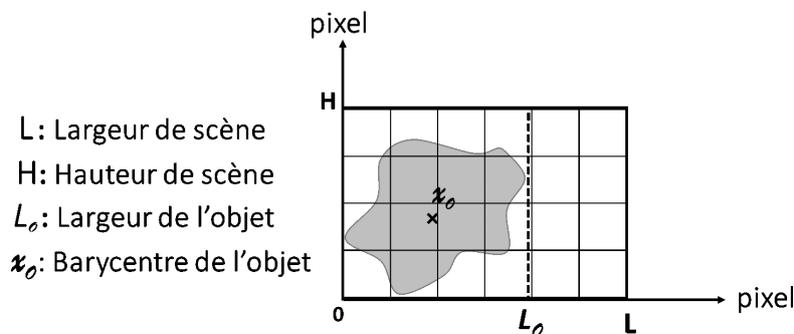


FIGURE I.16 – Caractérisation d'un objet à partir du capteur de vision

---

### Algorithme I.1 : Positionnement d'un objet dans une scène

---

**Entrées :**  $d_o$  (la distance séparant l'objet du système d'acquisition),  $x_o$  (le barycentre de l'objet détecté),  $L_o$  (la largeur de l'objet détecté), L (la largeur de l'image correspondante à la scène acquise),  $S_{dist}$  (seuil sur la distance choisi selon l'application),  $S_1$ ,  $S_2$  (choisis empiriquement)

**Sorties :** Objet à gauche, Objet en face, objet à droite.

Si  $d_o < S_{dist}$

Objet en face

Sinon Si  $x_o < S_1$  et  $L_o < S_1$

Objet à gauche

Sinon Si ( $S_1 < x_o < S_2$  et  $L_o < S_1$ ) ou  $L_o > S_1$

Objet en face

Sinon Si  $x_o > S_2$  et  $L_o < S_1$

Objet à droite

Fin Si

Fin Si

Fin Si

Fin Si

---

La Figure I.17 illustre quelques exemples de positionnement d'objets selon l'algorithme I.1, en choisissant  $S_1$  et  $S_2$  égaux à  $\frac{L}{2 \cdot \text{Nombre d'or}}$  et  $\frac{L}{\text{Nombre d'or}}$  respectivement, avec Nombre d'or=1.61803398875.

## I.3 Cadre théorique pour le développement du système d'interprétation proposé

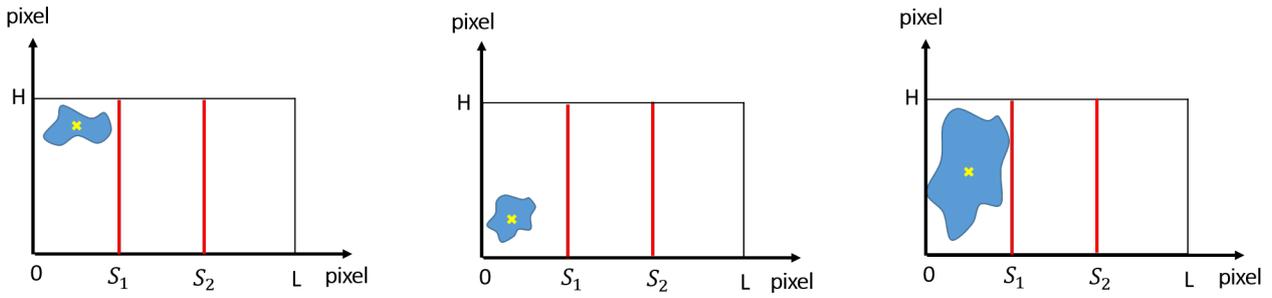
### I.3.1 Concepts fondamentaux

#### I.3.1.1 Donnée, Information, Connaissance

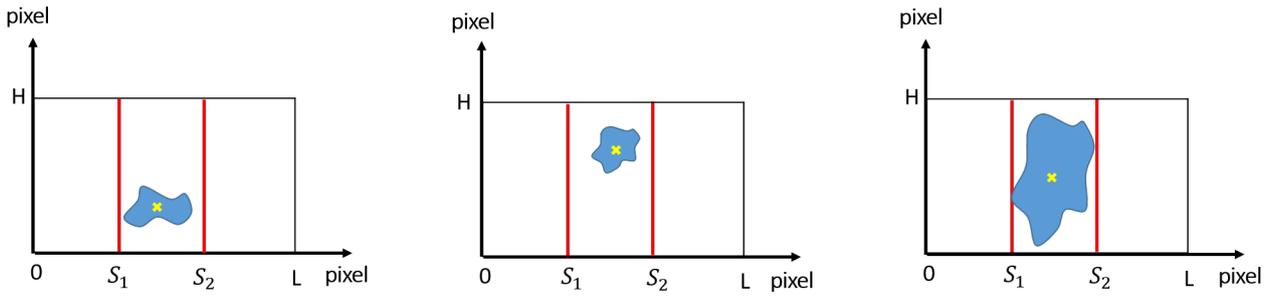
Dans le traitement d'information, il existe un chevauchement conceptuel entre les trois termes suivants : *donnée*, *information* et *connaissance* parce que leurs définitions sont inexactes et que les frontières entre eux sont floues et souvent subjectives. Dans la littérature, plusieurs définitions ont été proposées pour distinguer entre ces termes [73]. La différence entre ces termes incombe à leur niveau d'abstraction.

Une *donnée* est considérée comme le niveau d'abstraction le plus bas. En fait, tant que son espace de définition est inconnu, elle n'a aucun sens en soi. Une fois que le contexte, le modèle physique et le contenu informationnel sont connus, une *donnée* devient une *information*. Ainsi, *l'information* est de niveau d'abstraction plus élevé que de la donnée. La Figure I.19(a) et Figure I.19(b) présentent deux informations différentes obtenues à partir de mêmes données.

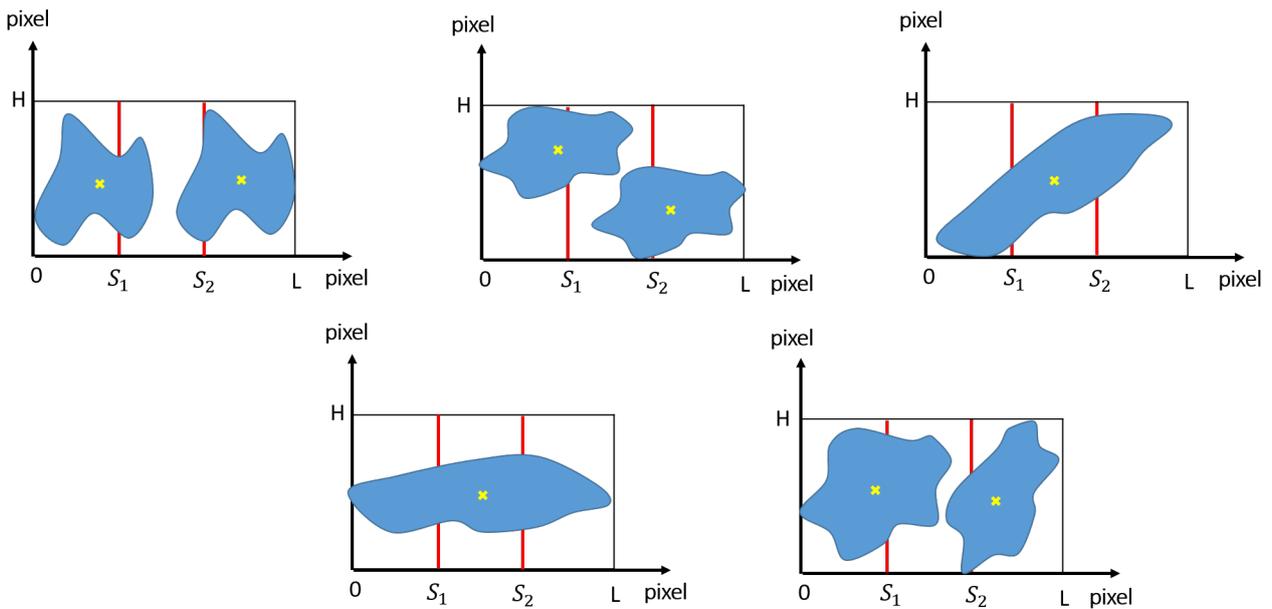
La *connaissance* est de niveau d'abstraction plus élevé que *l'information*. Elle peut être considérée comme le reflet de la compréhension d'un individu de *l'information* basée sur sa propre expérience, ses propres idées, etc.



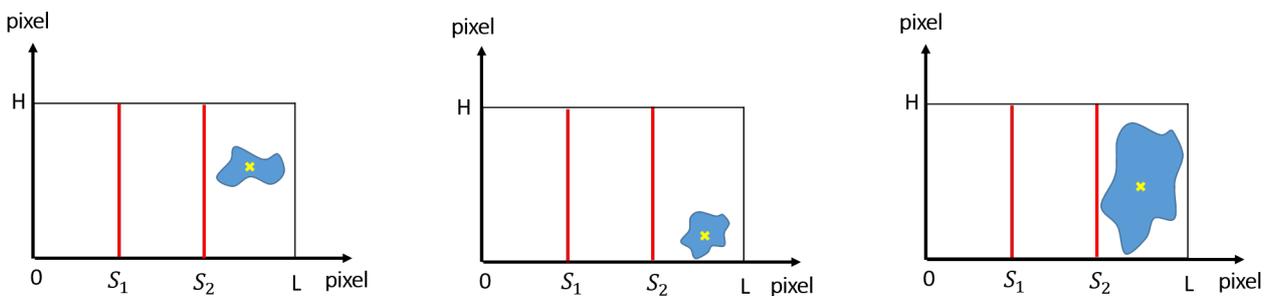
(a) – Objets considérés à gauche ( $x_o < S_1$  et  $L_o < S_1$ )



(b) – Objets considérés en face ( $S_1 < x_o < S_2$  et  $L_o < S_1$ )



(c) – Objets considérés en face ( $L_o > S_1$ )



(d) – Objets considérés à droite ( $x_o > S_2$  et  $L_o < S_1$ )

FIGURE I.17 – Positionnement des objets dans une scène

FIGURE I.18 – Obtention d'informations distinctes à partir des mêmes données confrontées à différents contextes [7] (illustration retirée afin de respecter les droits d'auteurs)

### I.3.1.2 Les imperfections dans l'information

Plusieurs formes d'imperfections peuvent affecter les informations. Dans la suite de cette section, ces différentes formes d'imperfection sont brièvement présentées.

1. **L'ambiguïté** : Si plusieurs contenus d'information peuvent apparaître en même temps, ou si deux interprétations ou plus peuvent être faites, alors l'information considérée est ambiguë.
2. **L'incomplétude** : L'incomplétude représente le manque d'informations fournies par la source. Elle peut être mesurée par la différence entre la quantité d'informations que la source fournit réellement et la quantité d'informations que la source devrait fournir.
3. **L'incertitude** : L'incertitude peut être dû à un manque d'informations (par exemple, un manque de connaissances sur les effets secondaires d'une maladie), ou à des informations contradictoires (par exemple, des opinions différentes des témoignages d'un accident) et cela se produit surtout lorsque des phénomènes aléatoires (tels que lancer un dé) ne peuvent pas être prédits pour chaque résultat expérimental. Il existe deux types d'incertitude [7] :
  - Incertitude aléatoire : elle est liée à des phénomènes aléatoires répétables ou à l'étude des caractéristiques au sein d'une population
  - Incertitude épistémique : elle est causée par un manque de connaissances.
4. **L'imprécision** : Lorsque nous ne connaissons pas la valeur exacte des variables, attributs, etc., afin de décrire clairement le contenu de l'information, nous parlons alors d'informations imprécises.
5. **Le conflit** : On parle de conflit dans des situations où l'information a des interprétations contradictoires et donc incompatibles.

## I.3.2 Choix du cadre théorique

Une bonne modélisation contribue à l'amélioration des performances et de la fiabilité du système d'interprétation de scène. En fait, elle adopte un mode de raisonnement qui prend en compte les imperfections de l'information, ce qui est obtenu en transférant les informations vers des représentations mathématiques au moyen d'une méthode ou d'une théorie spécifique. Dans

la Section IV du Chapitre 1 de [7], une revue sur les théories de modélisation d'informations imparfaites a été présentée. Le choix de l'une de ces théories dépend fortement de l'application et de sa complexité, des connaissances disponibles et du type d'imperfection [7]. Par exemple, lorsqu'il y a un grand nombre d'échantillons et qu'on souhaite modéliser uniquement des imperfections de type incertitude aléatoire, la théorie des probabilités est considérée comme un cadre théorique très approprié. Cependant, si la taille de l'échantillon est limitée et que d'autres types d'imperfections doivent être pris en compte, la théorie des probabilités n'est pas le bon choix.

Pour les systèmes d'interprétation de scène, l'imperfection des données est due à deux raisons : (i) les propriétés physiques des capteurs utilisés, (ii) la brièveté du signal représentatif des données télémétriques. Ces deux imperfections sont liées au capteur et à la représentation de l'information, et font donc partie de l'incertitude épistémique. Dans ce contexte, la théorie de possibilités est le cadre le plus approprié pour gérer ces imperfections. Pour cette raison, la théorie de possibilités a été adoptée comme un cadre théorique pour le développement du système d'interprétation proposé.

### I.3.3 La théorie de possibilités

La théorie de possibilités a été introduite par Zadeh en 1978 comme une extension de la théorie des ensembles flous [74]. En fait, tandis que la théorie des ensembles flous assouplit le cadre de la théorie classique des ensembles pour pouvoir traiter l'imprécision, la théorie de possibilités est destinée à gérer les incertitudes épistémiques. La théorie de possibilités a ensuite été développée par Dubois et Prade [75].

Dans ce qui suit, nous présentons les outils de base de cette théorie, à savoir la distribution de possibilités, les mesure de possibilité et de nécessité, les mesures d'incertitude, les mesures de similarité, les différents opérateurs de fusion d'informations et les critères de décision possibilistes.

Dans la suite de cette section, nous désignons :

- $\Omega = \{x_1, \dots, x_n, \dots, x_N\}$  est l'univers du discours
- L'ensemble de puissance  $2^\Omega$  est l'ensemble de tous les sous-ensembles de  $\Omega$ , y compris l'ensemble vide  $\emptyset$  et l'univers du discours  $\Omega$ .
- $A$  est un évènement ( $A \in 2^\Omega$  ou  $A \subseteq \Omega$ ).
- $A^c$  représente l'évènement complémentaire de  $A$

### I.3.3.1 Distribution de possibilités

La théorie de possibilités est basée sur le concept de distribution de possibilités, notée  $\pi(\cdot)$  qui attribue à chaque singleton  $x_n$  de  $\Omega$ , une valeur dans l'intervalle  $[0,1]$ . La valeur  $\pi(x_n)$  représente le degré de possibilité pour que le singleton  $x_n$  soit le seul singleton qui se produit.

$$\begin{aligned} \pi : \Omega &\longrightarrow [0, 1] \\ x_n &\longrightarrow \pi(x_n) \end{aligned} \tag{I.1}$$

Tenant compte de cette définition, deux cas extrêmes de connaissances peuvent être définis :

- Connaissance complète :  $\exists!x_n \in \Omega, \pi(x_n) = 1$  et  $\pi(x_m) = 0, \forall x_m \in \Omega, x_m \neq x_n$
- Ignorance totale :  $\forall x_n \in \Omega, \pi(x_n) = 1$  (tous les singletons de  $\Omega$  sont tout à fait possibles)

L'estimation de la distribution de possibilités est l'une des questions les plus discutées en théorie de possibilités [76]. Dans [7], les techniques d'estimation de distributions de possibilités sont regroupées en deux catégories : (i) techniques guidées par les connaissances d'expert pour représenter de manière directe, ou indirecte, la forme générale de distributions de possibilités, et (ii) techniques basées sur des fonctions caractéristiques d'autres théories de l'incertain pour inférer la distribution de possibilités, à savoir : la distribution de probabilités, la fonction de masse. Dans nos travaux, nous accordons une attention particulière à la technique d'estimation de distributions de possibilités basée sur une transformation des distributions de probabilités [77, 78, 79, 80, 81, 82, 83, 84]. Il est à noter que cette conversion représente une cohérence conceptuelle, vu que les distributions de probabilités et de possibilités s'appliquent aux informations qui sont entachées d'incertitude.

Une distribution  $\pi$  est caractérisée par sa hauteur  $h(\pi)$  qui est considérée comme sa valeur maximale. Cette valeur décrit le degré de consistance de réalisation des singletons, c'est-à-dire qu'elle détermine le ou les singletons ayant la compatibilité la plus élevée des connaissances disponibles :

$$h(\pi) = 1 - \max_{x \in \Omega} (\pi(x)) \tag{I.2}$$

Si la hauteur de la distribution de possibilités est inférieure à 1, la distribution de possibilités est considérée comme étant anormale ou inconsistante. Pour illustrer cette situation, on définit le degré d'inconsistance d'une distribution de possibilités par :

$$Inc(\pi) = 1 - h(\pi) \tag{I.3}$$

L'inconsistance et la hauteur sont deux connaissances qui peuvent être déduites d'une distribution de possibilités comme illustrée à la Figure I.19.

FIGURE I.19 – Les connaissances véhiculées par une distribution de possibilités [7]  
(illustration retirée afin de respecter les droits d'auteurs)

### I.3.3.2 Mesures de possibilité et de nécessité

Contrairement à la théorie des probabilités et à la théorie des ensembles flous, la théorie de possibilités repose sur deux mesures : la mesure de possibilité  $\Pi$ , et la mesure de nécessité  $N$  définies respectivement par l'Équation I.4 et l'Équation I.5, respectivement.

$$\Pi(A) = \max_{x_n \in A} (\pi(x_n)) \quad (\text{I.4})$$

$$N(A) = 1 - \Pi(A^c) = \min_{x_n \notin A} (1 - \pi(x_n)) \quad (\text{I.5})$$

À partir des connaissances représentées par la distribution de possibilités  $\pi$ , la mesure de possibilité  $\Pi(A)$  permet d'estimer le niveau de cohérence de la réalisation de l'événement  $A$ . En fait, elle représente le degré de possibilité que le seul singleton  $x_n$  de  $A$  se soit produit. Le simple fait d'utiliser la mesure de la possibilité ne suffit pas, car elle ne peut ni distinguer l'ignorance totale ( $\Pi(A) = 1$  et  $\Pi(A^c) = 1$ ), ni distinguer la certitude totale ( $\Pi(A) = 1$  et  $\Pi(A^c) = 0$ ) d'un événement  $A$ . Par conséquent, il est nécessaire de la compléter par une mesure de nécessité que l'événement soit réalisé [85].

À partir des connaissances représentées par la distribution de possibilités  $\pi$ , la mesure de nécessité  $N(A)$  permet de représenter le niveau de certitude de la réalisation de l'événement  $A$ . En fait, elle représente le degré de certitude que le seul singleton  $x_n$  de  $A$  se soit produit. Si  $N(A) = 1$ , cela signifie que l'occurrence de l'événement  $A$  est considérée comme s'être totalement produite, en contrepartie si  $N(A) = 0$ , cela signifie que l'occurrence de l'événement  $A$  est considérée comme incertaine de s'être produite.

Compte tenu de la mesure de la possibilité et de la nécessité, il peut y avoir trois situations extrêmes [7] :

- Événement impossible :  $[N(A), \Pi(A)] = [0, 0]$
- Événement possible et certain :  $[N(A), \Pi(A)] = [1, 1]$
- Événement possible mais incertain :  $[N(A), \Pi(A)] = [0, 1]$

Il convient de noter qu'un degré de possibilité est la borne supérieure d'un degré de probabilité, et le degré de nécessité est sa borne inférieure, étant donné qu'un événement certain doit être probable et celui probable doit être possible [86].

$$N(A) \leq P(A) \leq \Pi(A) \quad (\text{I.6})$$

### I.3.3.3 Mesures d'incertitude

Dans la littérature, plusieurs mesures d'incertitude ont été proposées afin d'évaluer la quantité d'information contenue dans une distribution de possibilités. Dans ce qui suit, nous nous intéressons aux mesures suivantes : la Spécificité  $S_p$  [87] et la mesure d'incertitude  $U$  [88].

Considérons un univers de discours composé de  $N$  singletons  $\Omega = \{x_1, \dots, x_n, \dots, x_N\}$  et une distribution de possibilités  $\pi : \Omega \rightarrow [0, 1]$  est arrangée par un ordre décroissant, telle que :  $\pi(x_1) \geq \pi(x_2) \geq \dots \geq \pi(x_N)$ .

#### 1. Mesure d'incertitude "U"

La mesure d'incertitude "U" est définie par [88] :

$$U(\pi) = [\sum_{i=1}^N (\pi_i + \pi_{i+1}) \log_2(i)] + (1 - \pi_1) \log_2(N) \quad (\text{I.7})$$

On remarque que la plage de  $U$  est  $[0, \log_2(N)]$ . Tenant compte de cette remarque, deux cas extrêmes peuvent être définis :

- Connaissance complète :  $U(\pi) = 0$
- Ignorance totale :  $U(\pi) = \log_2(N)$

#### 2. Mesure de spécificité de Yager " $S_p$ "

La mesure de spécificité " $S_p$ " [87] définie par l'Équation I.8 vérifiant les propriétés suivantes :

- $S_p(\pi) = 1$  si et seulement si  $\pi(x_o) = 1$ , et  $\pi(x) = 0$  pour tout  $x \in \Omega$  ( $x \neq x_o$ )
- $S_p(\pi_\Phi) = 0$ ;  $\pi_\Phi(x) = 0$  pour tout  $x \in \Omega$
- Si  $\pi_1 \leq \pi_2$  alors  $S_p(\pi_1) \leq S_p(\pi_2)$

$$S_p(\pi) = \pi_1 - \sum_{i=2}^N \omega_i \pi_i \quad (\text{I.8})$$

Avec  $\omega_i$  est un ensemble de poids satisfaisant les propriétés suivantes :

- $\omega_i \in [0, 1]$
- $\sum_{i=2}^N \omega_i = 1$
- $\omega_i \geq \omega_j$ , pour tout  $1 < j < i$

### I.3.3.4 Mesures de similarité

Dans de nombreux domaines de recherche, tels que l'apprentissage automatique, la science des données, la reconnaissance de formes, etc., les mesures de similarité ont suscité un grand

intérêt. En particulier, dans la théorie de possibilités, afin de comparer des informations incertaines, il est nécessaire de comparer leurs distributions de possibilités. Dans cette section, nous étudions les mesures de similarité les plus connues dans la littérature.

Dans la suite, on désigne par  $\pi_1$  et  $\pi_2$  deux distributions de possibilités appartenant au même univers du discours  $\Omega = \{x_1, \dots, x_N\}$

### 1. Distance de Minkowski et ses dérivés

Afin d'évaluer la similarité entre  $\pi_1$  et  $\pi_2$ , la métrique de distance point à point entre ces deux distributions est utilisée.

— La distance de Minkowski :

$$D_M(\pi_1, \pi_2) = \sqrt[p]{\sum_{i=1}^N |\pi_1(x_i) - \pi_2(x_i)|^p} \quad (\text{I.9})$$

— La similarité basée sur la distance de Manhattan normalisée :

$$S_M(\pi_1, \pi_2) = 1 - \frac{\sum_{i=1}^N |\pi_1(x_i) - \pi_2(x_i)|}{N} \quad (\text{I.10})$$

— La similarité basée sur la distance Euclidienne normalisée :

$$S_E(\pi_1, \pi_2) = 1 - \sqrt{\frac{\sum_{i=1}^N (\pi_1(x_i) - \pi_2(x_i))^2}{N}} \quad (\text{I.11})$$

— La similarité basée sur la distance Chebychev normalisée

$$S_C(\pi_1, \pi_2) = 1 - \max_{i=1}^N (\pi_1(x_i) - \pi_2(x_i)) \quad (\text{I.12})$$

### 2. Mesure de similarité "Info Closeness"

La mesure proposée par Higashi et Klir [89] est une des premières oeuvres visant à évaluer la similarité entre des distributions de possibilités. La similarité entre  $\pi_1$  et  $\pi_2$ , notée  $S_G$ , est définie comme suit :

$$S_G(\pi_1, \pi_2) = 1 - \frac{G(\pi_1, \pi_2) - G_{\min}}{G_{\max}} \quad (\text{I.13})$$

Avec :  $G_{\min} = 0$  ;  $G_{\max} = 2 * \log_2(|\Omega|) - \log_2(|\Omega| - 1)$  et  $G$  est défini comme suit :

$$G(\pi_1, \pi_2) = 2 * U(\pi_1 \vee \pi_2) - U(\pi_1) - U(\pi_2) \quad (\text{I.14})$$

Où :  $\vee$  est l'opérateur maximum et  $U$  est la mesure d'incertitude

### 3. Mesure de similarité "Info Affinity"

La mesure de similarité "Info Affinity" [90] est une mesure de distance combinée avec la valeur de l'inconsistance des deux distributions  $\pi_1$  et  $\pi_2$ .

$$\text{InfoAffinity}(\pi_1, \pi_2) = 1 - \frac{d(\pi_1, \pi_2) - \text{Inc}(\pi_1 \wedge \pi_2)}{2} \quad (\text{I.15})$$

Avec :

$d$  : représente une distance entre deux distributions (par exemple la distance de Manhattan).

*Inc* : le degré d'inconsistance

$$Inc(\pi_1 \wedge \pi_2) = 1 - \max_{x_n \in A} (\pi_1 \wedge \pi_2(x_n)) = 1 - h(\pi_1 \wedge \pi_2) \quad (I.16)$$

#### 4. Mesure de similarité " Info Specificity "

La mesure de similarité "Info specificity" [91] est basée sur deux mesures, l'une basée sur une métrique de distance (distance de Manhattan normalisée) qui compare la précision de l'information à travers d'un schéma "point à point", et l'autre sur une mesure de spécificité " $S_P$ ", définie par [92], qui permet de quantifier les changements dans la quantité d'informations. La mesure de similarité Info specificity est définie comme suit :

$$InfoSpecificity(\pi_1, \pi_2) = \begin{cases} 1 - \frac{\alpha * dist_{Manh}(\pi_1, \pi_2) + \beta * dist_{Sp}(\pi_1, \pi_2)}{\alpha + \beta} & Si \ dist_{Manh} \neq 1 \\ 0 & Si \ dist_{Manh} = 1 \end{cases} \quad (I.17)$$

Avec :

$dist_{Manh}$  : est la distance de Manhattan normalisée entre les distributions  $\pi_1$  et  $\pi_2$

$dist_{Sp}$  : est la différence entre les spécificités des deux distributions de possibilité, en utilisant la mesure de spécificité proposée par Yager [87], tel que :  $dist_{Sp} = |S_p(\pi_1) - S_p(\pi_2)|$

$\alpha$  et  $\beta$  : sont deux coefficients, tel que  $\alpha \in [0,1]$  et  $\beta \in [0,1]$

#### I.3.3.5 Fusion d'informations possibilistes

La théorie de possibilités offre une variété d'opérateurs pour fusionner les informations provenant de plusieurs sources possibles. Soient  $\pi_1$  et  $\pi_2$  deux distributions de possibilités représentant des informations fournies par deux sources  $S_1$  et  $S_2$ , respectivement. Dans le reste de cette section, trois modes de fusion de ces deux distributions de possibilités sont présentées.

##### 1. Comportement conjonctif

Les opérateurs de fusion conjonctifs sont généralement utilisés pour fusionner des sources concordantes (c'est-à-dire toutes les sources considérées comme fiables). Il convient de noter que l'effet de ces opérateurs est de réduire l'incertitude globale et de fournir des résultats plus précis que de considérer chaque source séparément. La Figure I.20 illustre un exemple de fusion conjonctive utilisant l'opérateur "min".

FIGURE I.20 – Fusion conjonctive par l'opérateur min (Fig.2.4 de [6]) (illustration retirée afin de respecter les droits d'auteurs)

La fusion conjonctive, notée  $\pi_\wedge$ , de  $\pi_1$  et  $\pi_2$ , est définie par Dubois [81] :

$$\forall x \in \Omega, \pi_\wedge(x) = \pi_1(x) \wedge \pi_2(x) \quad (I.18)$$

L'opérateur conjonctif peut prendre différentes formes :

- Minimum :  $\pi_{\wedge}(x) = \min(\pi_1(x), \pi_2(x))$
- Produit :  $\pi_{\wedge}(x) = \pi_1(x) \cdot \pi_2(x)$
- Lukasiewicz T-norme :  $\pi_{\wedge}(x) = \max(0, \pi_1(x) + \pi_2(x) - 1)$
- etc.

## 2. Comportement disjonctif

Les opérateurs de fusion disjonctifs sont généralement utilisés pour fusionner des sources discordantes (c'est-à-dire que les sources sont partiellement ou totalement en conflit). Il convient de noter que l'effet de ces opérateurs est d'augmenter l'incertitude globale et de fournir des résultats fiables mais moins précis que de considérer chaque source séparément. Cela réduit, sans aucun doute, sa praticité. La Figure I.21 illustre un exemple de fusion disjonctive utilisant l'opérateur "max".

La fusion disjonctive, notée  $\pi_{\vee}$ , de  $\pi_1$  et  $\pi_2$ , est définie par Dubois [81] :

$$\forall x \in \Omega, \pi_{\vee}(x) = \pi_1(x) \vee \pi_2(x) \quad (\text{I.19})$$

L'opérateur disjonctif peut prendre différentes formes :

- Maximum :  $\pi_{\vee}(x) = \max(\pi_1(x), \pi_2(x))$
- Somme :  $\pi_{\vee}(x) = \pi_1(x) + \pi_2(x)$
- Lukasiewicz T-norme :  $\pi_{\vee}(x) = \min(\pi_1(x) + \pi_2(x), 1)$
- etc.

FIGURE I.21 – Fusion disjonctive par l'opérateur max (Fig.2.3 de [6])(illustration retirée afin de respecter les droits d'auteurs)

## 3. Comportement de compromis

S'il n'y a pas de consensus sur la fiabilité des différentes sources (c'est-à-dire que les sources sont partiellement en conflit), un opérateur adaptatif (de compromis) est appliqué. La fusion adaptative est utilisée pour conserver et renforcer les informations de consensus, et pour conserver les informations conflictuelles entre les sources d'informations à une plus petite échelle (Cf. Figure 3). Dans [81], Dubois et Prade ont présenté un exemple d'opérateur adaptatif.

### I.3.3.6 Critères de décision dans un cadre possibiliste

Dans le cadre possibiliste, la prise de décision réside dans l'identification du singleton  $x_n$  le plus vraisemblable à un échantillon de test T en fonction de sa distribution de possibilités  $\pi(T)$ .

Bien que la théorie de possibilités soit utilisée pour traiter l'incertitude et que son but ultime soit de déterminer le contenu informationnel, il est surprenant de constater qu'il y ait très peu de travaux dans la littérature qui analysent la problématique de la prise de décision dans un tel contexte [6]. Dans ce qui suit, nous présentons, un résumé de ces travaux.

### 1. Mesure de possibilité maximale

Dans cette règle de décision, l'alternative C qui fournit une mesure de possibilité la plus élevée sera considéré comme le plus vraisemblable à l'échantillon T.

### 2. Mesure de nécessité maximale

Cette règle de décision est basée sur la mesure de nécessité la plus élevée. En fait, l'alternative C qui fournit une mesure de nécessité la plus élevée est considérée comme la plus vraisemblable à l'échantillon T. Dans [93], Tversky a montré une relation d'équivalence entre cette règle de décision et celle basée sur la possibilité maximale. Ainsi, cette règle n'apporte aucun autre avantage par rapport à la première règle de décision.

### 3. Indice de confiance maximal

Sur la base de la mesure de l'incertitude, d'autres règles de prise de décision possibilistes ont été basées sur l'indice de confiance [94]. Cet indice  $Ind(C|T)$  est considéré comme une combinaison de la mesure de possibilité  $\Pi(\cdot)$  et de la mesure de nécessité  $N(\cdot)$ . Il est défini comme suit :

$$Ind : 2^\Omega \longrightarrow [-1, +1] \quad (I.20)$$

$$Ind(C|T) = \Pi(C|T) + N(C|T) - 1$$

Tenant compte de cette définition, deux cas extrêmes peuvent être définis :

- Réalisation complètement invalide de T :  $Ind(C|T) = -1$ , si  $N(C|T) = \Pi(C|T) = 0$
- Réalisation complètement valide de T :  $Ind(C|T) = +1$ , si  $N(C|T) = \Pi(C|T) = 1$

### 4. Moyenne maximale

Dans [95], une règle de décision basée sur la moyenne de la mesure de possibilité et de nécessité a été proposée. Cette moyenne, notée  $\psi$ , est définie par l'Équation I.21. Dans le cadre de cette règle, l'alternative C qui fournit une  $\psi$  la plus élevée est considérée comme la plus vraisemblable à l'échantillon T.

$$\psi(C|T) = \frac{\Pi(C|T) + N(C|T)}{2} \quad (I.21)$$

### 5. Mesure de similarité maximale

Cette règle de décision est basée sur le fait que l'alternative C la plus similaire à l'échantillon de test T sera considérée comme la plus vraisemblable.

## I.4 Conclusion

Diverses sont les systèmes existants dans le domaine de navigation dans un environnement intérieur. Le choix du (ou des) capteur(s) doit prendre en compte plusieurs critères tel que son domaine de fonctionnement, son coût, sa portée, le type de scène, les objets à détecter, etc. De plus, le choix du modèle, selon lequel le robot va interpréter son environnement, est aussi important que le choix des capteurs. Comme les modèles proposés dans la littérature sont plus ou moins complexes, nous avons proposé un modèle simple pour interpréter la scène. En fait, il considère qu'une scène ne se compose que de trois parties (gauche, face et droite).

Dans ce chapitre, nous avons également présenté les concepts de base liés à l'information ainsi que les différents types d'imperfections capables de l'affecter. Compte tenu de ces imperfections au cours du processus d'interprétation, les informations doivent être bien représentées selon un modèle. En fait, la modélisation est une étape primordiale du traitement de l'information afin d'obtenir des systèmes d'interprétation fiables et performants à travers une transformation de la représentation de données radiométriques issues de sources physiques en une représentation mathématique. À l'issue de nos recherches sur les différentes théories connues pour leur capacité et leur efficacité à modéliser des données imparfaites, on peut conclure que le choix de l'une de ces théories dépend fortement du cadre de l'application et notamment de l'aspect qualitatif et quantitatif des données. Notre intérêt a été particulièrement porté sur la théorie de possibilités, car elle montre une meilleure capacité à représenter et à traiter des informations imprécises, ambiguës et contradictoires provenant de différentes sources de connaissances. Comme la distribution de possibilités étant le modèle mathématique utilisé par cette théorie, pour traiter l'incertitude de l'information, une attention particulière a été accordée aux outils qui peuvent leur être appliqués.

# Chapitre II

## Développement d'outils de traitement possibilistes des données acquises pour la conception du système de navigation

### Sommaire

---

II.1 Proposition d'approches pour la sélection d'attributs dans un cadre possibiliste . . . . .	37
II.2 Proposition de classifieurs binaires hiérarchiques . . . . .	49
II.3 Évaluation du CBH(1) et CBH(2) . . . . .	58
II.4 Conclusion . . . . .	66

---

## Introduction

Les applications de classification multi-classes sont généralement définies par une quantité massive de données qui peuvent être incertaines, imprécises ou même manquantes en raison de sources d'informations peu fiables, de bruit, etc. Ces formes d'imperfections font partie de l'incertitude épistémique. Il a été prouvé, dans le chapitre précédent, que le cadre possibiliste est bien approprié pour la modélisation de ce type de données. Généralement, dans la phase de modélisation, on fait appel à la caractérisation des données par usage de primitives (attributs). Il faut souligner que l'utilisation d'un nombre important d'attributs augmente la dimensionnalité, augmentant ainsi la complexité du système. Par exemple, dans [96], les auteurs ont utilisé 4 millions d'attributs dans leur système. Certains d'entre eux sont redondants ou inutiles. Par conséquent, ils peuvent être des sources de divers bruits pouvant entraîner une dégradation des performances du système.

Dans la première partie de ce chapitre, trois approches de sélection d'attributs dans un cadre possibiliste ont été proposées. La sélection est basée sur la détermination du pouvoir discriminant des attributs. Pour la première approche, la définition du pouvoir discriminant, notée "DP\_SUMF", intègre des informations de niveau perceptif (la variance) et de niveau sub-sémantique (la mesure de spécificité d'une distribution de possibilités et la dissimilarité entre les distributions de possibilités). Pour les deux autres approches de sélection d'attributs, les pouvoirs discriminants ne sont basés que sur des informations de niveau sub-sémantique, à savoir : la mesure de spécificité de la distribution de possibilités, l'aire du consensus des distributions de possibilités et la hauteur maximale des points d'intersection des distributions de possibilités.

Dans la deuxième partie ce chapitre, deux classifieurs binaires hiérarchiques l'un monocritère l'autre multicritère sont développés pour : (i) simplifier le problème d'entrée à haute dimension pour la classification multi-classes et (ii) faire face à l'incertitude liée aux mesures d'attributs de l'ensemble de données en appliquant une des approches de sélection d'attributs proposée dans la première partie de ce chapitre. Enfin, les deux classifieurs binaires hiérarchiques conçus sont évalués et comparés à d'autres classifieurs conventionnels, suite à leur application à différentes bases de données de référence (Benchmarks).

## II.1 Proposition d'approches pour la sélection d'attributs dans un cadre possibiliste

La sélection d'attributs devient une étape de plus en plus importante dans les systèmes de prise de décision. Son objectif principal est de réduire la dimensionnalité du système, réduisant ainsi sa complexité. Dans [97], les auteurs présentent différentes approches de sélection d'attributs, telles que le modèle en régression linéaire (linear forward selection LFS), l'analyse en composantes principales (PCA), Relief, le score de Fisher, etc. Certaines de ces approches fonctionnent bien dans certaines situations, mais échouent dans d'autres situations car ils ne prennent pas en compte (i) l'incertitude liée aux attributs, et (ii) que les attributs n'ont pas la même importance envers la différenciation entre les classes. En fait, certains d'entre eux sont utiles pour la séparation de classe, tandis que d'autres sont plus fiables pour la représentation de classes. Ainsi, il est intéressant de développer d'approches de sélection d'attributs en se basant sur le calcul des pouvoirs discriminants d'attributs afin de surpasser les performances des algorithmes de sélection d'attributs traditionnels.

### II.1.1 App 1 : Estimation du pouvoir discriminant par la fusion de mesures statistiques et mesures d'incertitude (DP\_SUMF : Discriminant power based on statistics and uncertainty measure fusion)

Dans cette section, une nouvelle approche basée sur le calcul du pouvoir discriminant "DP\_SUMF" est proposée pour sélectionner les attributs les plus pertinents dans un cadre possibiliste, où l'incertitude est représentée par des modèles possibilistes. La détermination du pouvoir "DP\_SUMF" est produite par la fusion de mesures statistiques et mesures d'incertitude normalisées, à savoir : le pouvoir discriminant basé sur la mesure de variance " $d_{var}$ ", le pouvoir discriminant basé sur la mesure de spécificité de Yager " $d_{sp}$ ", et la dissimilarité " $d_{si}$ " entre les N distributions de possibilités de l'univers du discours  $\Omega = \{C_1, C_2, \dots, C_N\}$ . Il est à préciser que le pouvoir "DP\_SUMF" varie entre [0, 3].

#### — Pouvoir discriminant basé sur la mesure de variance " $d_{var}$ " :

Comme le montre l'étape 1 de l'algorithme II.1, ce pouvoir est défini comme le rapport entre la variance inter-classe " $var_{inter}$ " entre les N classes de l'univers de discours  $\Omega$  et la somme de la variance inter-classe et intra-classe " $var_{intra}$ ". En fait, la variance inter-classe est utilisée pour déterminer la capacité d'un attribut à séparer les classes, alors

que la variance intra-classe est utilisée pour déterminer la capacité d'un attribut à les représenter. Il est à noter que plus la variance inter-classe est élevée plus le " $d_{var}$ " est élevé. Il est à noter que dans l'algorithme II.1, on désigne par  $\bar{x}$  la valeur moyenne de  $x$ .

— **Pouvoir discriminant basé sur la mesure de spécificité de Yager " $d_{sp}$ " :**

L'approche de sélection d'attributs proposée est principalement consacrée à la gestion de l'incertitude liée aux attributs. Par conséquent, il est utile de s'appuyer sur une mesure qui évalue la capacité d'attributs incertains à discriminer entre les classes dans un cadre possibiliste. Pour cette raison, la mesure de spécificité de Yager est utilisée [87].

Selon l'étape 2 de l'algorithme II.1, le pouvoir " $d_{sp}$ " est défini comme le rapport entre la spécificité intra-classe et la somme des spécificités intra-classe et inter-classe. Il est à noter que la définition du " $d_{sp}$ " n'est pas analogue à la définition de " $d_{var}$ ". En fait, nous cherchons à avoir un pouvoir discriminant élevée. Une classe bien représentée selon un attribut  $f_k$  présente une spécificité intra-classe importante. Ainsi, nous avons tendance à avoir une spécificité intra-classe importante pour chaque distribution de possibilités  $\pi_{C_i}^{f_k}$ ,  $i \in \{1, \dots, N\}$ , et une spécificité inter classe faible. La détermination de la spécificité intra-classe et inter-classe est illustré à la Figure II.1.

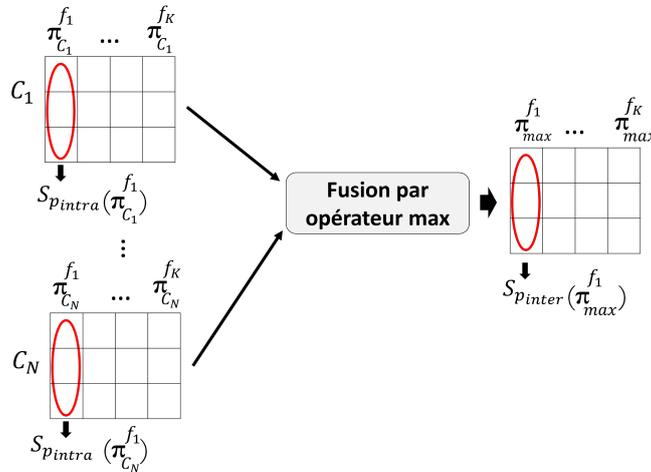


FIGURE II.1 – Détermination de la spécificité intra-classe et inter-classe

— **Dissimilarité " $d_{si}$ " :**

La dissimilarité, définie par " $d_{si}$ ", est utilisée pour évaluer, le degré de différence entre les  $N$  distributions de possibilités de  $\Omega$  (Étape 3 de l'algorithme II.1).

Une fois les trois mesures d'informations " $d_{var}$ ", " $d_{sp}$ " et " $d_{si}$ " sont calculées, nous les normalisons, puis nous les fusionnons afin de générer le pouvoir discriminant "DP\_SUMF". Enfin, les attributs dont le pouvoir discriminant est supérieur à  $\delta$  (seuil choisi empiriquement), sont sélectionnés pour la discrimination des classes.

---

**Algorithme II.1 : App 1\_DP\_SUMF**


---

**Entrées :**  $\Omega = \{C_1, C_2, \dots, C_N\}$ ,  $F = \{f_1, f_2, \dots, f_K\}$  : l'ensemble d'attributs,  $\delta$  : seuil choisi empiriquement

**Sorties :** DP\_1 : matrice à deux colonnes dont la première contenant des valeurs booléennes (1 : si l'attribut est sélectionné, 0 : si l'attribut est rejeté) et la deuxième colonne représente les pouvoirs discriminants d'attributs

bb  $\leftarrow$  [], DP  $\leftarrow$  []

**Pour** k = 1 à K **Faire :**

*Étape 1 :* Calcul du pouvoir discriminant basé sur la mesure de variance " $d_{var}$ "

**Pour** i=1 à N **Faire :**

$x_{C_i}^{f_k} \leftarrow [x_{C_i}^{f_k}(1), \dots, x_{C_i}^{f_k}(m)]$  : l'ensemble des mesures de l'attribut  $f_k$  sur la classe  $C_i$

$$\overline{x_{C_i}^{f_k}} \leftarrow \sum_{m=1}^{card(C_i)} \frac{x_{C_i}^{f_k}(m)}{card(C_i)}$$

**Fin Pour**

$$\overline{x^{f_k}} \leftarrow \sum_{i=1}^N \frac{\overline{x_{C_i}^{f_k}}}{N}$$

$$var_{inter}^{f_k} \leftarrow \sum_{i=1}^N card(C_i) * (\overline{x_{C_i}^{f_k}} - \overline{x^{f_k}})^2$$

$$var_{intra}^{f_k} \leftarrow \sum_{i=1}^N \sum_{m=1}^{card(C_i)} (x_{k,m} - \overline{x_{C_i}^{f_k}})^2$$

$$d_{var}^{f_k} \leftarrow \frac{var_{inter}^{f_k}}{var_{inter}^{f_k} + var_{intra}^{f_k}}$$

*Étape 2 :* Calcul du pouvoir discriminant basé sur la mesure de spécificité " $d_{sp}$ "

**Pour** i=1 à N **Faire :**

$p_{C_i}^{f_k} \leftarrow$  Détermination de la distribution de probabilité pour  $C_i$  selon  $f_k$

$\pi_{C_i}^{f_k} \leftarrow$  Transformation de  $p_{C_i}^{f_k}$  en une distribution de possibilités [80]

**Fin Pour**

$$sp_{intra}^{f_k} \leftarrow \sum_{i=1}^N \text{Spécificité Yager } (\pi_{C_i}^{f_k}) \text{ [87]}$$

$$\pi_{max}^{f_k} \leftarrow \max(\pi_{C_i}^{f_k}), i \in \{1, \dots, N\} \text{ (voir Figure II.1)}$$

$$sp_{inter}^{f_k} \leftarrow \text{Spécificité Yager } (\pi_{max}^{f_k}) \text{ [87]}$$

$$d_{sp}^{f_k} \leftarrow \frac{sp_{intra}^{f_k}}{sp_{intra}^{f_k} + sp_{inter}^{f_k}}$$

*Étape 3 :* Calcul de la dissimilarité " $d_{si}$ " entre les différentes distributions de possibilités

$$S_{i,j}^{f_k} \leftarrow \text{similarité}(\pi_{C_i}^{f_k}, \pi_{C_j}^{f_k}) \text{ } i, j \in \{1, \dots, N\}; i \neq j \text{ [92]}$$

$$d_{si}^{f_k} \leftarrow 1 - S_{i,j}^{f_k}$$

*Étape 4 :* Normalisation

**Pour** j=1 à card( $\pi_{C_1}^{f_k}$ ) **Faire :**

$$d_{Nvar}^{f_k}(j) \leftarrow \frac{d_{var}^{f_k}(j) - \min(d_{var}^{f_k})}{\max(d_{var}^{f_k}) - \min(d_{var}^{f_k})}$$

$$d_{Nsp}^{f_k}(j) \leftarrow \frac{d_{sp}^{f_k}(j) - \min(d_{sp}^{f_k})}{\max(d_{sp}^{f_k}) - \min(d_{sp}^{f_k})}$$

$$d_{Nsi}^{f_k}(j) \leftarrow \frac{d_{si}^{f_k}(j) - \min(d_{si}^{f_k})}{\max(d_{si}^{f_k}) - \min(d_{si}^{f_k})}$$

**Fin Pour**

*Étape 5* : Détermination du pouvoir discriminant "DP\_SUMF"

$$DP\_SUMF^{f_k} \leftarrow d_{Nvar}^{f_k} + d_{Nsp}^{f_k} + d_{Nsi}^{f_k}$$

$$DP \leftarrow [DP \ DP\_SUMF]$$

*Étape 6* : Sélection d'attributs

Si  $DP\_SUMF^{f_k} > \delta$

$$b \leftarrow 1$$

Sinon

$$b \leftarrow 0$$

Fin Si

$$bb \leftarrow [bb \ b]$$

**Fin Pour**

$$DP\_1 \leftarrow [bb \ DP]$$


---

## II.1.2 App 2 : Estimation du pouvoir discriminant en se basant sur des mesures possibilistes de spécificité et de conflit (DP\_PMSC : Discriminant power based on possibilistic measures of specificity and conflict)

Cette approche de sélection d'attributs est basée sur le calcul du pouvoir discriminant, noté "DP\_PMSC", qui est défini comme la valeur moyenne de la capacité " $C_{prep}^{f_k}$ " d'un attribut  $f_k$  de représenter les classes et sa capacité " $C_{sep}^{f_k}$ " de les séparer. Comme illustré dans l'algorithme II.2, la capacité " $C_{prep}^{f_k}$ " d'un attribut  $f_k$  utilise la mesure de spécificité de Yager. En fait, plus un modèle est spécifique, mieux il est informatif et représentatif de sa classe. La Figure II.2 montre que la séparation entre deux classes est d'autant plus marquée que leurs modèles respectifs soient plus spécifiques (de forme plus étroite). Les deux cas illustrés par Figure II.2(c) et (d) montrent deux classes représentées par des modèles non spécifiques, ne présentent aucune utilité pour la séparation.

Il est impératif qu'au moins l'un des modèles soit relativement spécifique pour pouvoir discuter de la séparabilité entre les classes. Les hauteurs des points d'intersection entre les modèles, sont informatives sur la séparabilité entre les classes. Moins cette valeur est élevée,

mieux les classes sont séparables. Les Figures II.2(a) et (e) illustrent des situations favorisant la séparabilité entre les deux classes, alors que les Figures II.2(b) et (f) illustrent les situations défavorables pour la séparation. Le cas idéal, illustré par la Figure II.2(a), correspondant à une hauteur nulle.

---

**Algorithme II.2 : App 2\_DP\_PMSC**

---

**Entrées :**  $\Omega = \{\pi_{C_1}, \pi_{C_2}, \dots, \pi_{C_N}\}$ ,  $F = \{f_1, f_2, \dots, f_K\}$  : l'ensemble d'attributs,  $\delta$  : seuil choisi empiriquement

**Sorties :** DP\_2 : matrice à deux colonnes dont la première contenant des valeurs booléennes (1 : si l'attribut est sélectionné, 0 : si l'attribut est rejeté) et la deuxième colonne représente les pouvoirs discriminants d'attributs

bb  $\leftarrow$  [], DP  $\leftarrow$  []

**Pour** k=1 à K **Faire**

*Étape 1* : Détermination du " $C_{prep}^{fk}$ "

**Pour** i=1 à N **Faire**

$S_p(\pi_{C_i}^{fk}) \leftarrow$  Specificité Yager ( $\pi_{C_i}^{fk}$ )

**Fin Pour**

$C_{prep}^{fk} \leftarrow \frac{\sum_{i=1}^N S_p(\pi_{C_i}^{fk})}{N}$

*Étape 2* : Détermination du " $C_{psep}^{fk}$ "

" $H_k$ " : l'ensemble des hauteurs des points d'intersection de toutes les distributions de possibilités

$h_{i,j}^{fk} \leftarrow \pi_{C_i}^{fk} = \pi_{C_j}^{fk}$   $i, j \in \{1, \dots, N\}; i \neq j$

$H_k \leftarrow \{h_{i,j}^{fk} \neq 0\}$

$M_h^{fk} \leftarrow \max(H_k)$  : Calcul du consensus maximal (Cf. Figure 3(a))

$C_{psep}^{fk} \leftarrow 1 - M_h^{fk}$  : Calcul du conflit minimal (Cf. Figure 3(a))

*Étape 3* : Détermination du pouvoir discriminant "DP\_PMSC"

$DP\_PMSC^{fk} \leftarrow \frac{C_{prep}^{fk} + C_{psep}^{fk}}{2}$

DP  $\leftarrow$  [DP DP\_PMSC]

*Étape 4* : Sélection d'attributs

Si  $DP\_PMSC^{fk} > \delta$

b  $\leftarrow$  1

Sinon

b  $\leftarrow$  0

Fin Si

bb  $\leftarrow$  [bb b]

**Fin Pour**

DP\_2  $\leftarrow$  [bb DP]

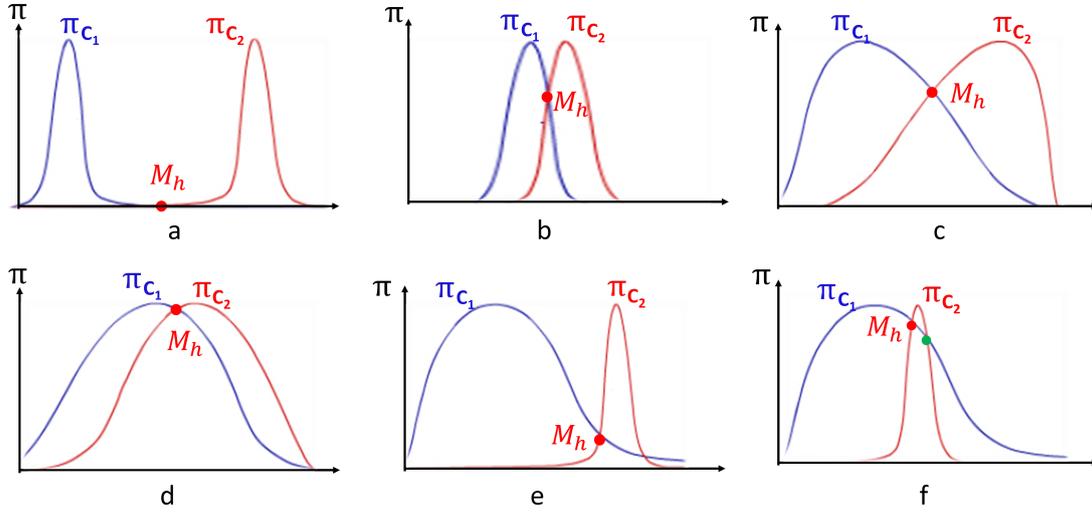


FIGURE II.2 – Différentes situations de représentation de deux classes  $C_1$  et  $C_2$  selon les formes et les dispositions de leurs modèles respectifs

### II.1.3 App 3 : Estimation du pouvoir discriminant en se basant sur des mesures possibilistes de consensus et de conflit (DP\_PMCC : Discriminant power based on possibilistic measures of consensus and conflict)

Cette approche de sélection d'attributs est basée sur le calcul du pouvoir discriminant, noté "DP\_PMCC", qui est défini comme la valeur moyenne de la capacité " $C_{prep}^{f_k}$ " d'un attribut  $f_k$  de représenter les classes et sa capacité " $C_{psep}^{f_k}$ " de les séparer. Comme illustré dans l'algorithme II.3, la capacité " $C_{prep}^{f_k}$ " d'un attribut  $f_k$  est définie comme la somme des aires des consensus entre les différentes distributions de possibilités de  $\Omega$ , tandis que sa capacité " $C_{psep}^{f_k}$ " est définie comme le conflit minimal entre elles.

Le calcul de l'air du consensus  $A_r$  dépend fortement des valeurs des axes des distributions de possibilités. Étant donné que nous ne traitons que des distributions normalisées, l'axe des ordonnées est toujours compris entre 0 et 1. Ainsi, cela ne pose pas de problème lors du calcul de  $A_r$ . Cependant, sur l'axe des abscisses, nous représentons les valeurs d'attributs qui n'ont pas généralement les mêmes marges de valeurs. En fait, une classe peut être caractérisée par un attribut  $f_1$  dont les valeurs comprises entre 0.001 et 0.1, et par un autre attribut  $f_2$  dont les valeurs comprises entre 10 et 50. Pour résoudre ce problème, nous avons décidé d'exprimer l'axe

des abscisses en fonction du nombre de bins (le nombre de classes d'histogramme représentant la taille de la distribution de probabilité, et donc la taille de la distribution de possibilité).

---

### Algorithme II.3 : App 3\_DP\_PMCC

---

**Entrées :**  $\Omega = \{\pi_{C_1}, \pi_{C_2}, \dots, \pi_{C_N}\}$ ,  $F = \{f_1, f_2, \dots, f_K\}$  : l'ensemble d'attributs,  $\delta$  : seuil choisi empiriquement

**Sorties :** DP\_3 : matrice à deux colonnes dont la première contenant des valeurs booléennes (1 : si l'attribut est sélectionné, 0 : si l'attribut est rejeté) et la deuxième colonne représente les pouvoirs discriminants d'attributs

bb  $\leftarrow$  [], DP  $\leftarrow$  []

**Pour** k=1 à K **Faire**

*Étape1* : Détermination du " $C_{prep}^{f_k}$ "

$A_r^{f_k} \leftarrow$  aire de consensus entre  $\{\pi_{C_1}, \dots, \pi_{C_N}\}$

$$C_{prep}^{f_k} \leftarrow \frac{(card(\pi_{C_1}^{f_k})-1)*1 - A_r^{f_k}}{(card(\pi_{C_1}^{f_k})-1)*1}$$

*Étape2* : Détermination du " $C_{psep}^{f_k}$ "

" $H_k$ " : l'ensemble des hauteurs des points d'intersection de toutes les distributions de possibilités

$$h_{i,j}^{f_k} \leftarrow \pi_{C_i}^{f_k} = \pi_{C_j}^{f_k} \quad i, j \in \{1, \dots, N\}; i \neq j$$

$$H_k \leftarrow \{h_{i,j}^{f_k} \neq 0\}$$

$M_h^{f_k} \leftarrow \max(H_k)$  : Calcul du consensus maximal (Cf. Figure 3(a))

$C_{psep}^{f_k} \leftarrow 1 - M_h^{f_k}$  : Calcul du conflit minimal (Cf. Figure 3(a))

*Étape3* : Détermination du pouvoir discriminant "DP\_PMCC"

$$DP\_PMCC^{f_k} \leftarrow \frac{C_{prep}^{f_k} + C_{psep}^{f_k}}{2}$$

DP  $\leftarrow$  [DP DP\_PMCC]

*Étape 4* : Sélection d'attributs

Si  $DP\_PMCC^{f_k} > \delta$

b  $\leftarrow$  1

Sinon

b  $\leftarrow$  0

Fin Si

bb  $\leftarrow$  [bb b]

**Fin Pour**

DP\_3  $\leftarrow$  [bb DP]

---

La Figure II.3 illustre six cas pour discriminer deux classes en fonction de la forme de consensus entre les distributions de possibilités. Les classes des Figures II.3(a), (c), et (e) sont mieux discriminées que les classes des Figures II.3(b), (d), et (f) en raison de leurs faibles valeurs de  $M_h$  et  $A_r$ . La situation idéale, illustrée à la Figure II.3(a) est caractérisée par  $M_h$  et  $A_r$  correspondants aux  $\pi_{C_1}$  et  $\pi_{C_2}$  nuls.

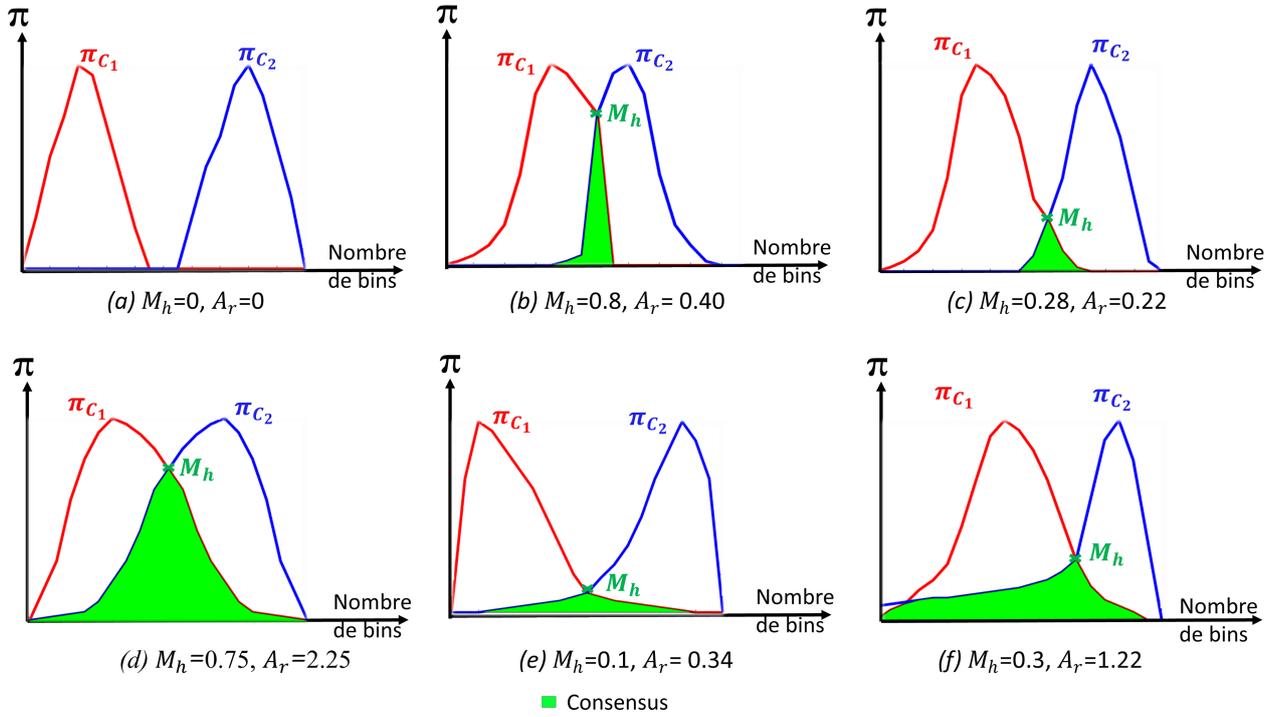


FIGURE II.3 – Différentes situations de représentation deux distributions de possibilités selon la forme de leur consensus

## II.1.4 Résultats et discussions

Deux expériences ont été menées pour évaluer nos approches de sélection d'attributs proposées. La première expérience a été menée sur une base de données synthétiques, qui a été construite de façon que nous connaissons à l'avance la capacité de chaque attribut à discriminer et à représenter les classes. Cette expérience nous aide à vérifier si nous avons atteint les résultats escomptés. La deuxième expérience consiste à évaluer nos approches sur différentes bases de données de référence et à les comparer avec d'autres approches de sélection d'attributs. Le critère d'évaluation est le taux de correctes classifications, effectué par usage des attributs sélectionnés en utilisant le classifieur SVM.

### 1. Base de données synthétiques

La base de données synthétiques est formée par un ensemble d'attributs pertinents et non pertinents représentés sous forme de distributions de possibilités (Cf. Figure II.4). L'univers

des discours de cette base de données est composé de trois classes  $\Omega = \{C_1, C_2, C_3\}$ , et chaque classe est représentée par neuf attributs  $F = \{f_1, f_2, \dots, f_9\}$ .

Comme le montre la Figure II.6, pour un seuil  $\delta$  égale à 0.6 (choisi empiriquement) sur les différents pouvoirs discriminants, présentés dans le Tableau II.1, cinq attributs  $\{f_1, f_2, f_5, f_6, f_9\}$  sont sélectionnés.

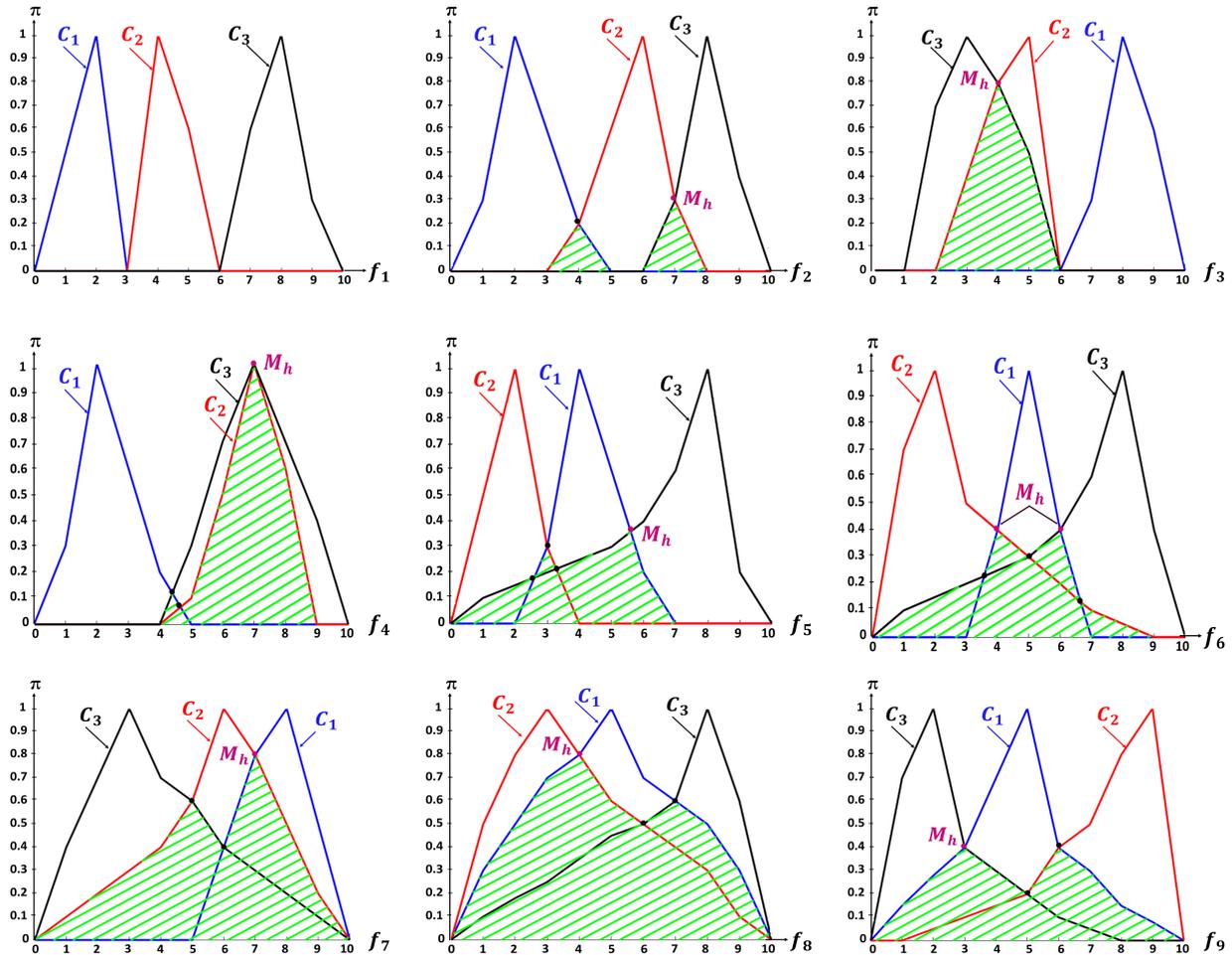


FIGURE II.4 – Base de données synthétiques

TABLE II.1 – Pouvoirs discriminants déterminés selon les différentes approches de sélection d'attributs proposées.

	Ensemble de consensus $\{H_k\}$	Consensus maximal $M_h$	$A_r$	DP_SUMF	DP_PMSC	DP_PMCC
$f_1$	$\{\emptyset\}$	$\emptyset$	0	<b>0.92</b>	<b>0.896</b>	<b>1</b>
$f_2$	$\{0.2, 0.3\}$	0.3	0.05	<b>0.753</b>	<b>0.724</b>	<b>0.825</b>
$f_3$	$\{0.8\}$	0.8	0.17	0.498	0.479	0.515
$f_4$	$\{0.1, 0.05, 1\}$	1	0.222	0.386	0.381	0.389
$f_5$	$\{0.15, 0.3, 0.2, 0.35\}$	0.35	0.129	<b>0.721</b>	<b>0.678</b>	<b>0.76</b>
$f_6$	$\{0.2, 0.4, 0.3, 0.4, 0.1\}$	0.4	0.162	<b>0.693</b>	<b>0.626</b>	<b>0.719</b>
$f_7$	$\{0.6, 0.4, 0.8\}$	0.8	0.35	0.407	0.365	0.425
$f_8$	$\{0.8, 0.5, 0.6\}$	0.8	0.48	0.33	0.313	0.36
$f_9$	$\{0.4, 0.2, 0.4\}$	0.4	0.22	<b>0.647</b>	<b>0.602</b>	<b>0.69</b>

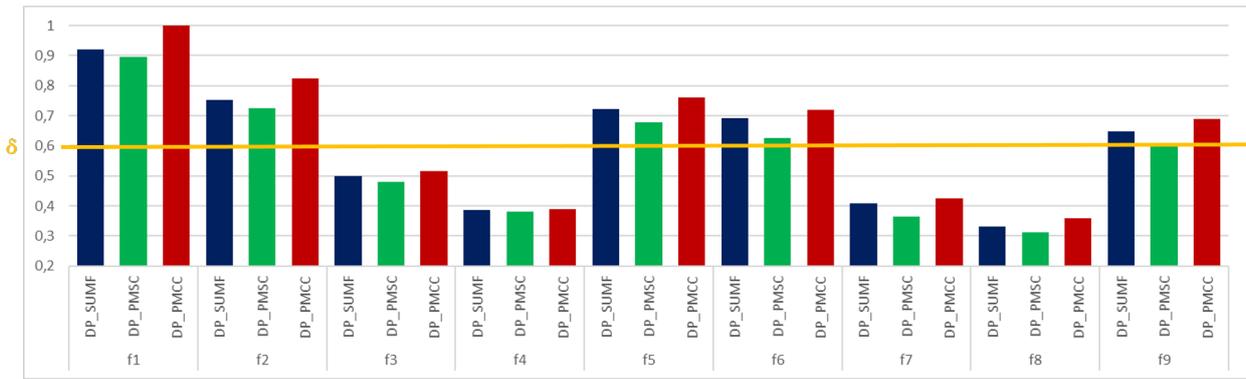


FIGURE II.5 – Pouvoirs discriminants déterminés selon les différentes approches de sélection d'attributs proposées.

## 2. Bases de données de références (Benchmark)

Pour évaluer nos approches de sélection d'attributs, nous avons les testées sur quatre bases de données de référence [98], à savoir la base de données de verres, la base de données de fleurs d'iris, la base de données de diabète Indien de Pima et la base de données de silhouettes de véhicules. Les résultats obtenus sont comparés à ceux obtenus à partir de cinq approches de sélection d'attributs de la littérature, à savoir 4 approches classiques PCA, LFS, le score de Fisher, Relief et une approche "Poss\_FS" proposée dans [97] basée sur des modèles possibilistes. L'évaluation des approches de sélection d'attributs est basée sur deux concepts : le nombre d'attributs sélectionnés (NAS) et le taux de correcte classifications (TCC). Le Tableau II.2 illustre les résultats obtenus pour chaque approche de sélection d'attributs testée sur les différentes bases de données considérées.

Comme le montre le Tableau II.2, le TCC est très dépendant du NAS, ce qui rend difficile de prouver la contribution de nos approches. Pour éviter ces dépendances, dans [7], il a été proposé de quantifier le gain obtenu par chaque approche de sélection à travers deux taux : i) le taux de sélection d'attributs (TSA) défini par l'Équation II.1 et ii) le taux d'influence sur la classification (TIC) défini par l'Équation II.2. Une valeur TIC supérieure à 1 signifie que la sélection a amélioré le taux de classification correct avant la sélection d'attributs. Une valeur TIC inférieure à 1 signifie que la sélection a dégradé la reconnaissance. Plus  $|1-TIC|$  est élevé, plus l'impact de la sélection d'attributs sur le taux de classification est important.

$$TSA = \frac{\text{Nombre d'attributs sélectionnés}}{\text{Nombre de tous les attributs}} \quad (\text{II.1})$$

$$TIC = \frac{TCC (\text{après sélection d'attributs})}{TCC (\text{avec tous les attributs})} \quad (\text{II.2})$$

TABLE II.2 – Résultats de la sélection d'attributs obtenus pour différentes approches.

Approche	Iris		Verre		Diabètes		Véhicule	
	NAS	TCC	NAS	TCC	NAS	TCC	NAS	TCC
Tous les attributs	4	96.2	9	95.3	8	78.1	18	82.8
PCA	2	81.3	6	93.9	4	75.6	7	74.7
LFS	2	95.3	7	93.9	4	77.8	10	70.6
Fisher	2	81.3	4	91.5	6	76.4	10	74.2
Relief	2	81.3	4	93.4	4	77.2	10	72.8
Poss_FS [97]	2	95.3	7	94.4	4	76.3	8	72.8
DP_SUMF	2	95.3	5	94.6	4	77.6	6	76.8
DP_PMSC	2	95.3	6	93.9	4	77.2	7	74.7
DP_PMCC	2	95.3	5	94.8	4	77.8	6	77.5

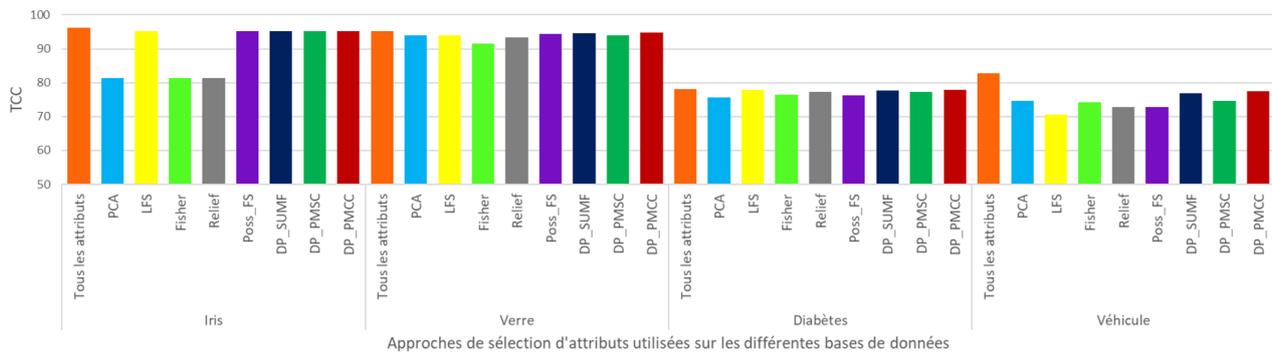


FIGURE II.6 – Taux de correcte classifications obtenus par les différentes approches de sélection d'attributs sur les différentes bases de données

Chaque point de la Figure II.7 correspond aux performances (TSA, TIC) obtenues par une approche de sélection d'attributs pour une base de données. Il est tout à fait visible qu'il n'y a pas de couples (approche / base de données) qui ont un TIC supérieur à 1. Cela signifie que pour les cas étudiés (approche/base de données), l'utilisation de tous les attributs conduit au meilleur taux. Les attributs utilisés sont donc bien appropriés pour les applications en question. Il se peut que ces attributs ont été déjà présélectionnés. Il convient de noter que par rapport à d'autres approches, nos approches proposées montrent une stabilité remarquable avec des valeurs TIC très proches de 1. Ceci est visible par la forme linéaire, horizontale, serrée, rectiligne, proche de l'unité de leurs courbes.

Puisque nous cherchons à évaluer la capacité des approches de sélection d'attributs à maintenir la stabilité du taux de reconnaissance, nous proposons de calculer la moyenne et l'écart type correspondants à chaque approche de sélection d'attributs. L'approche ayant la valeur de l'écart type la plus faible est celle qui a la plus grande stabilité de reconnaissance en ce qui concerne le nombre d'attributs sélectionnés. Comme le montre la Figure II.9, l'approche basée sur le pouvoir discriminant "DP\_SUMF" et l'approche basée sur le pouvoir discriminant "DP\_PMCC" obtiennent les meilleurs scores de stabilité.

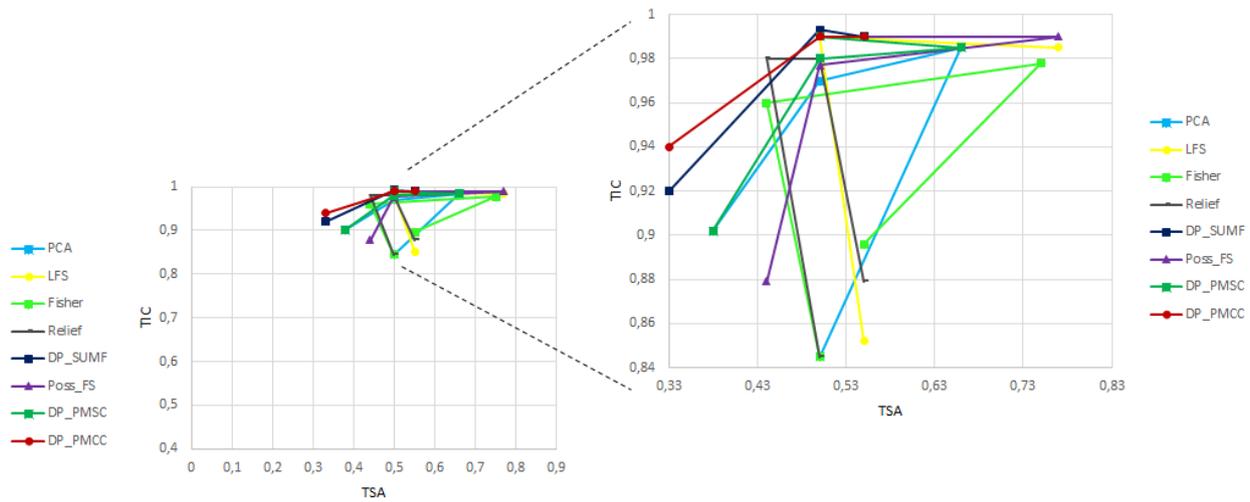


FIGURE II.7 – Variation du TIC en fonction du TSA

Nos approches permettent d'obtenir de bons résultats de classification sur n'importe quelle base de données tout en poussant au mieux la réduction des attributs. Contrairement à nos approches et à "Poss\_FS" [97], d'autres approches de la littérature, ont des performances aléatoires. Ainsi, nous pouvons conclure que l'inférence des capacités de représentation / séparation aboutit à un bon compromis entre le taux de réduction des attributs et le taux correct de classification sur n'importe quelle base de données. Il convient de noter que nos approches de sélection d'attributs sont moins complexes que "Poss\_FS" [97]. En fait, leur approche dépend principalement de la complexité de l'indice de Shapley qui croît de façon exponentielle à mesure que le nombre d'attributs augmente.

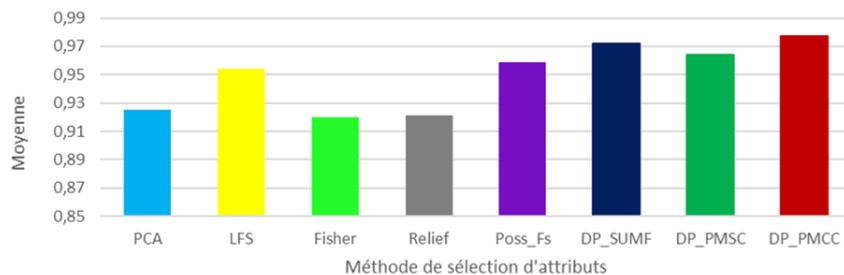


FIGURE II.8 – Moyenne des mesures TIC calculés pour les approches de sélection d'attributs évalués sur 4 bases de données

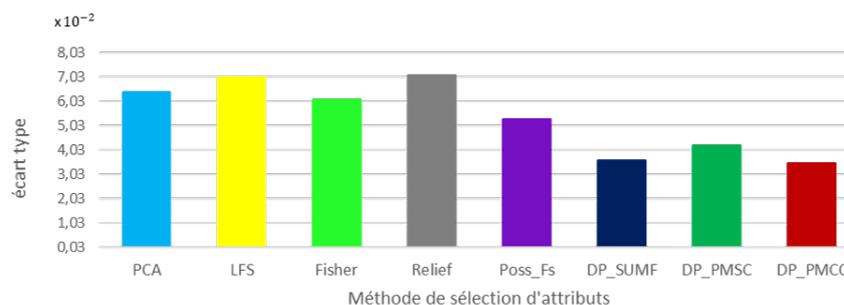


FIGURE II.9 – Écart type des mesures TIC calculés pour les approches de sélection d'attributs évalués sur 4 bases de données

## II.2 Proposition de classifieurs binaires hiérarchiques

### II.2.1 État de l'art

Dans les applications d'analyse et d'interprétation de données, la classification multi-classes est une tâche importante. Elle est couramment utilisée dans de nombreuses applications du monde réel, notamment la prévision des catastrophes naturelles [99], la sécurité [100], la médecine [101], la psychologie [102], etc. Les applications de classification multi-classes sont généralement définies par une quantité massive de données. Plusieurs algorithmes ont été proposés pour résoudre ce problème, qui peuvent être divisés en deux catégories [103].

*Première catégorie* : elle gère directement le problème multi-classes en utilisant des algorithmes tels que les réseaux de neurones artificiels, les réseaux Naive-Bayes, les machines vectorielles de support (SVM) [104, 105], les k-plus proches voisins, les arbres de décision [106], etc. La principale limite de cette première catégorie est de déterminer les attributs qui permettent de distinguer les classes lorsque le nombre de classe est important [107]. Par conséquent, les algorithmes définis dans cette catégorie peuvent produire une précision médiocre.

*Deuxième catégorie* : elle convertit le problème de classification multi-classes en plusieurs problèmes de classification binaires. Dans [103, 104, 106, 107, 109, 110, 111, 127], les auteurs développent des classifieurs binaires hiérarchiques qui décomposent récursivement un problème de N-classes en un problème de N-1 classes. Hosenie *et al.* [112] ont proposé un classifieur hiérarchique afin d'atténuer le problème de déséquilibre d'apprentissage mentionné dans [113]. Dans [114], les auteurs ont développé un SVM hiérarchique (HSVM). Le classifieur proposé a rapporté de bonnes performances lors de l'utilisation des données de télédétection.

Parmi les stratégies les plus souvent utilisées pour la décomposition d'un problème multi-classes en problèmes binaires, nous citons le codage de sortie avec correction d'erreur, un contre tous (One Versus All : OVA), tous contre tous (All Versus All : AVA), etc. La stratégie OVA est largement utilisée pour traiter les problèmes de classification multi-classes car elle est simple et intuitive. Bien que simple, elle offre des résultats efficaces au moins aussi précis que d'autres stratégies. Comparée au codage correcteur d'erreurs et à l'AVA, l'OVA a la plus faible complexité de calcul. Soient  $m$  et  $n$  représentent respectivement le nombre de classes et le nombre d'observations, la quantité de calcul requise par l'AVA est  $O(m.n)$  qui est potentiellement beaucoup plus grande que  $O(n)$  requise par OVA [115, 116].

La plupart des algorithmes de classification conventionnels supposent que les données d'ap-

prentissage sont définies avec précision, ce qui n'est généralement pas réaliste. En fait, dans les applications pratiques, en raison de sources d'informations peu fiables, de bruit, d'erreurs de mesure, diversité de données, etc., l'incertitude dans l'apprentissage est omniprésente. Dans le processus de classification, ces formes d'imperfections doivent être prises en compte et soigneusement résolues afin de bénéficier du contenu informationnel caché dans ces données imparfaites. De nombreuses recherches en apprentissage automatique se sont concentrées sur la classification des données incertaines. Dans [117], les auteurs ont proposé un classifieur de vecteur de support total (TSVC) basé sur un classifieur SVM pour traiter l'incertitude des données. Dans [118, 119], les auteurs ont proposé des arbres de décision classiques étendus dans lesquels l'incertitude des données est gérée par la théorie des probabilités. D'autres travaux [120, 121] ont été proposés pour traiter des données d'apprentissage de nature floue en combinant la théorie d'évidence et les arbres de décision. Javid *et al.* [122] ont proposé un classifieur semi-supervisé capable de gérer les zones incertaines de l'OVA en utilisant la fonction de croyance. Agrawal *et al.* [123] ont intégré une nouvelle mesure de distance dans K-NN afin de gérer l'incertitude des attributs numériques.

La théorie des possibilités a été efficacement employée dans des techniques de classification. Dans [124, 125], les auteurs ont proposé des classifieurs possibilistes pour gérer des données imprécises. Un arbre de décision possibiliste est également utilisé pour traiter des informations incertaines dans la procédure d'apprentissage [126], ainsi que pour faire face à l'incertitude épistémique liée aux classes [127].

Dans la suite de cette partie, deux classifieurs binaires hiérarchiques sont développés, l'un est monocritère et l'autre multicritère. Enfin, après avoir appliqué les deux classificateurs binaires hiérarchiques conçus à différentes bases de données de référence (benchmarks), ils sont évalués et comparés avec d'autres classificateurs conventionnels.

## II.2.2 Contexte de l'étude et motivation

Soit  $\Omega = \{C_1, C_2, \dots, C_N\}$  l'univers de discours,  $\{Obj_1, Obj_2, \dots, Obj_N\}$  un ensemble d'objets appartenant respectivement aux classes  $\{C_1, C_2, \dots, C_N\}$  et acquis par un capteur. Soit  $\{f_1, \dots, f_K\}$  l'ensemble d'attributs extraits des données acquises. Ces valeurs d'attributs mesurées sur les objets, permettent l'attribution de ces objets à l'une des classes de  $\Omega$ . Il arrive que l'on puisse mesurer les mêmes valeurs d'un attribut sur des objets appartenant à des classes différentes, tel qu'il est illustré par la Figure II.10. De telles situations créent de l'ambiguïté quant à l'attribution d'une étiquette classe unique aux objets présentant cette ambiguïté. Ceci

affecte l'efficacité du classifieur. Élucider cette ambiguïté s'avère alors d'une grande utilité. C'est l'objectif de la proposition ci-dessous.

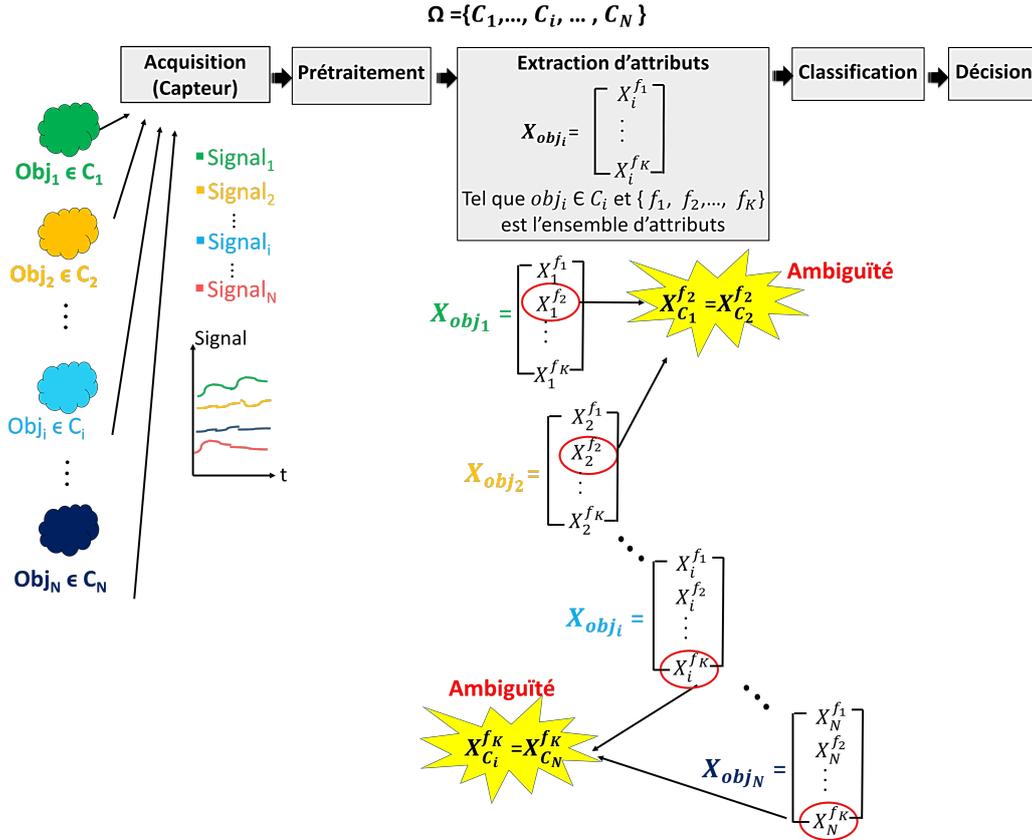


FIGURE II.10 – Incertitude liée aux attributs dans un processus de classification

## II.2.3 Proposition d'un classifieur binaire hiérarchique monocritère "CBH(1)"

### II.2.3.1 Principe

Le classifieur CBH(1) proposé est itératif (Cf. Figure II.12). Son principe consiste à considérer, à l'itération  $t$ , tous les couples de classes  $(C_i, \Omega_i^t)$ ;  $\Omega_i^t = \Omega^{t-1} / C_i$ ,  $i \in \{1, \dots, \text{card}(\Omega^{t-1})\}$  et  $\Omega^0 = \Omega$ , et à évaluer, sur chaque couple, les pouvoirs discriminants de tous les attributs  $f_k$ ;  $k \in \{1, \dots, K\}$ . La détermination du pouvoir discriminant est fait selon l'algorithme II.1. La Figure II.11 illustre les différentes mesures de dissimilarité utilisées par cet algorithme.

L'attribut qui permet de réaliser la meilleure discrimination parmi l'un des couples  $(C_i, \Omega_i^t)$ , est jugé optimal pour le fait qu'il a permis de minimiser l'ambiguïté dans l'attribution de l'étiquette classe  $C_i$  en question. Celle-ci est aussi retenue comme étant  $C_{op}^t$ . Le couple  $(f_{op}^t, C_{op}^t)$  est alors utilisé dans l'itération  $t$  pour discriminer la classe  $C_{op}^t$ . Pour l'itération suivante  $\Omega^t = \Omega^{t-1} / C_{op}^t$ .

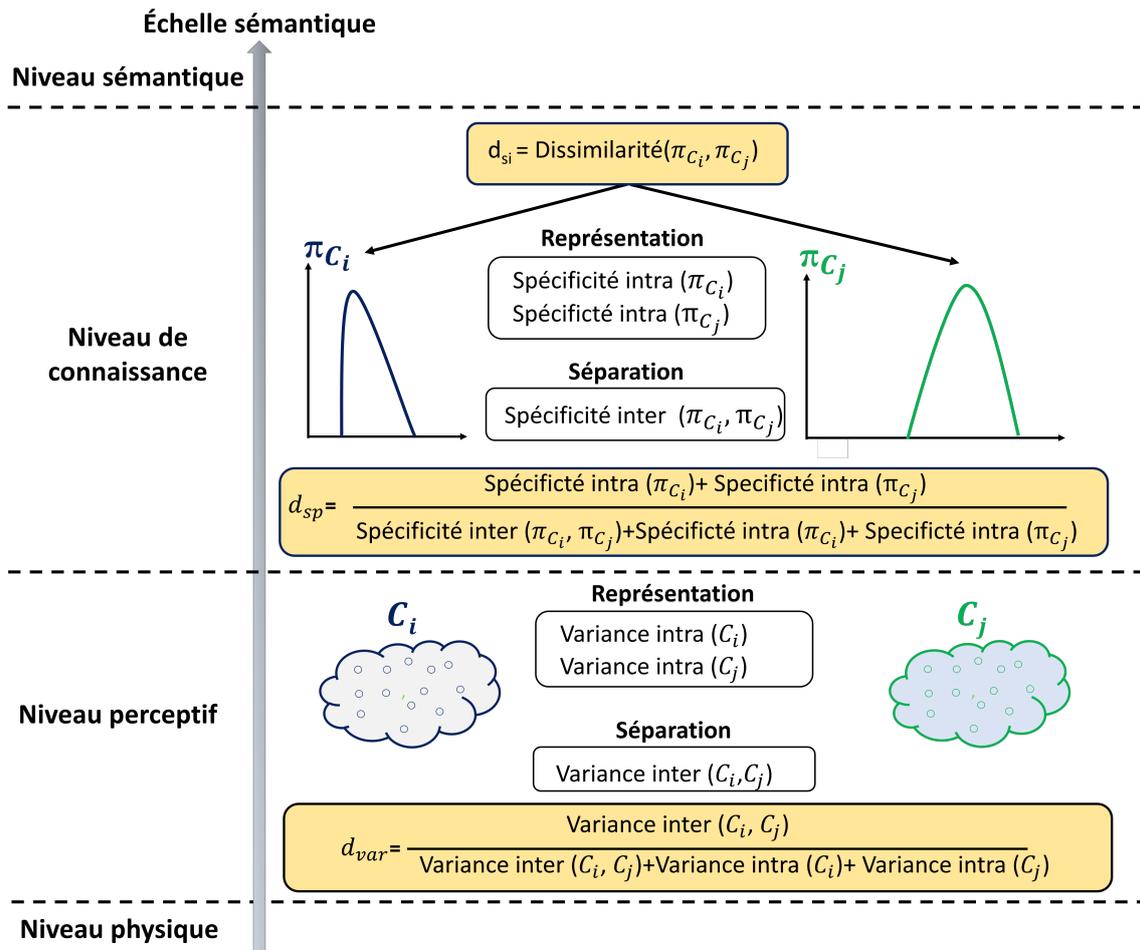


FIGURE II.11 – Situation des mesures de dissimilarité utilisées sur une échelle sémantique

### II.2.3.2 Architecture du CBH(1)

L'architecture du CBH(1) suit un processus descendant détaillé dans l'algorithme II.4. Comme le montre la Figure II.12, le noeud racine indexé "1" représente l'univers fini de discours  $\Omega^0 = \Omega = \{C_1, C_2, \dots, C_N\}$ . À partir de ce noeud racine, le CBH(1) essaie de décomposer  $\Omega^0$  en deux groupes disjoints. Le premier groupe représente la classe  $C_{op}^1$  la plus discriminée parmi les autres classes selon l'attribut  $f_{op}^1$ . L'ensemble  $\Omega^1 = \Omega^0 / C_{op}^1$  est associé au second groupe. Ce processus est appliqué de manière récursive jusqu'à ce que toutes les classes soient discriminées les unes des autres.

---

**Algorithme II.4 :** Processus de sélection de  $C_{op}^t$  et  $f_{op}^t$  pour une itération t

---

**Entrées :**  $\Omega^t = \Omega^{t-1} / C_{op}^t$ ,  $N = \text{card}(\Omega^0)$ ,  $F = \{f_1, f_2, \dots, f_K\}$  : ensemble d'attributs,  $\delta$  : seuil choisi empiriquement

**Sorties :**  $C_{op}^t$  et  $f_{op}^t$  représentent respectivement la classe la plus discriminée et l'attribut optimal qui favorise sa discrimination

pouv ← [ ]

Étape 1 : Définition des (N-t) décompositions possibles de  $(C_i, \Omega_i^t)$

Étape 2 : Détermination des pouvoirs discriminants d'attributs pour chaque décomposition possible

**Pour**  $j = 1$  à (N-t) **Faire** :

$$X_j = (C_i, \Omega_i^t)$$

$$DP\_1 \leftarrow \text{App 1\_DP\_SUMF}(X_j, F, \delta)$$

$$\text{pouv} \leftarrow [\text{pouv } DP\_1( :, 2)]$$

**Fin Pour**

Étape 3 : Recherche de la valeur maximale de la matrice pouv

Étape 4 : Extraction  $C_{op}^t$  et  $f_{op}^t$

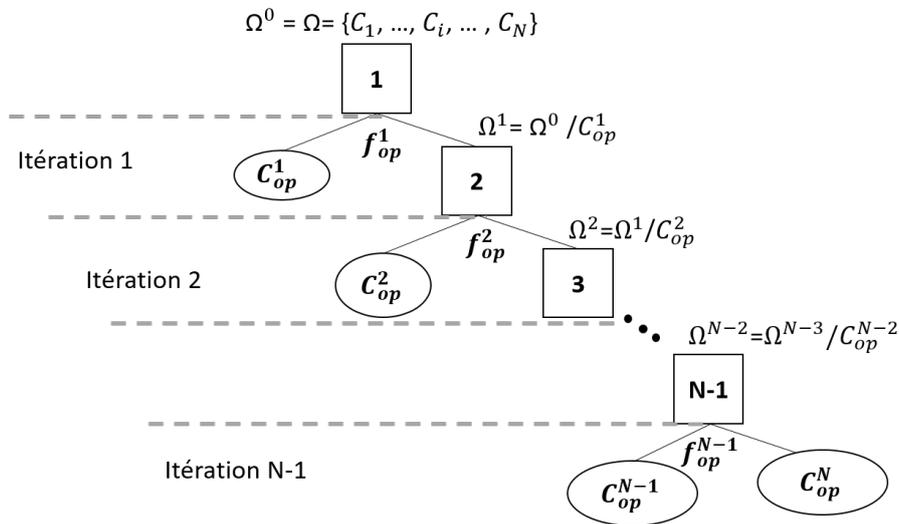


FIGURE II.12 – Architecture du CBH(1)

### II.2.3.3 Phase de classification

Lors de la phase de classification, une étiquette doit être attribuée à chaque nouvel échantillon qui arrive. Considérons un objet inconnu X représenté uniquement par l'ensemble d'attributs optimaux  $\{f_{op}^1, f_{op}^2, \dots, f_{op}^{N-1}\}$ . Afin de définir sa classe, deux stratégies basées sur un raisonnement hiérarchique sont proposées.

#### 1. Classification par maximum de possibilité

La première stratégie de classification repose sur la comparaison de deux entités de granularité différente (un échantillon et une distribution de possibilité), qui nous mène d'évaluer notre classifieur par projection.

Cette stratégie consiste à comparer récursivement la valeur de projection de l'échantillon de l'objet inconnu X sur la distribution de possibilités  $\pi_{C_{op}^t}$  et sa valeur de projection sur la

distribution de possibilités  $\pi_{\Omega^t}$  selon l'attribut  $f_{op}^t$ . Il est à noter que  $\pi_{\Omega^t}$  est obtenue par la fusion de tous les  $\pi_{C_j}$ ,  $C_j \in \Omega_i^t$ , par l'opérateur max. La projection s'arrête à l'itération  $t$ , et l'échantillon de l'objet inconnu  $X$  est considéré comme un échantillon appartenant à un objet de la classe  $C_{op}^t$  si et seulement si  $\alpha_1^t > \alpha_2^t$ , avec :

$\alpha_1^t$  : sa valeur de projection sur  $\pi_{C_{op}^t}$  selon l'attribut  $f_{op}^t$ .

$\alpha_2^t$  : sa valeur de projection sur  $\pi_{\Omega^t}$  selon l'attribut  $f_{op}^t$ .

La Figure II.13 illustre un exemple d'univers fini de discours composé de 4 classes  $\Omega = \{C_1, C_2, C_3, C_4\}$  et un échantillon de test  $X$  qui est supposé un objet de la classe  $C_3$ .

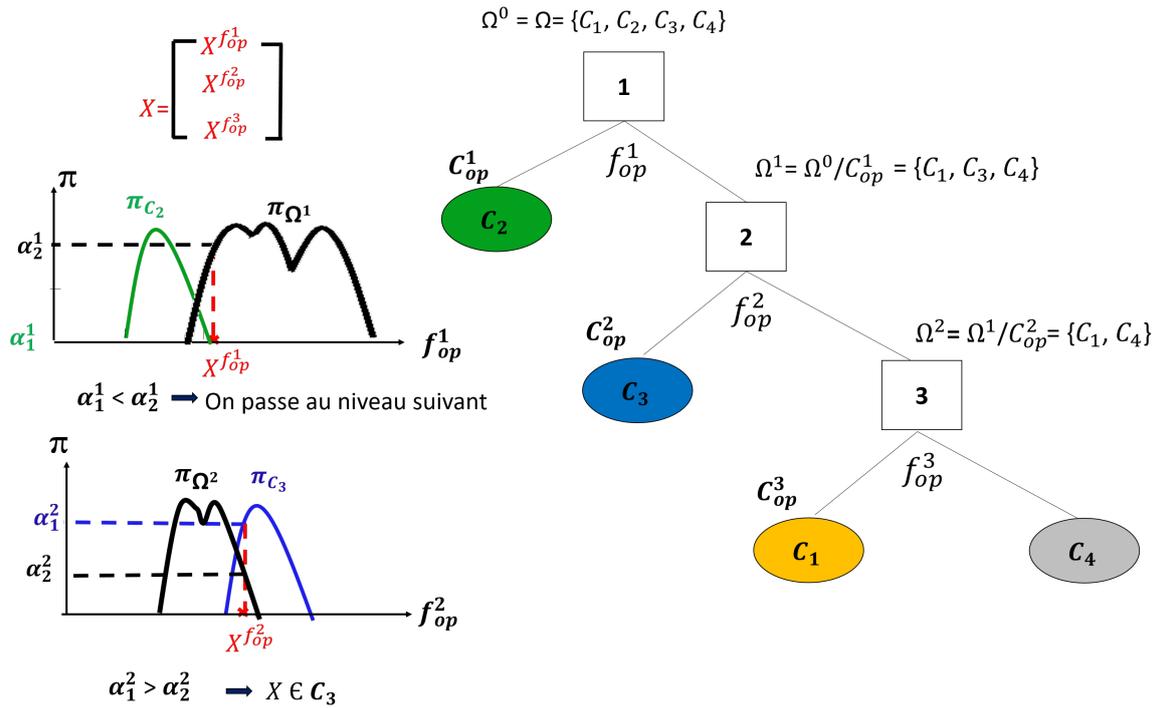


FIGURE II.13 – Classification par maximum de possibilité

## 2. Classification par estimation de similarité entre le modèle de la classe et un modèle d'échantillons

Contrairement à la classification par maximum de possibilité, cette stratégie est basée sur la comparaison de deux entités de même granularité sémantique (distribution de possibilités et une autre distribution de possibilité). En fait, il s'agit de mesurer récursivement le degré de similarité entre la distribution de possibilités  $\pi_X$  correspondant à un objet inconnu  $X$  avec la distributions de possibilistés  $\pi_{C_{op}^t}$  et avec  $\pi_{\Omega^t}$  selon l'attribut  $f_{op}^t$ . La classification s'arrête à l'itération  $t$  et l'objet  $X$  est considéré comme un objet appartenant à la classe  $C_{op}^t$  si et seulement si  $s_1^t > s_2^t$ , avec :

$s_1^t$  : mesure de similarité entre  $\pi_X$  et  $\pi_{C_{op}^t}$  selon l'attribut  $f_{op}^t$ .

$s_2^t$  : mesure de similarité entre  $\pi_X$  et  $\pi_{\Omega^t}$  selon l'attribut  $f_{op}^t$ .

La Figure II.14 illustre un exemple d'univers fini de discours composé de 4 classes  $\Omega = \{C_1, C_2, C_3, C_4\}$  et un échantillon de test  $X$  qui est supposé un objet de la classe  $C_2$ .

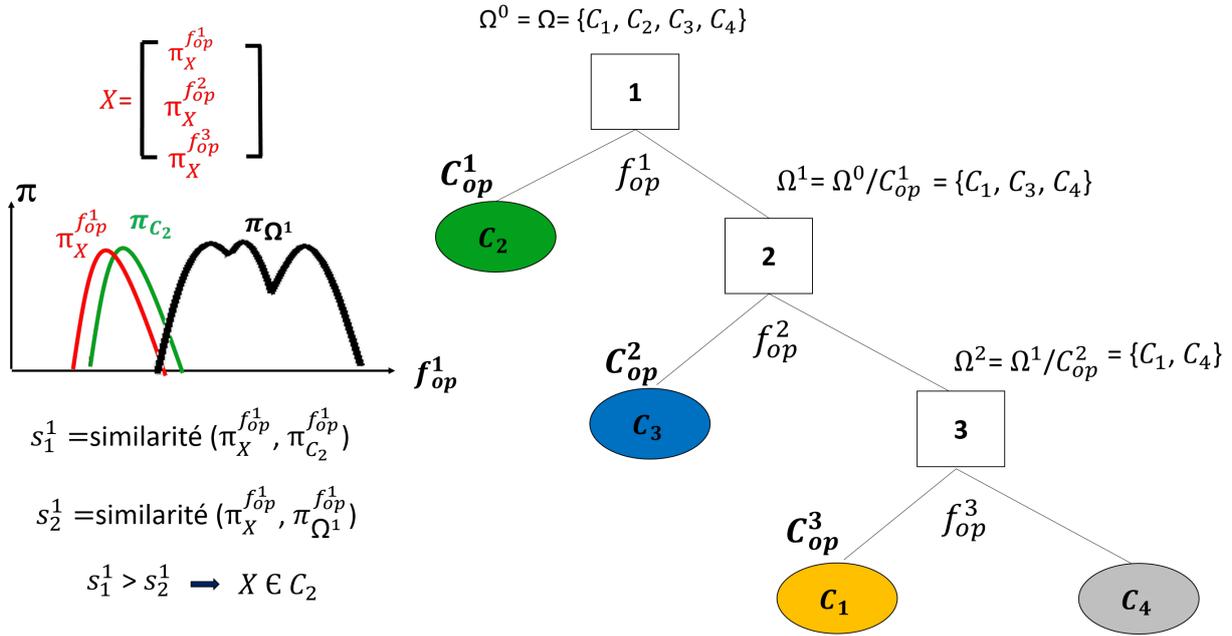


FIGURE II.14 – Classification par maximum de similarité entre le modèle de l'objet  $X$  et les modèles des classes

## II.2.4 Proposition d'un classifieur binaire hiérarchique multicritère "CBH(2)"

### II.2.4.1 Principe

Le classifieur CBH(2) proposé est itératif (Cf. Figure II.15) tout comme CBH(1). Son principe consiste à considérer, à l'itération  $t$ , tous les couples de classes  $(C_i, \Omega_i^t)$ ,  $i \in \{1, \dots, \text{card}(\Omega^{t-1})\}$  et à évaluer, sur chaque couple, les pouvoirs discriminants de tous les attributs  $f_k$ ;  $k \in \{1, \dots, K\}$ .

En définissant pour une itération  $t$ , tous les couples possibles  $(C_i, \Omega_i^t)$ , nous attribuons un score à chaque couple. Ce score est défini comme la somme des pouvoirs discriminants de l'ensemble d'attributs  $F_{op_i}^t$  divisé par son cardinal. Comme le montre l'algorithme II.5, le couple  $(C_i, \Omega_i^t)$  ayant le score le plus élevé est alors considéré comme étant la classe  $C_{op}^t$  la plus distincte de l'ensemble  $\Omega^t$  selon  $F_{op_i}^t$ . Le couple  $(F_{op}^t, C_{op}^t)$  est alors utilisé dans l'itération  $t$  pour discriminer la classe  $C_{op}^t$ . Pour l'itération suivante  $\Omega^t = \Omega^{t-1} / C_{op}^t$ .

Comme nous l'avons vu lors de la construction du CBH(1), la sélection d'un attribut optimal intègre de multiples informations de différents niveaux sémantiques, à savoir : calcul de variance (niveau perceptif), mesure de similarité (niveau de connaissance), etc. Pour le classifieur CBH(2) proposé, nous avons choisi de le construire et de l'évaluer en ne considérant que des modèles

possibilistes représentant les différentes classes de  $\Omega$ , sans recourir au niveau perceptif. Ainsi, nous avons repris l'approche de sélection d'attributs proposée dans l'algorithme II.3.

---

**Algorithme II.5** : Sélection de couple  $(C_{op}^t, \Omega^t)$  pour une itération  $t$

---

**Entrées** :  $\Omega^t = \Omega^{t-1} / C_{op}^t$ ,  $N = \text{card}(\Omega^0)$ ,  $F = \{f_1, f_2, \dots, f_K\}$  : ensemble d'attributs,  $\delta$  : seuil choisi empiriquement

**Sorties** :  $(C_{op}^t, \Omega^t)$  : la classe  $C_{op}^t$  la plus distincte de l'ensemble  $\Omega^t$

*Étape 1* : Définition des  $(N-t)$  décompositions possibles de  $(C_i, \Omega_i^t)$

*Étape 2* : Sélection d'attributs optimaux pour chaque décomposition possible

**Pour**  $j = 1$  à  $(N-t)$  **Faire** :

$$X_j = (C_i, \Omega_i^t)$$

**Pour**  $k=1$  à  $K$  **Faire**

$p_{C_i}^{f_k} \leftarrow$  Détermination de la distribution de probabilité pour la classe  $C_i$  selon  $f_k$

$p_{C_j}^{f_k} \leftarrow$  Détermination de la distribution de probabilité pour chaque classe  $C_j \in \Omega_i^t$ ;  $j \in \{1, \dots, \text{card}(\Omega_i^t)\}$  selon  $f_k$

$\pi_{C_i}^{f_k} \leftarrow$  Transformation de  $p_{C_i}^{f_k}$  en une distribution de possibilités [80]

$\pi_{C_j}^{f_k} \leftarrow$  Transformation de  $p_{C_j}^{f_k}$  en une distribution de possibilités [80]

$\pi_{\Omega_i^t}^{f_k} \leftarrow$  Fusion de tous les  $\pi_{C_j}^{f_k}$  par un opérateur max

**Fin Pour**

$$Y_j = (\pi_{C_i}, \pi_{\Omega_i^t})$$

$$DP\_3 \leftarrow \text{App 3\_DP\_PMCC}(Y_j, F, \delta)$$

$$F_{op_i}^t \leftarrow DP\_3(:, 2)$$

$$\text{score}(Y_j) \leftarrow \frac{\sum_{f_l \in F_{op_i}^t} DP\_3(l, 2)}{\text{card}(F_{op_i}^t)} : \text{Calcul du score pour chaque décomposition possible}$$

**Fin Pour**

$$\text{score}(Y_{op}) \leftarrow \max_{j=1}^{N-t} (\text{score}(Y_j))$$

$$(C_{op}^t, \Omega^t) \leftarrow \text{le couple } (C_i, \Omega_i^t) = Y_{op}$$


---

### II.2.4.2 Architecture du CBH(2)

Comme le montre la Figure II.15, l'architecture du classifieur CBH(2) est similaire à CBH(1), mais la discrimination de la classe  $C_{op}^t$  de l'ensemble  $\Omega^t = \Omega^{t-1} / C_{op}^t$ , est faite selon un ensemble d'attributs  $F_{op}^t$ .

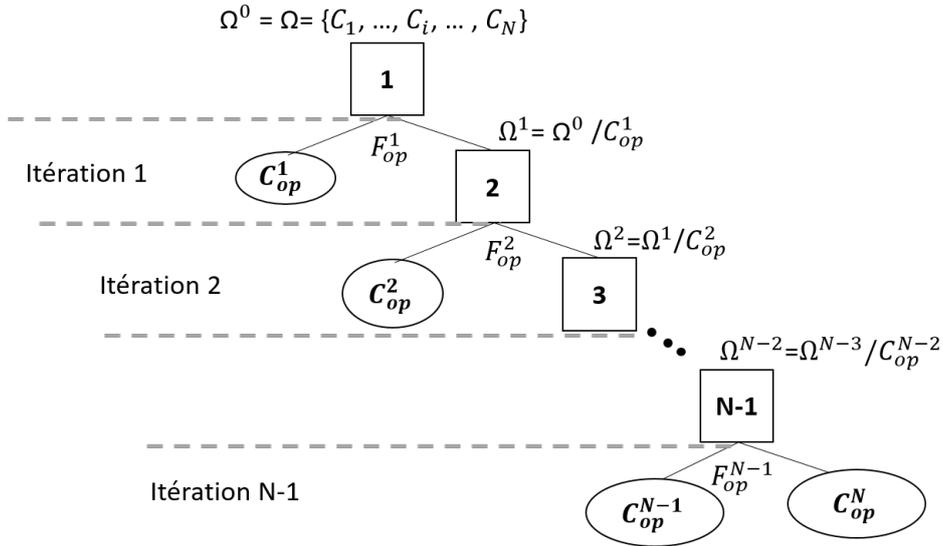


FIGURE II.15 – Architecture du CBH(2)

### II.2.4.3 Phase de classification

Le principe du CBH(2) est hiérarchique tout comme CBH(1) avec la différence de l'utilisation d'un ensemble d'attributs optimaux au niveau de couche au lieu d'un seul attribut optimal.

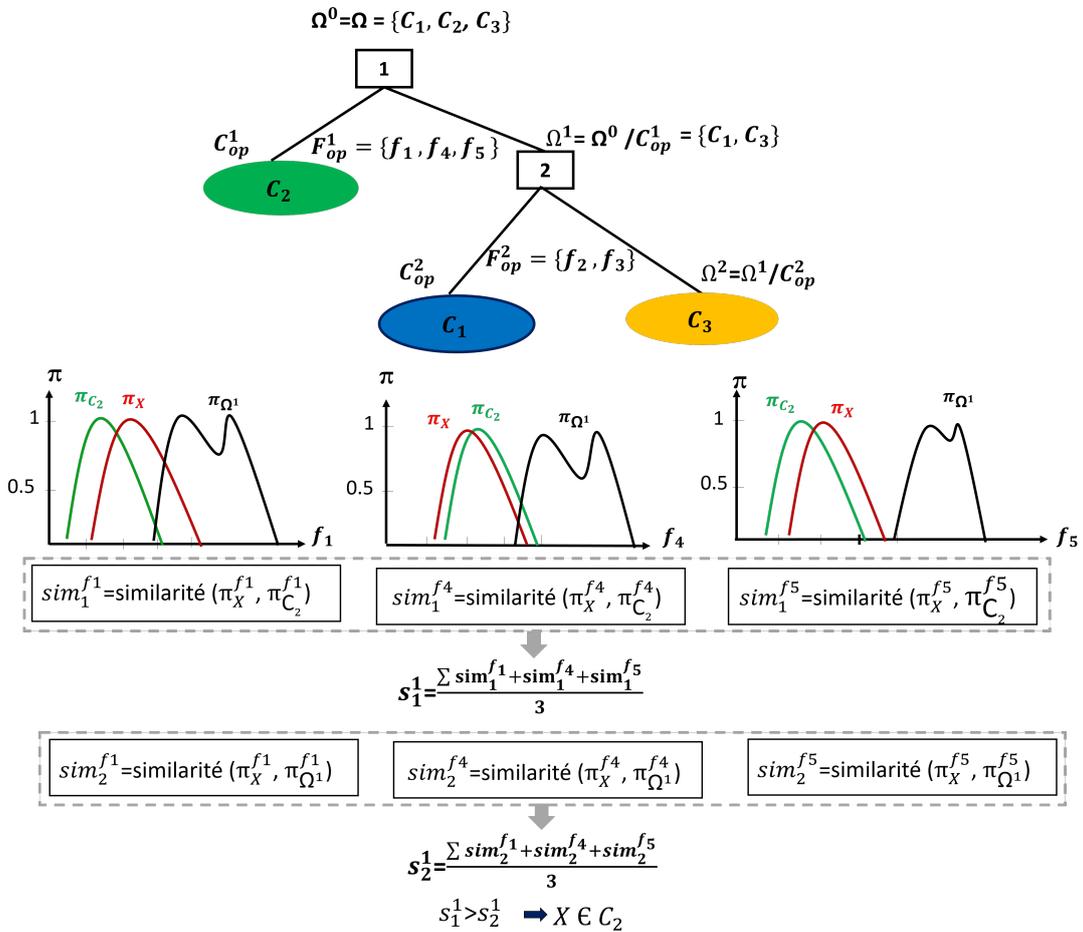


FIGURE II.16 – Classification par maximum de similarité entre le modèle de l'objet X et les modèles des classes

La Figure II.16 illustre un exemple d'univers fini de discours composé de 3 classes  $\Omega = \{C_1, C_2, C_3\}$ . Supposons qu'à la première itération, la classe  $C_2$  est la classe  $C_{op}^1$  la plus discriminée de l'ensemble  $\Omega^1 = \{C_1, C_3\}$  selon l'ensemble d'attributs  $F_{op}^1 = \{f_1, f_4, f_5\}$  et que à la deuxième itération, la classe  $C_1$  est discriminée de la classe  $C_3$  selon l'ensemble d'attributs  $F_{op}^2 = \{f_2, f_3\}$ . Afin de déterminer la classe d'un objet inconnu  $X$ , nous calculons d'abord le degré de similarité " $s_1^1$ " qui est égale à la moyenne des degrés de similarité entre  $\pi_X$  et  $\pi_{C_{op}^1}$  selon l'ensemble d'attributs  $F_{op}^1$ . Ensuite, nous calculons le degré de similarité " $s_2^1$ " qui est égale à la moyenne des degrés de similarité entre  $\pi_X$  et  $\pi_{\Omega^1}$  selon l'ensemble d'attributs  $F_{op}^1$ . Comme  $s_1^1 > s_2^1$ , l'objet  $X$  est évalué comme un objet de la classe  $C_{op}^1$ .

## II.3 Évaluation du CBH(1) et CBH(2)

### II.3.1 Classifieurs utilisés

Afin d'évaluer CBH(1) et CBH(2), nous nous concentrons d'abord sur la comparaison de leurs performances avec quatre classifieurs de nature différentes.

1. **Machines à vecteur de support (Support Vector Machine SVM)** : Il est très utilisé car il prend en compte à la fois l'erreur empirique et la complexité du classifieur. Dans notre stratégie, le noyau RBF est utilisé, et qui est défini comme suit :

$$RBF(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (\text{II.3})$$

Où :

$x_i$  : le vecteur de support

$x_j$  : un point de test

$\gamma$  : détermine la zone d'influence

2. **Le perceptron multicouche (Multilayer Perceptron MLP)** : C'est l'un des réseaux de neurones composé de plusieurs couches. Le flux d'informations de la couche d'entrée vers la couche de sortie est unidirectionnel.
3. **Classifieur J48** : C'est un arbre de décision implémenté en Java à partir de l'arbre de décision C4.5 [106]. La construction de l'arbre est basée sur l'entropie d'information. J48 utilise des critères de fractionnement pour trouver l'attribut offrant la meilleure distribution des données d'apprentissage. Généralement, l'attribut dont le gain d'information normalisé le plus élevé est utilisé.
4. **Classifieur bayésien naïf (Naïve Bayes NB)** : Il s'agit d'une classification probabiliste basée sur le théorème de Bayes. Il est dit naïf parce qu'il est basé sur une forte hypothèse

d'indépendance entre les attributs.

Ensuite, nous comparons les performances de CBH(1) et CBH(2) avec un classifieur possibiliste conventionnel. Considérons un univers de discours fini composé de  $N$  classes  $\Omega = \{C_1, \dots, C_N\}$ , et que chaque classe est représentée par  $K$  attributs. Comme le montre l'Équation II.4, la classification de possibilité conventionnelle consiste d'abord à calculer tous les degrés de possibilité  $\alpha_{k,i}$  pour qu'un objet  $X$  soit de classe  $C_i$  selon l'attribut  $f_k$ . Ensuite, pour chaque classe  $C_i$ , nous calculons la valeur minimale des  $\alpha_{k,i}$  afin d'attribuer un degré  $\Pi_i$  à l'objet  $X$ , qui représente sa degré de possibilité d'être un objet de la classe  $C_i$ . La classe ayant le degré de possibilité le plus élevé est retenue comme la classe à laquelle appartient  $X$ .

$$\Pi_i (X) = \max_{i=1}^N (\min_{k=1}^K (\alpha_{k,i})) \quad (\text{II.4})$$

### II.3.2 Bases de données utilisées

Huit base de données du référentiel UCI ont été considérées [98]. Elles ont été choisies en fonction de la nature numérique de leurs attributs. Les détails de chaque base de données sont présentés, par ordre alphabétique, dans le Tableau II.3. Pour chaque base de données, 60 % de ses échantillons sont utilisés pour la phase d'apprentissage et 40 % de ses échantillons sont utilisés pour la phase de test.

TABLE II.3 – Détails des bases de données utilisées.

Base de données	Instances	Attributs	Classes
Colonne vertébrale	310	6	3
Diabètes	768	9	2
Graines	210	8	3
Ionosphère	351	34	2
Iris	150	5	3
Véhicule	946	18	4
Verre	214	10	7
Vin	178	13	3

1. **Colonne vertébrale** : Il s'agit d'une base de données biomédicale établie par Dr Henrique da Mota. Les données sont organisées en trois catégories : normal (100 cas), spondylolisthésis (150 cas) et hernie discale (60 cas).
2. **Diabètes** : Certaines restrictions sont imposées au choix des diabètes. En fait, tous les diabétiques présentés dans cette base de données sont des femmes d'origine indienne Pima, âgées d'au moins 21 ans.
3. **Graines** : Cette base de données comprend trois variétés de blé différentes : Canada, Rosa, et Kama, dont chacune est représentée par 70 instances. La structure interne du

noyau cellulaire a été identifiée par le biais de la technologie des rayons X. Par rapport à d'autres technologies d'imagerie plus sophistiquées (comme le microscope à balayage ou la technologie laser), elle est non destructive et beaucoup moins chère.

4. **Ionosphère** : Il s'agit d'une base de données d'échos radar, qui ne contient que deux classes : "bon écho" et "mauvais écho". Un "bon" écho est un écho indiquant l'existence des électrons libres dans l'ionosphère, tandis qu'un "mauvais" écho ne fournit pas cette information. Les échos radar ont été obtenus à partir du système de Goose Bay, qui se compose de 16 antennes haute fréquences.
5. **Iris** : Cette base de données comprend 3 espèces de plantes d'iris, chacune contenant 50 instances.
6. **Véhicule** : La base de données de véhicules est utilisée pour classer une silhouette donnée dans l'un des quatre types de véhicules à l'aide d'un ensemble d'attributs extraits.
7. **Verre** : La recherche sur la classification des types de verre est guidée par des enquêtes criminologiques. Sur les scènes du crime, si les verres laissés sont correctement identifiés, ils peuvent être utilisés comme preuve.
8. **Vin** : Cette base est le résultat de l'analyse chimique de trois variétés de vin de la même région en Italie. La teneur de chacun des trois est caractérisée par 13 ingrédients.

### II.3.3 Logiciel utilisé

WEKA est une application open source développée par l'Université de Waikato en Nouvelle-Zélande, qui contient un ensemble d'outils avancés de prétraitement de données, de sélection d'attributs, ainsi que de classifieurs d'apprentissage automatique [128]. Les fichiers utilisés pour cette plateforme sont des fichiers ARFF, qui se composent de deux parties : l'entête et la partie données. La plateforme Weka présente de nombreux avantages :

- Elle est livrée avec une licence libre sous la licence publique générale GNU.
- Elle offre divers outils d'exploration de données, tels que le prétraitement des données, la sélection d'attributs, la classification, le regroupement et la régression.
- Elle est portable car elle est implémentée en Java.
- Elle est pratique en raison de son interface graphique.

### II.3.4 Mesures statistiques de pertinence

Les résultats ont été évalués en fonction de 6 mesures statistiques de pertinence telles que le taux des vrais positives "VP", le taux des faux positives "FP", la précision, le taux de rappel,

la mesure F et le coefficient MCC.

Considérons que l'univers de discours est composé de N classes  $\Omega = \{C_1, \dots, C_N\}$ , nous notons  $VP_i$ ,  $FP_i$ ,  $FN_i$  et  $VN_i$  définissent respectivement "Vrai Positives", "Faux Positives", "Faux Négatives" et "Vrai Négatives" d'une classe  $C_i$ .

— *Précision* :

$$Précision = \frac{1}{N} \sum_{i=1}^N \frac{VP_i}{VP_i + FP_i} \quad (II.5)$$

— *Rappel* :

$$Rappel = \frac{1}{N} \sum_{i=1}^N \frac{VP_i}{VP_i + FN_i} \quad (II.6)$$

— *Mesure F* : La mesure F inclut à la fois la précision et le taux de rappel. Elle peut être considérée comme la moyenne pondérée de ces deux mesures.

$$F = \frac{2}{N} \sum_{i=1}^N \frac{Précision_i * Rappel_i}{Précision_i + Rappel_i} \quad (II.7)$$

— Coefficient de Matthew (MCC) : Ce coefficient est utilisé pour mesurer la qualité de la classification.

$$MCC = \frac{1}{N} \sum_{i=1}^N \frac{VP_i * VN_i - FP_i * FN_i}{\sqrt{(VP_i + FP_i)(VP_i + FN_i)(VN_i + FP_i)(VN_i + FN_i)}} \quad (II.8)$$

### II.3.5 Résultats et discussions

Afin d'évaluer l'efficacité de CBH(1) et CBH(2), nous avons comparé leurs performances à d'autres classifieurs. La comparaison est faite selon les mesures statistiques susmentionnées.

À partir des tableaux et figures présentés ci-dessous, nous pouvons tirer les conclusions suivantes :

- Le CBH(2) performe généralement tous les autres classifieurs.
- Le CBH(2) et le CBH(1) évalué par similarité ont les mêmes mesures de performance pour les bases de données : colonne vertébrale, graines et iris.
- Le CBH(1) évalué par similarité est le deuxième classifieur le plus performant, après le CBH(2) pour les bases de données : diabète, inosphère, véhicule, verre et vin.
- Le CBH(1) évalué par similarité est plus performant que le CBH(1) évalué par projection.
- Les classifieurs CBH(2) et CBH(1) (évalué par similarité et celui évalué par projection) performant l'arbre de décision conventionnel C4.5. Ainsi, l'approche de décomposition OVA, qu'on a adoptée, est meilleure que celle d'un arbre de décision.

- Les classifieurs CBH(2) et CBH(1) (évalué par similarité) performant les classifieurs conventionnels SVM et MLP. En fait nous avons considéré l'incertitude des données le long de la phase de construction des classifieurs.
- Le classifieur CBH(1) évalué par projection performe généralement les classifieurs MLP et SVM. Cependant, pour la base de données ionosphère et verre, les performances de CBH(1) sont légèrement inférieures à celles de SVM.
- Les classifieurs CBH(2) et CBH(1) performant le NB. Ainsi, la modélisation possibiliste est plus adéquate qu'une modélisation probabiliste pour gérer des données incertaines (incertitude épistémique).
- La classification binaire hiérarchique utilisant uniquement des attributs sélectifs donne de meilleurs résultats que la classification possibiliste conventionnelle (Cf. Figure II.25).

TABLE II.4 – Évaluation de la base de données Colonne vertébrale.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,831	0,847	0,871	0,855	0,893	0,932	0,932
FP	0,079	0,056	0,057	0,063	0,031	0,058	0,058
Précision	0,840	0,870	0,885	0,868	0,905	0,932	0,932
Rappel	0,831	0,847	0,871	0,855	0,893	0,932	0,932
Mesure F	0,818	0,853	0,869	0,852	0,891	0,932	0,932
MCC	0,762	0,786	0,822	0,800	0,837	0,88	0,88

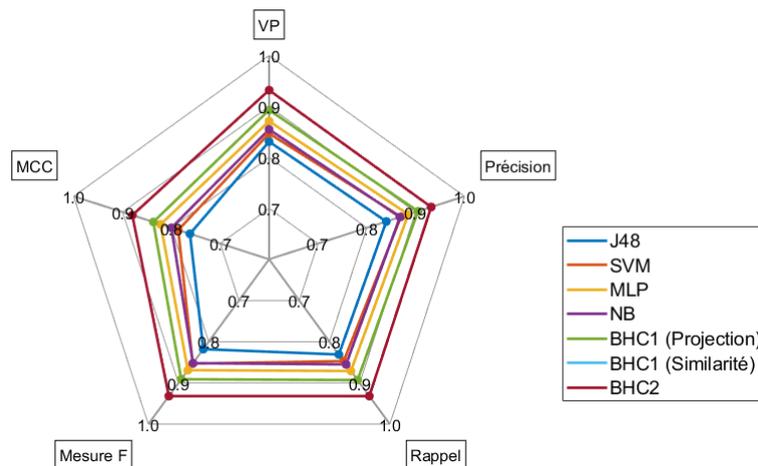


FIGURE II.17 – Performances des classifieurs sur la base de données colonne vertébrale

TABLE II.5 – Évaluation de la base de données Diabètes.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,723	0,781	0,795	0,775	0,921	0,992	1
FP	0,474	0,286	0,262	0,290	0,022	0,018	0
Précision	0,733	0,778	0,792	0,772	0,92	0,992	1
Rappel	0,723	0,781	0,795	0,775	0,921	0,992	1
Mesure F	0,682	0,778	0,793	0,773	0,920	0,992	1
MCC	0,345	0,51	0,542	0,496	0,703	0,982	1

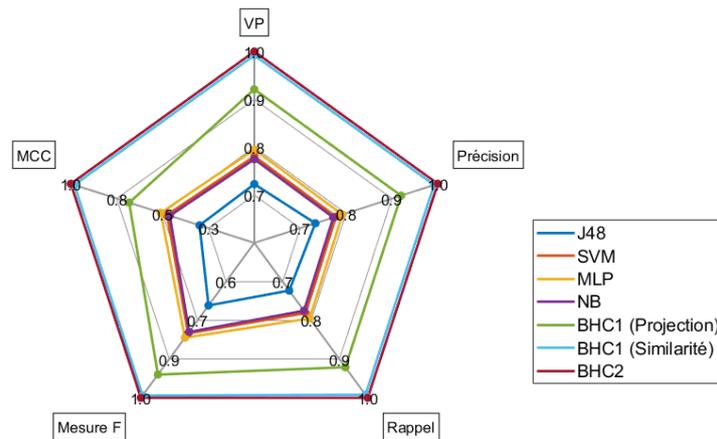


FIGURE II.18 – Performances des classifieurs sur la base de données Diabètes

TABLE II.6 – Évaluation de la base de données Graines.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,881	0,893	0,881	0,821	0,903	0,921	0,921
FP	0,060	0,054	0,060	0,089	0,054	0,042	0,042
Précision	0,885	0,900	0,903	0,853	0,904	0,931	0,931
Rappel	0,881	0,893	0,881	0,821	0,903	0,921	0,921
Mesure F	0,882	0,895	0,884	0,826	0,903	0,920	0,920
MCC	0,823	0,842	0,833	0,748	0,860	0,887	0,887

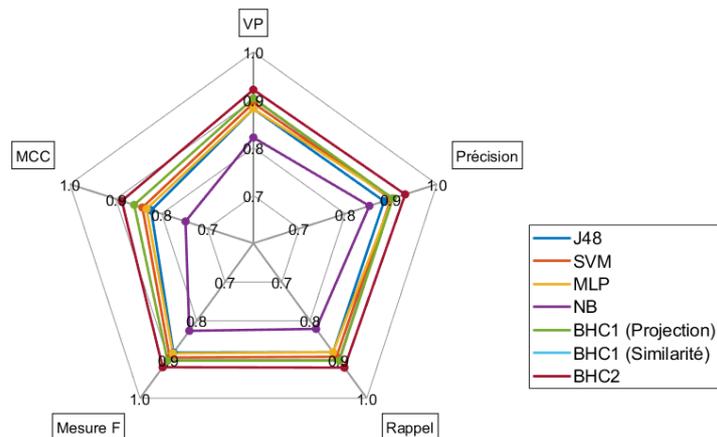


FIGURE II.19 – Performances des classifieurs sur la base de données Graines

TABLE II.7 – Évaluation de la base de données Ionosphère.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,929	0,957	0,921	0,943	0,953	0,96	0,973
FP	0,049	0,068	0,141	0,058	0,047	0,041	0,057
Précision	0,937	0,958	0,930	0,944	0,953	0,96	0,974
Rappel	0,929	0,957	0,921	0,943	0,953	0,96	0,973
Mesure F	0,930	0,957	0,919	0,943	0,953	0,96	0,973
MCC	0,856	0,907	0,834	0,877	0,886	0,92	0,939

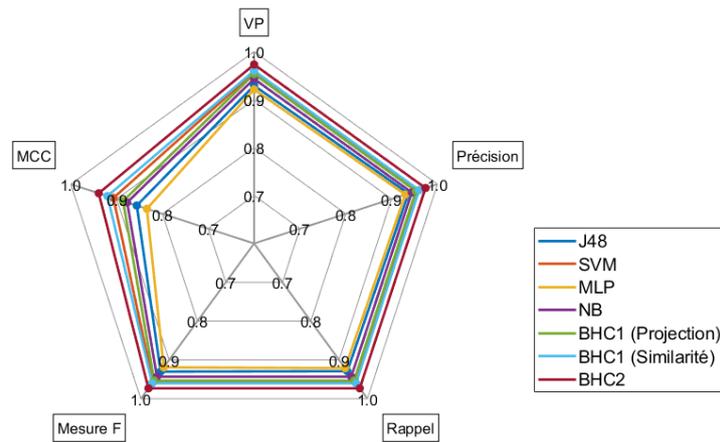


FIGURE II.20 – Performances des classifieurs sur la base de données Ionosphère

TABLE II.8 – Évaluation de la base de données Iris.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,950	0,96	0,967	0,983	0,994	1	1
FP	0,025	0,02	0,017	0,008	0,002	0	0
Précision	0,957	0,96	0,967	0,984	0,995	1	1
Rappel	0,95	0,96	0,967	0,983	0,994	1	1
Mesure F	0,951	0,96	0,967	0,983	0,994	1	1
MCC	0,928	0,94	0,950	0,975	0,988	1	1

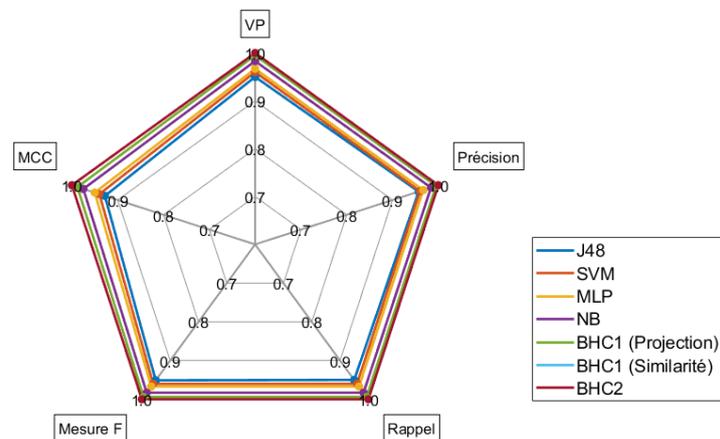


FIGURE II.21 – Performances des classifieurs sur la base de données Iris

TABLE II.9 – Évaluation de la base de données Véhicule.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,728	0,828	0,817	0,731	0,834	0,887	0,907
FP	0,092	0,058	0,062	0,091	0,055	0,038	0,031
Précision	0,721	0,828	0,822	0,718	0,831	0,895	0,909
Rappel	0,728	0,828	0,817	0,731	0,834	0,887	0,907
Mesure F	0,719	0,826	0,817	0,717	0,832	0,885	0,903
MCC	0,633	0,782	0,757	0,633	0,775	0,852	0,863

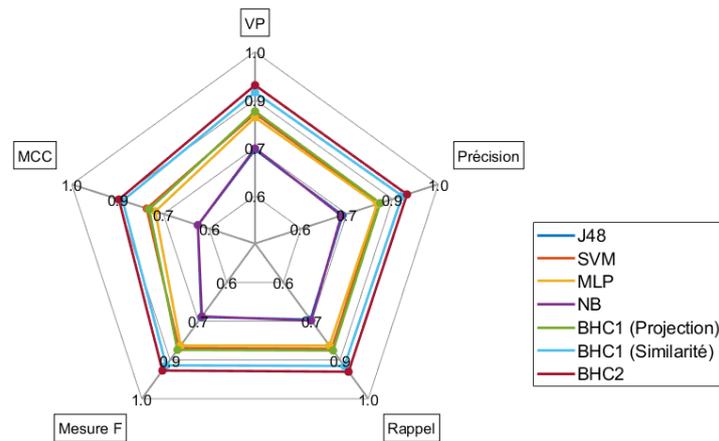


FIGURE II.22 – Performances des classifieurs sur la base de données Véhicule

TABLE II.10 – Évaluation de la base de données Verre.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,841	0,953	0,907	0,860	0,944	0,953	0,964
FP	0,064	0,014	0,042	0,058	0,018	0,014	0,050
Précision	0,840	0,953	0,910	0,857	0,944	0,953	0,965
Rappel	0,841	0,953	0,907	0,860	0,944	0,953	0,964
Mesure F	0,838	0,952	0,905	0,856	0,944	0,952	0,965
MCC	0,782	0,940	0,873	0,804	0,925	0,940	0,954

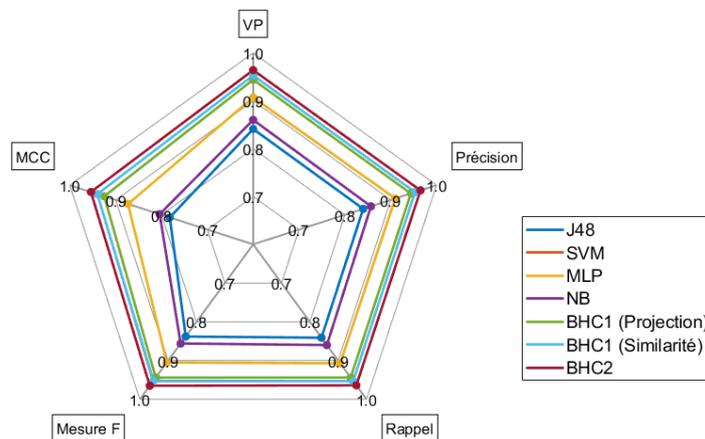


FIGURE II.23 – Performances des classifieurs sur la base de données Verre

TABLE II.11 – Évaluation de la base de données Vin.

	J48	SVM	MLP	NB	Évaluation de CBH(1) par projection	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,857	0,957	0,971	0,961	0,971	0,979	0,986
FP	0,091	0,022	0,014	0,19	0,015	0,014	0,004
Précision	0,886	0,957	0,974	0,962	0,972	0,980	0,987
Rappel	0,857	0,957	0,971	0,961	0,971	0,979	0,986
Mesure F	0,854	0,957	0,971	0,961	0,971	0,979	0,986
MCC	0,791	0,933	0,956	0,946	0,956	0,968	0,978

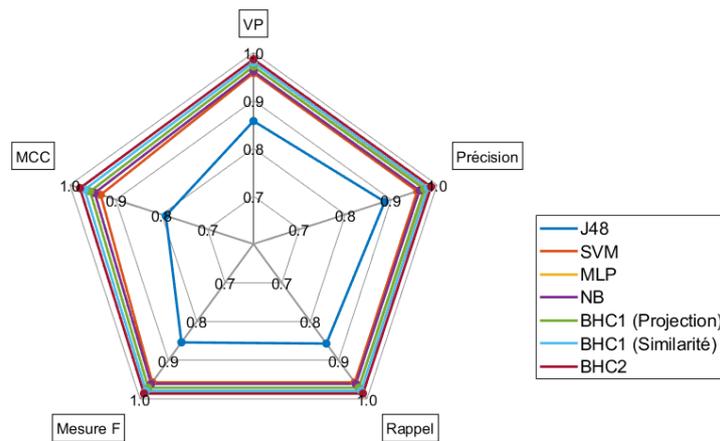


FIGURE II.24 – Performances des classifieurs sur la base de données Vin

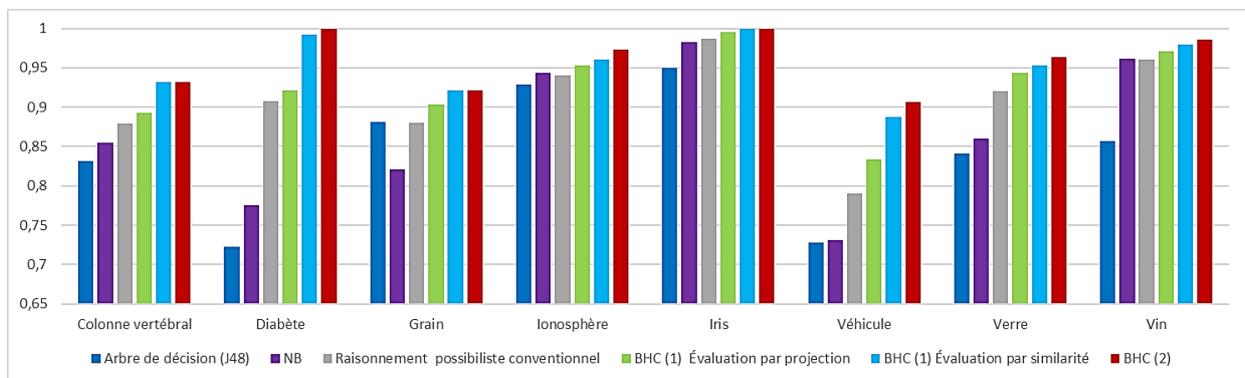


FIGURE II.25 – Analyse comparative basée sur le taux correct de classification

## II.4 Conclusion

Il convient de noter que lors de la modélisation des données, l'utilisation de tous les attributs augmente la dimensionnalité, augmentant ainsi la complexité du système. Par conséquent, il est intéressant de proposer de approches efficaces pour sélectionner les attributs dans un cadre possibiliste. Pour cette raison, dans ce chapitre, nous proposons en premier lieu trois approches de sélection d'attributs. Pour ces trois approches, le choix d'attributs est basé sur

leur pouvoir discriminant. Pour la première approche de sélection d'attributs, le pouvoir discriminant "DP\_SUMF" est déterminé en fonction des informations de niveau perceptif (variance inter, variance intra) et de niveau sub-sémantique (spécificité inter, spécificité intra, dissimilarité entre les distributions de possibilités), tandis que les pouvoirs discriminants "DP\_PMSC" et "DP\_PMCC" sont uniquement basés sur des informations de niveau sub-sémantique (la hauteur maximale des points d'intersection des distributions de possibilités, la mesure de spécificité d'une distribution de possibilité, l'aire du consensus des distributions de possibilités).

Dans la suite de ce chapitre, un classifieur binaire hiérarchique monocritère CBH(1) et un classifieur binaire hiérarchique multicritère CBH(2) ont été proposés pour simplifier le problème d'entrée massive pour les applications de classification multi-classes, tout en tenant en compte de l'incertitude liée aux attributs de l'ensemble de données. En fait, dans les applications pratiques, ce type d'imperfection est inévitable. La méthode de décomposition utilisée pour la construction de CBH(1) et CBH(2) est basée sur la technique OVA grâce à sa simplicité. Afin d'évaluer leurs performances, nous avons utilisé 6 mesures statistiques, à savoir le taux des vrais positives, le taux des faux positives, la précision, le taux de rappel, la mesure F et le coefficient MCC. En termes de précision de classification, la modélisation possibiliste avec une architecture binaire hiérarchique est bien meilleure que la modélisation probabiliste, encore meilleure que la méthode de classification possibiliste conventionnelle. De plus, nous pouvons conclure que la modélisation possibiliste est plus appropriée pour gérer l'incertitude épistémique que la modélisation probabiliste.

Dans le chapitre suivant, les deux classifieurs CBH(1) et CBH(2) sont appliqués à des données ultrasonores réelles.

# Chapitre III

## Interprétation de scène basée sur un capteur ultrason

### Sommaire

---

III.1 Approche de détection et d'évitement d'obstacles . . . . .	69
III.2 Classification binaire : Catégorisation de la rigidité de l'objet détecté en Dur/Mou . . . . .	71
III.3 Classification multi-classes : Reconnaissance de cinq matériaux . .	86
III.4 Conclusion . . . . .	99

---

## Introduction

Les systèmes à ultrasons sont largement utilisés dans les environnements intérieurs pour assurer une navigation autonome et sécurisée. La popularité de ces systèmes est due à leur prix modéré, leur robustesse et leur fiabilité. Généralement, les systèmes à ultrasons sont des télémètres mesurant la distance en se basant sur le temps d'arrivée de l'onde ultrasonore.

Afin de mieux comprendre les échos ultrasonores, plusieurs travaux ont été menés pour extraire plus d'informations que la simple mesure de la distance. Win *et al.* [129] ont proposé un système, pour identifier la forme des obstacles, basé sur trois capteurs à ultrasons, tandis que Ohtani et Baba [130] ont développé un robot équipé de 16 capteurs ultrasons disposés linéairement. Dans [131, 132], les auteurs ont montré qu'à partir de l'analyse de l'amplitude de l'écho ultrasonore, le coin peut être distingué du mur. Leur système a un taux de reconnaissance de 88 % sur les courtes distances. Cependant, le taux diminue à mesure que la distance augmente. Visvanathan *et al.* [133] ont proposé un système, équipé de cinq capteurs à ultrasons placés le long d'une ligne droite, pour extraire les caractéristiques communes du mur, à savoir les segments et les bords.

Ces systèmes à ultrasons, utilisent généralement plusieurs capteurs et opèrent sur des trajets prédéfinis. Les tâches assurées dans les travaux de l'état de l'art concernent en premier lieu la détection et la localisation des obstacles et la détermination de leur forme, en deuxième lieu. Nous proposons dans ce travail de tester la capacité de l'onde ultrasonore dans d'autres terrains non explorés, spécialement celui de la reconnaissance de la rigidité des matériaux.

Dans ce chapitre, en plus de la détection et de la localisation d'objets en profondeur, nous présentons deux approches l'une statistique et l'autre possibiliste pour identifier la rigidité des matériaux des objets détectés à partir de l'écho ultrasonore (Cf. Figure III.1).

### III.1 Approche de détection et d'évitement d'obstacles

Suite à un balayage horizontal d'une scène par un capteur ultrason, celui-ci produit un écho, qui dépend des objets qui l'ont réfléchi. À partir de cet écho et en se basant sur la mesure de distance déterminée en fonction du temps d'arrivée de l'écho ultrasonore, un signal de profil peut être construit, offrant les différents niveaux de profondeur de la scène scannée.

Intuitivement, un changement brusque du signal de profil, doit correspondre à un passage d'un objet à un autre. Par conséquent, il est recommandé de segmenter ce signal en se basant sur la détection de ses points de changement brusque [134]. La segmentation du signal peut se

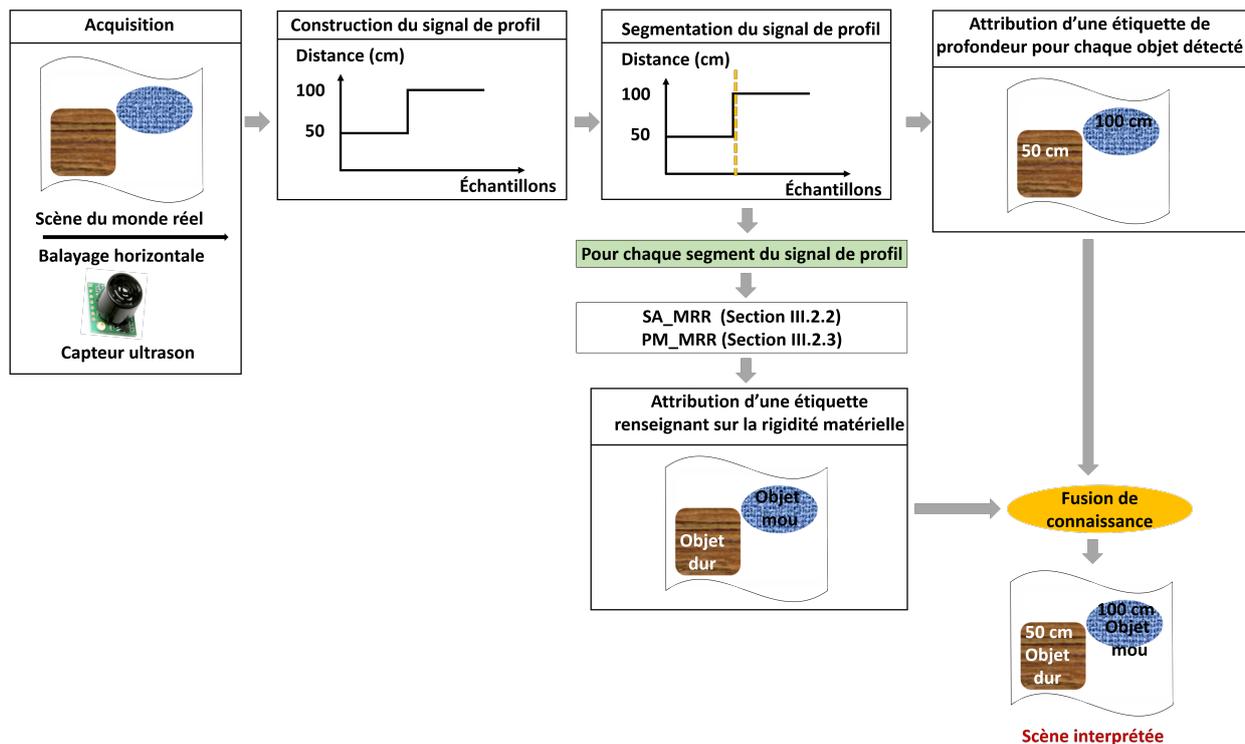


FIGURE III.1 – Interprétation de scène observée par un capteur ultrason

faire par un simple seuillage. Cependant, cette méthode s'avère inefficace car le seuil n'est pas dynamique. La segmentation peut être réalisée en appliquant le gradient, mais cette technique nécessite une étape de prétraitement du signal pour éliminer le bruit. Parmi les techniques de segmentation, il existe également, la transformée en ondelettes. Le résultat de cette technique est de générer deux signaux, un signal "Approximation" ne représente que les informations de basse fréquence, et un autre signal "Détail" représentent les informations de haute fréquence. Généralement, le bruit se trouve dans les basses fréquences du signal, tandis que les changements brusques du signal apparaissent dans les hautes fréquences.

La Figure III.2 illustre 4 scènes contenant un obstacle placé à différentes distances (à gauche, à droite, en face) où le fond de chaque scène est un mur à un mètre de cet obstacle. La distance entre le capteur ultrason et l'obstacle est d'environ vingt centimètres. La segmentation du signal de profil est basée sur la technique de la transformation en ondelettes. Ensuite, un seuil a été appliqué au signal "Détail" pour restaurer les positions des pics (les maximum locaux). Après plusieurs tests sur différentes formes d'ondelettes, l'ondelette mère utilisée dans ce travail est Daubechies db2 [135]. De la Figure III.2(a) et III.2(b), on constate que le signal de profil construit suite d'un balayage de gauche à droite est le même que le signal de profil construit suite d'un balayage de droite à gauche, ce qui est dû à l'utilisation d'un accéléromètre et d'un gyroscope.

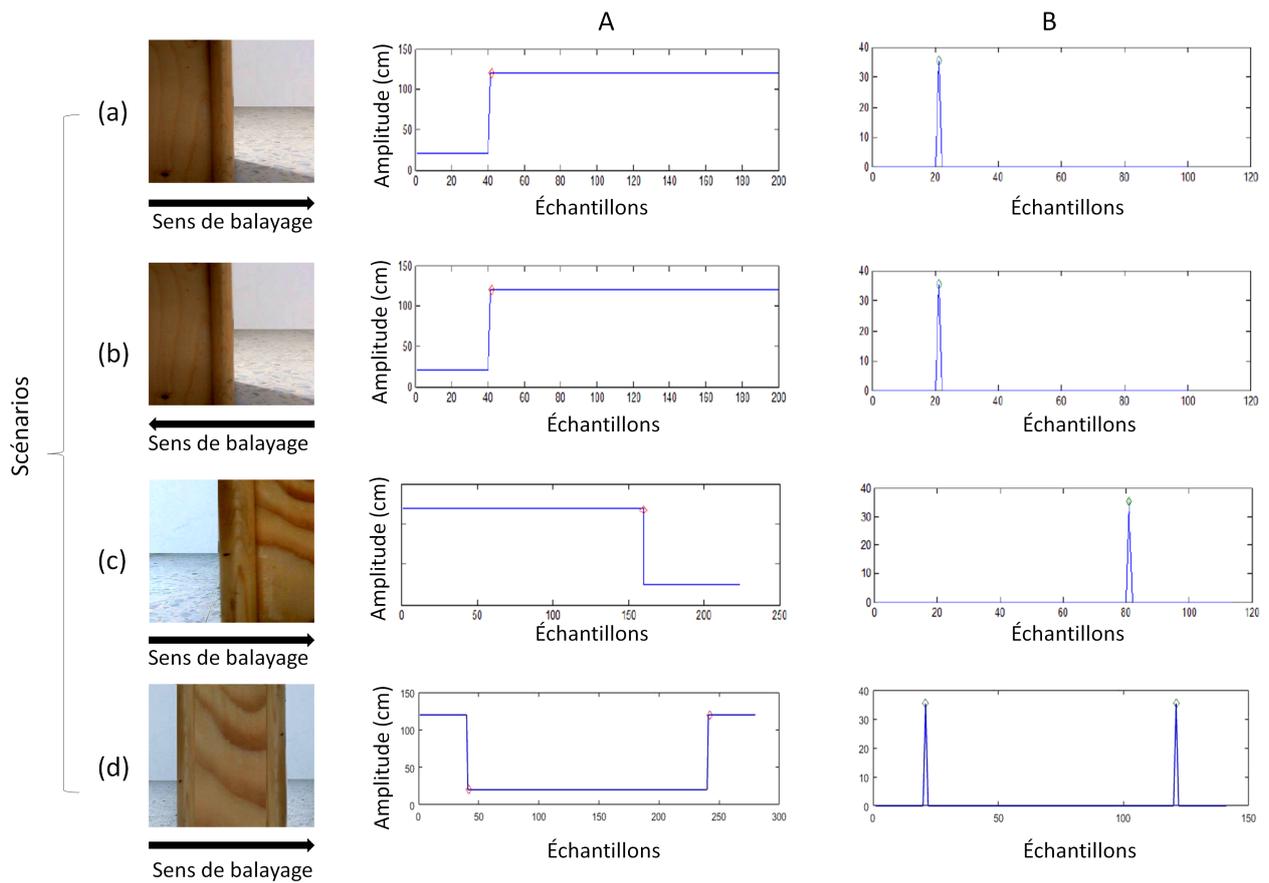


FIGURE III.2 – Différents scénarios de balayage d'une même scène (a,b,c,d), (A) leurs signaux de profil et (B) leurs détails d'ondelettes respectifs

Notre approche consiste à générer, en premier lieu, un signal de profil correspondant aux profondeurs des objets de la scène, puis à segmenter ce signal de façon à séparer les segments correspondants aux différents objets interceptés par l'onde ultrasonore. Enfin, chaque segment est analysé afin d'attribuer, à l'objet auquel il correspond, une étiquette spécifiant sa rigidité matérielle (Cf. Section III.2) et plus encore, le type de son matériaux (Cf. Section III.3).

## III.2 Classification binaire : Catégorisation de la rigidité de l'objet détecté en Dur/Mou

Les environnements intérieurs sont remplis d'objets de rigidités différentes à différents endroits. Afin de naviguer en toute sécurité dans un tel environnement, une simple détection d'obstacle sans connaître sa rigidité matérielle semble inefficace. En effet, les dégâts de collision dépendent fortement de la rigidité de matériaux des objets. En d'autre termes les dégâts causés par une collision avec un objet en laine sont très différents des dégâts causés par une collision avec un objet métallique. Ainsi, nous avons considéré deux classes pour spécifier la rigidité de matériaux des objets : une classe d'objets durs et une classe d'objets mous.

- **Approche statistique** : Les classes sont représentées par un vecteur d'attributs (Attributs de Haralick) extraits à partir d'une Distribution de Wigner Ville (DWV) des données brutes du capteur ultrason.
- **Approche possibiliste** : Les classes sont représentées par des modèles possibilistes (distributions de possibilités).

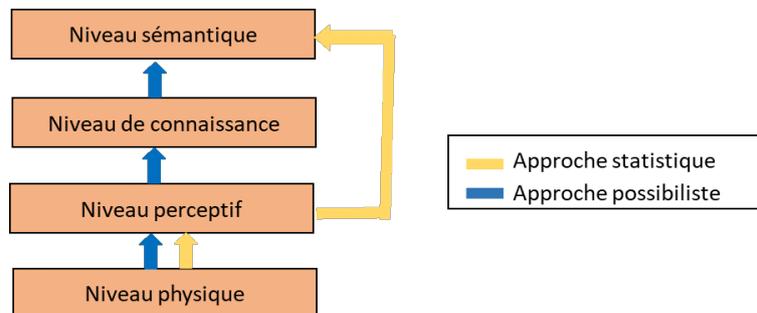


FIGURE III.3 – Approches proposées pour la reconnaissance de la rigidité de matériaux des objets

### III.2.1 Impact physique de la rigidité de matériaux des objets détectés sur l'écho ultrasonore

Lors de la détection d'un objet, une partie de l'onde incidente est diffusée, une autre partie est absorbée, et une autre partie est réfléchi (Cf. Figure III.4).

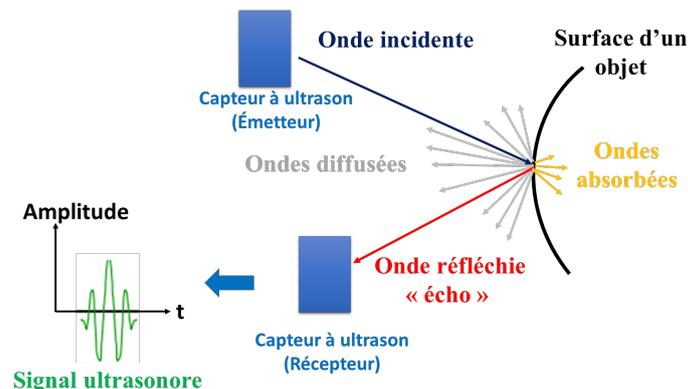


FIGURE III.4 – Différents phénomènes d'interaction de l'onde ultrasonore

L'amplitude de l'écho ultrasonore dépend fortement de la rigidité du matériau de l'objet détecté. En fait, plus l'objet est dur, plus l'amplitude de l'onde réfléchi est importante. En revanche, si l'objet détecté est mou, alors une grande partie de l'onde incidente est absorbée et diffusée, alors que l'onde restante est réfléchi avec une petite amplitude. La Figure III.5 présente l'impact de la rigidité de matériaux des objets détectés sur les ondes ultrasonores.

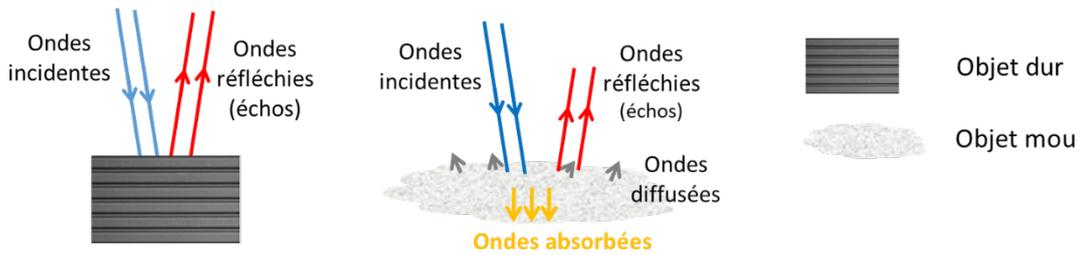


FIGURE III.5 – Impact de la rigidité de matériaux des objets détectés sur l'écho ultrasonore

L'onde ultrasonore est caractérisée par sa longueur d'onde " $\lambda$ ", sa fréquence " $f$ ", et son amplitude " $A$ ". Dans [136, 137], les auteurs ont montré respectivement dans la Figure 8 et la Figure 5 que l'amplitude " $A$ " décroît de façon exponentielle avec la distance " $x$ ". Selon le centre NDT, cette décroissance peut être définie comme [138] :

$$A_x = A_0 e^{-Kx} \quad (\text{III.1})$$

Où :

$A_0$  : Amplitude maximale au récepteur

$A_x$  : Amplitude maximale d'une particule à la distance  $x$  du récepteur

$K$  : Coefficient d'absorption

Sur la base de ces définitions, on peut conclure que l'amplitude de l'onde ultrasonore réfléchie dépend toutefois de la distance entre l'objet et le capteur ultrason et de la rigidité de matériaux des objets détectés.

## III.2.2 Approche statistique pour la reconnaissance de la rigidité matérielle (SA\_MRR : Statistical approach for material rigidity recognition)

### III.2.2.1 Description

Cette approche est dite statistique, car la prise de décision (attribution d'étiquette) est basée sur des vecteurs d'attributs statistiques. Comme le montre Figure III.6, l'approche proposée est basée sur deux phases :

**Phase d'apprentissage :** Cette phase consiste en premier lieu à convertir le signal ultrasonore en une représentation temps-fréquence selon DWV [139]. La Figure III.7 présente une DWV reflétée par un objet dur (Cf. Figure III.7(a)) et une autre reflétée par un objet mou (Cf. Figure III.7(b)). À partir de ces deux DWVs, il est clair qu'il existe une dissemblance de

texture entre eux. Néanmoins, il est impossible de tirer des conclusions significatives par une simple visualisation. Par conséquent, une étape d'extraction d'attributs s'avère nécessaire.

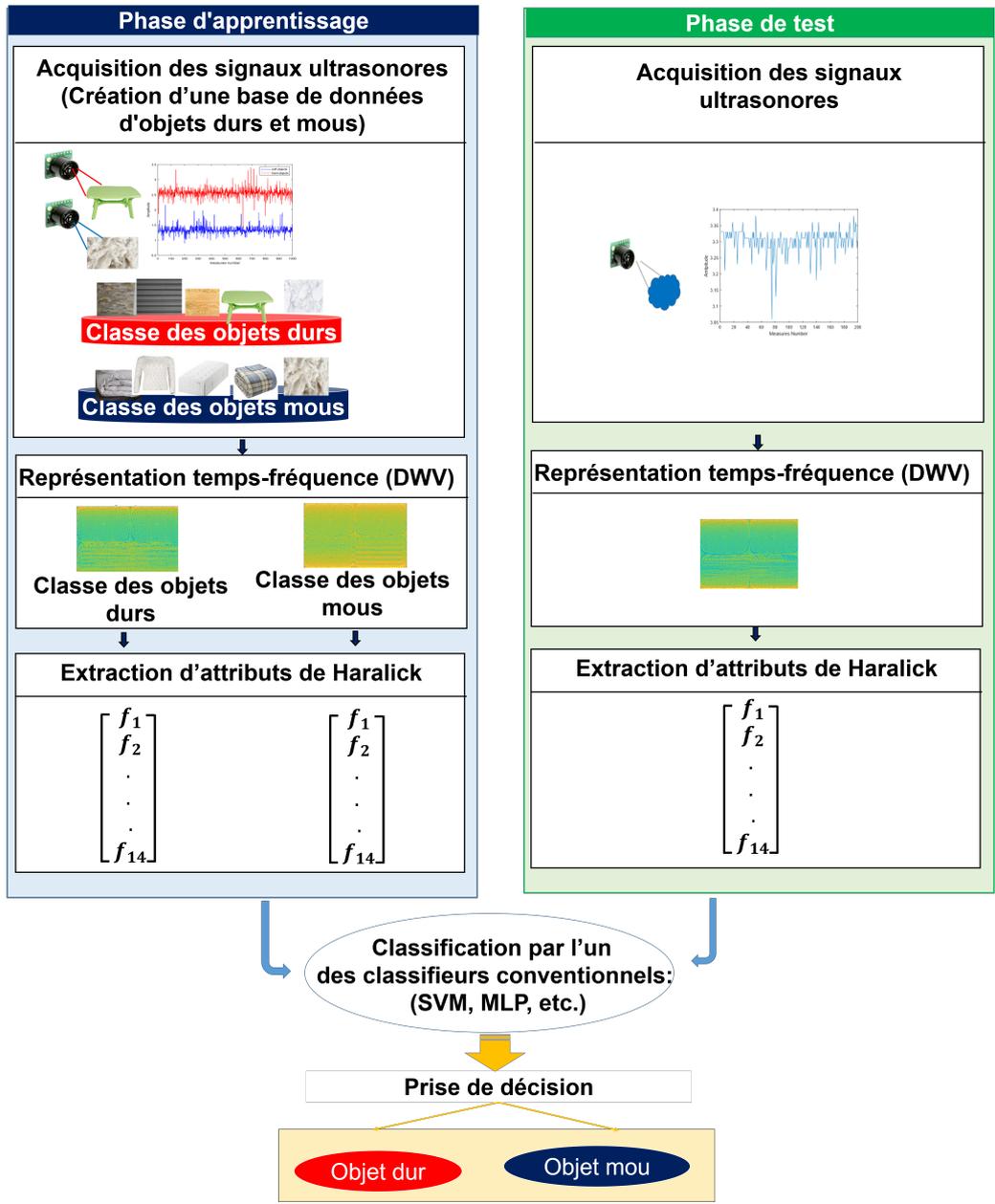


FIGURE III.6 – Approche statistique pour la reconnaissance de la rigidité des matériaux des objets détectés

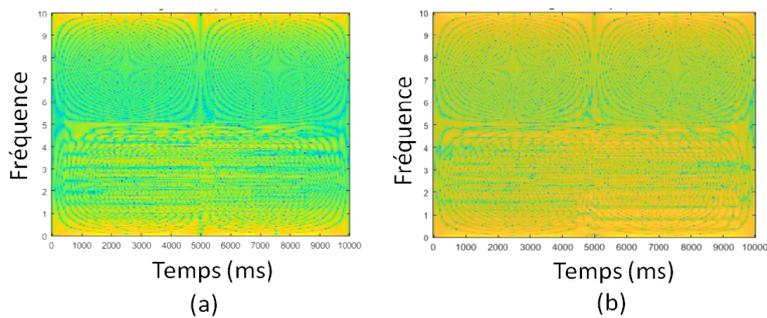


FIGURE III.7 – Exemples de représentation temps-fréquence de l'écho ultrasonore réfléchi par (a) un objet dur et (b) un objet mou

La méthode de Haralick est couramment utilisée pour l'extraction d'attributs liés à la texture de l'image. Cette méthode est généralement utilisée dans de nombreuses applications pratiques, notamment la caractérisation des types de cancer du poumon [140], la détection et la classification des cellules cancéreuses du côlon [141], la classification des cancers dans les images mammographies [142], etc.

La méthode de Haralick est basée sur l'extraction d'informations à partir de la matrice de co-occurrence de niveau de gris (GLCM) [143]. Le résultat de cette méthode est un vecteur de 14 attributs [144].

**Phase de test :** Comme la phase d'apprentissage, la phase de test consiste d'abord à convertir le signal ultrasonore en une représentation temps-fréquence, puis d'extraire les attributs de Haralick à partir de cette représentation. La prise de décision consiste à attribuer une étiquette "Objet dur" ou "Objet mou" à l'échantillon de test en se basant sur l'un des classifieurs conventionnels (SVM, MLP, etc.).

### III.2.2.2 Évaluation et discussions

L'évaluation de notre stratégie de détection des objets et de reconnaissance de leur rigidité matérielle est faite selon différents critères de classification, à savoir : la matrice de confusion, le taux de précision, la spécificité et la sensibilité.

#### 1. Construction d'une base de données :

À notre connaissance, aucune base de données de signaux ultrasonores n'est fournie pour identifier la rigidité matérielle des objets. Par conséquent, nous avons établi notre propre base de données en utilisant un capteur ultrason EZ2 et EZ4 de la famille LV-MaxSonar. Les largeurs de faisceau de ces deux capteurs sont présentées à la Figure III.8.

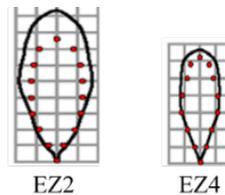


FIGURE III.8 – Largeurs des faisceaux des capteurs à ultrasons utilisés

Notre base de données est formée par 1000 échantillons, dont 500 échantillons sont réfléchis par des objets durs et 500 échantillons sont réfléchis par des objets mous. L'ensemble d'objets utilisés est résumé dans la Figure III.9.

La taille des échantillons acquis est de 200 mesures. La Figure III.10 présente l'amplitude d'un échantillon réfléchi par un objet dur et un autre réfléchi par un objet mou. Il est

bien clair que l'amplitude de l'écho ultrasonore réfléchi par un objet dur est supérieure à l'amplitude de l'écho ultrasonore réfléchi par un objet mou.

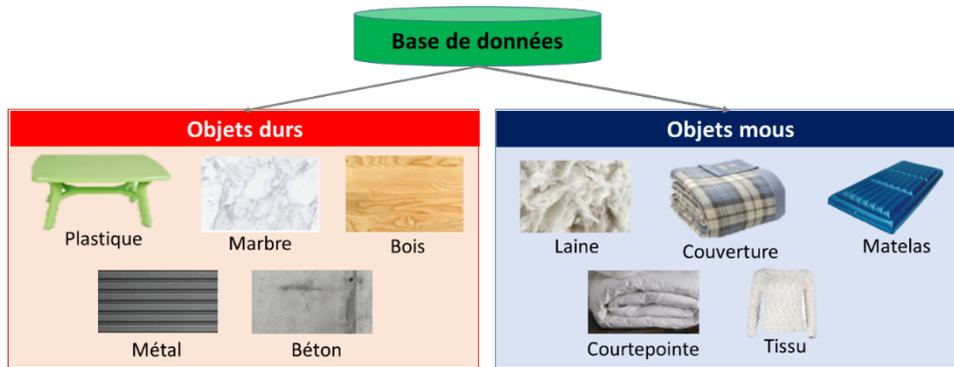


FIGURE III.9 – Objets utilisés pour la construction de la base de données

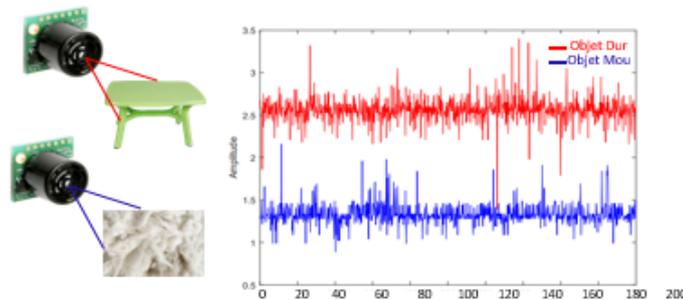


FIGURE III.10 – Exemples d'échos ultrasonores réfléchis par un objet dur et un objet mou

## 2. Expériences et résultats :

Pour évaluer notre approche, nous avons réalisé trois différents tests : (i) le premier consiste à montrer l'effet de la distance sur le signal ultrasonore, (ii) le deuxième test est utilisé pour montrer l'effet de la largeur du faisceau ultrasonore sur la détection des objets, et (iii) le troisième test consiste à montrer l'effet de la rigidité matérielle de l'obstacle détecté sur l'écho ultrasonore. Les trois classifieurs MLP, SVM et NB sont utilisés pour réaliser ces différents tests.

Afin de mettre en évidence l'objectif du premier test, nous avons construit une base de données à trois différentes distances, à savoir : 20 cm, 50 cm et 1 mètre. Comme le montre la Figure III.11, cette base de données est formée par 5 objets durs et 5 objets mous. Pour chaque objet à chaque distance, 60 échantillons sont utilisés dans la phase d'apprentissage et 40 échantillons sont utilisés dans la phase de test. En d'autres termes, pour chaque distance, 600 échantillons sont utilisés dans la phase d'apprentissage et 400 échantillons sont utilisés dans la phase de test.

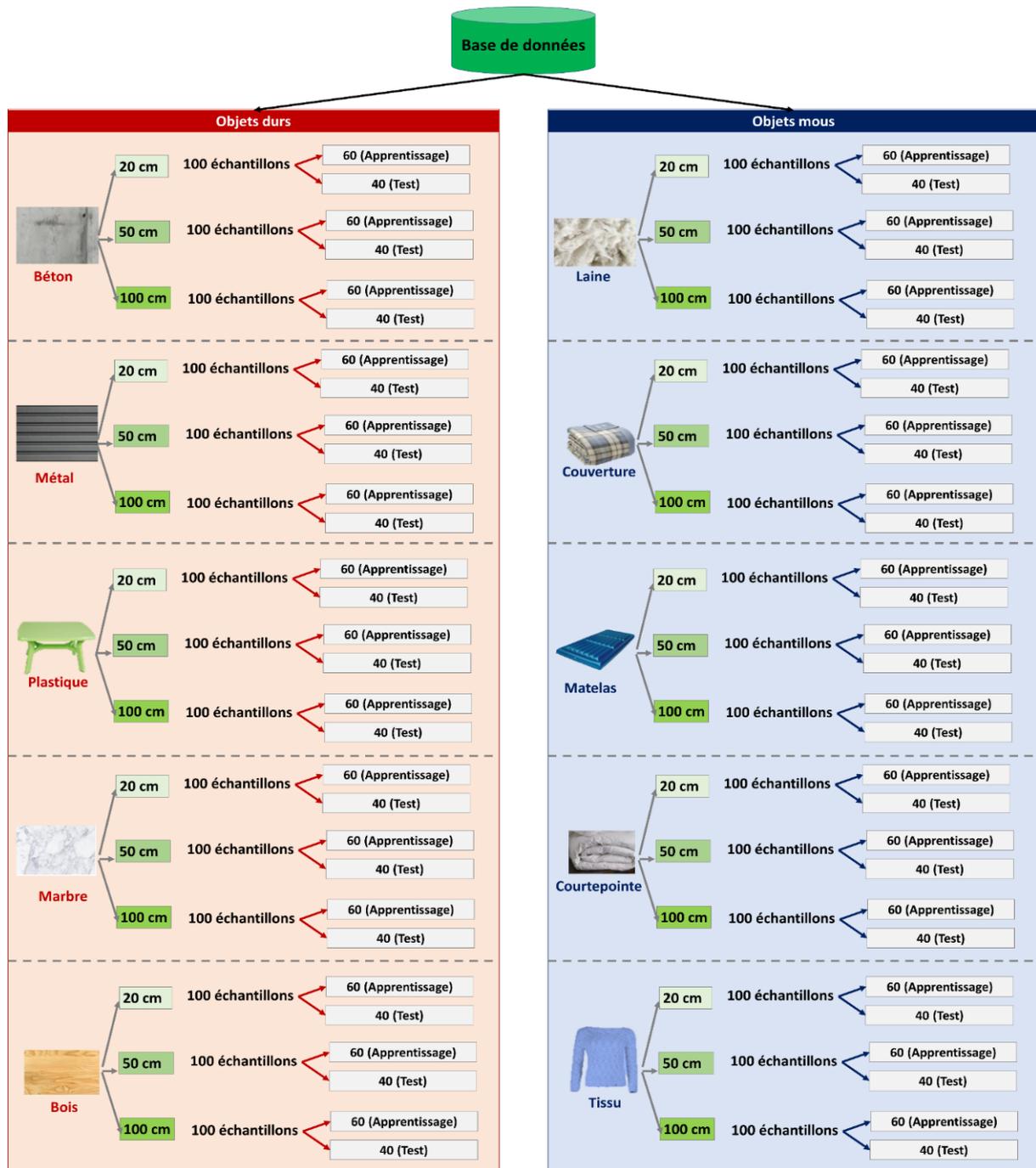


FIGURE III.11 – Structure de la base de données

Comme le montre la Figure III.12, les résultats obtenus permettent de déduire que :

- Plus la distance entre l'obstacle et le capteur ultrason est grande, plus la précision du système est faible. Cela confirme que l'amortissement de l'écho ultrasonore change en fonction de l'augmentation de la distance.
- Pour les courtes distances, le capteur ultrason EZ4 offre une meilleure précision que EZ2. Cependant, pour les longues distances, le capteur ultrason EZ2 offre une meilleure précision que EZ4. Cela confirme qu'il existe une forte corrélation entre la largeur du faisceau du capteur ultrason et la distance qui lui sépare de l'obstacle.

- Le capteur ultrason EZ2 est plus adapté à la détection d'objets distants, tandis que le capteur EZ4 est plus adapté à la détection d'objets à proximité.

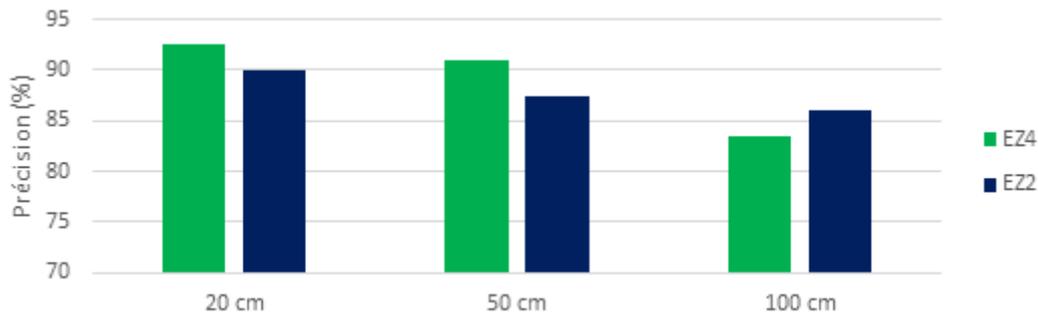


FIGURE III.12 – Impact de la distance sur l'écho ultrasonore

Un deuxième test, basé sur des classifieurs conventionnels, a été réalisé pour évaluer la performance de notre stratégie en utilisant les capteurs ultrasons EZ2 et EZ4 à une distance égale à 50 cm. À partir de la Figure III.13, on constate que pour les différents classifieurs utilisés, le capteur ultrason EZ4 offre une précision plus élevée que le capteur ultrason EZ2.



FIGURE III.13 – Comparaison de précision de EZ2 et EZ4 à une distance 50 cm

Le Tableau III.1 présente la matrice de confusion de différents classifieurs testés sur un ensemble de 400 échantillons (200 échantillons d'objets durs, 200 échantillons d'objets mous) acquis du capteur EZ4. Nous pouvons constater à partir de cette matrice de confusion que :

- Généralement, le taux de reconnaissance de la rigidité matérielle des objets mous est plus élevé que celui des objets durs.
- Le classifieur MLP donne le plus haut taux de reconnaissance d'objets mous. Cependant, ce classifieur a le taux de reconnaissance le plus bas pour les objets durs.
- La performance du classifieur SVM est relativement fiable.

TABLE III.1 – Matrice de confusion de chaque classifieur

Classifieur	Classe prédite	Dur	Mou
SVM	Dur	180	20
	Mou	16	184
MLP	Dur	170	30
	Mou	9	191
NB	Dur	175	25
	Mou	12	188

Étant donné que le meilleur cas est d'avoir une meilleure précision de reconnaissance des objets durs et des objets mous, et que cette situation n'est généralement pas offerte, nous utilisons d'autres métriques statistiques pour évaluer la performance de la stratégie proposée. En fait, afin d'évaluer la capacité du classifieur à prédire correctement les objets durs, nous définissons une métrique appelée taux de prédiction des objet durs (TPD) :

$$TPD = \frac{VD}{VD + FD} \quad (\text{III.2})$$

Où :

VD : représente les "Vrai Durs"

FD : représente les "Faux Durs"

De même, une autre métrique appelée taux de prédiction d'objets mous (TPM) est définie pour évaluer la capacité du classifieur à prédire correctement les objets mous.

$$TPM = \frac{VM}{VM + FM} \quad (\text{III.3})$$

Où :

VM : représente les "Vrai Mous"

FM : représente les "Faux Mous"

Sur la base du TPD et TPM, représentés dans le Tableau III.2, nous pouvons tirer les conclusions suivantes :

- Le meilleur TPD est fourni par le classifieur SVM.
- Le meilleur TPM est fourni par le classifieur MLP.

Les différents métriques d'évaluation utilisées (i.e. TPD et TPM, le taux de précision, la matrice de confusion) montrent la possibilité d'utiliser le classifieur SVM pour reconnaître les objets durs et les objets mous.

D'autres métriques statistiques peuvent également être utilisées, à savoir la spécificité et

TABLE III.2 – TPD et TPM pour chaque classifieur

Classifieur	TPD	TPM
SVM	0.9	0.92
MLP	0.85	0.955
NB	0.875	0.94

la sensibilité. La sensibilité détermine la capacité du classifieur à détecter les vrais positifs, tandis que la spécificité définit la capacité du classifieur à exclure les vrais négatifs.

$$\text{Spécificité} = \frac{VD}{VD + FM} \quad (\text{III.4})$$

$$\text{Sensibilité} = \frac{VM}{VM + FD} \quad (\text{III.5})$$

À partir des résultats illustrés à la Figure III.14, nous pouvons tirer les conclusions suivantes :

- Le classifieur SVM est le plus approprié pour la détection d'objets durs. En fait, il présente la sensibilité et le TPD le plus élevés.
- Le classifieur MLP permet d'obtenir le meilleur degré de spécificité et le meilleur TPM, ce qui confirme sa capacité à identifier correctement les objets mous.

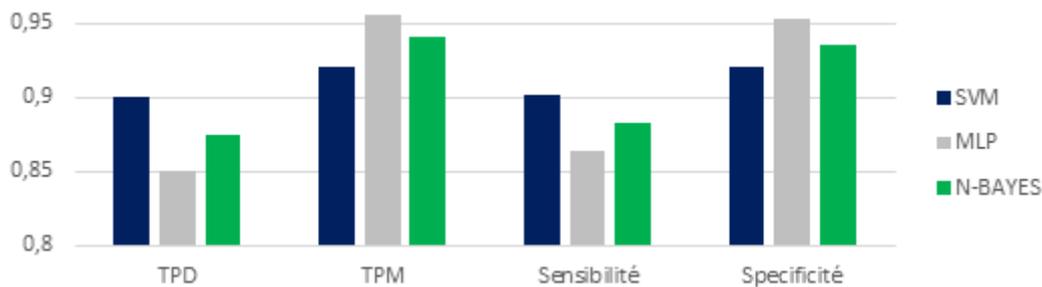


FIGURE III.14 – Taux des différentes métriques pour chaque classifieur

### III.2.3 Modélisation possibiliste pour la reconnaissance de la rigidité matérielle (PM\_MRR : Possibilistic modeling for material rigidity recognition)

Un système d'aide à la navigation doit averti l'utilisateur en temps réel pour éviter les collisions avec des objets nuisibles. Par conséquent, le temps de traitement doit être réduit autant que possible pour avoir un système en temps réel. Pour cela, nous avons considéré deux contraintes pour réduire le temps de traitement. Tout d'abord, un seul capteur ultrason est utilisé. Deuxièmement, une petite taille d'échantillon est acquise pour représenter l'écho

ultrasonore. Cela signifie que notre processus de reconnaissance de la rigidité matérielle n'est basé que sur un petit nombre de valeurs de données. Néanmoins, plus la taille d'échantillon est petite, plus l'information est ambiguë, inexacte, incertaine et incomplète.

L'utilisation d'attributs de Haralick a montré leur capacité à faire la distinction entre les objets durs et les objets mous. En fait, comme on a vu dans l'approche statistique, le taux de reconnaissance en utilisant le classifieur SVM avec tous les attributs de Haralick a dépassé 91%. Dans [145], il a été convenu d'utiliser l'entropie, la somme d'entropie et la variance après avoir observé que l'attribut " Variance" étant un attribut couramment utilisé pour mesurer le degré de dispersion statistique d'un ensemble de données et que l'attribut "Entropie" permet d'évaluer la stabilité et la quantité des données. Selon notre approche, l'entropie est déterminée à partir de DWV. L'entropie d'une image est un indicateur de sa complexité. En fait, plus l'image est "hétérogène", plus l'entropie est importante. De la Figure III.15, nous observons que :

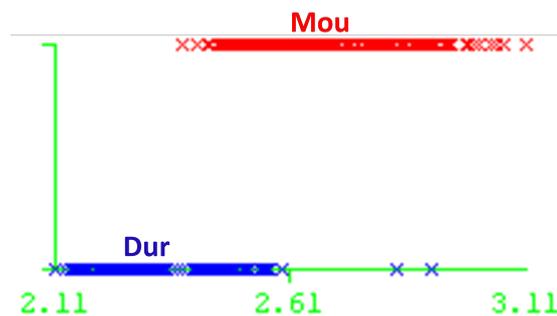


FIGURE III.15 – Les valeurs de mesures d'entropie

- Plus l'obstacle détecté est dur, plus l'écho ultrasonore est stable. Par conséquent, les valeurs mesurées d'entropie pour les matériaux durs sont faibles.
- Plus l'obstacle détecté est mou, plus l'écho ultrasonore est moins stable. Par conséquent, les valeurs mesurées d'entropie pour les matériaux mous sont plus élevées.

Il est à noter que le taux de reconnaissance est réduit à 89,5% en utilisant seulement trois attributs : variance, entropie et la somme d'entropie [145]. La diminution du taux de reconnaissance tout en réduisant le nombre d'attributs est tout à fait compréhensible puisque l'incertitude augmente en conséquence.

Afin de réduire la dimensionnalité tout en conservant des performances acceptables, il est nécessaire de prendre en compte les imperfections des données acquises. Dans notre cas, l'imperfection des données est due à deux raisons : (i) les propriétés physiques des ondes ultrasonores qui ne sont pas aussi bonnes que les ondes laser ou infrarouge dans la transmission d'informations, (ii) la brièveté du signal représentatif de l'écho ultrasonore acquis. Ces deux imperfections

sont liées au capteur et à la représentation de l'information, et font donc partie de l'incertitude épistémique (Cf. Section I.3.1.2). Dans ce contexte et comme il est montré dans le premier chapitre, la théorie de possibilités est le cadre le plus approprié non pas pour éliminer les imperfections de l'information, mais plutôt pour les garder à l'esprit lors de la modélisation des données. Pour cette raison, notre deuxième approche de reconnaissance de la rigidité de matériaux des objets est basée sur des modèles possibilistes (distributions de possibilités).

### III.2.3.1 Description

Comme le montre la Figure III.16, la deuxième approche proposée pour identifier la rigidité de matériaux des objets repose sur deux phases :

**Phase d'apprentissage :** Cette phase comprend cinq étapes. La première consiste à convertir les signaux ultrasonores acquis à partir des objets durs et des objets mous en des DWVs. Ensuite, nous extrayons, à partir des DWV, les attributs de Haralick. La troisième étape consiste à estimer la distribution de possibilités de chaque attribut de Haralick. Ensuite, une étape de sélection d'attributs est requise (Cf. Algorithme II.3). Enfin, deux cartes de possibilités " $\pi_{C_{Dur}}$ " et " $\pi_{C_{Mou}}$ " sont construites pour associer une signature à la classe des objets durs et à la classe des objets mous, respectivement.

**Phase de test :** Cette phase suit généralement les mêmes étapes que la phase d'apprentissage jusqu'à ce que la carte possibiliste relative à l'échantillon de test " $\pi_{C_{Test}}$ " est construite. La prise de décision consiste à attribuer une étiquette "Objet dur" ou "Objet mou" à la carte possibiliste correspondante au nouveau échantillon. Pour ce faire, nous mesurons le degré de similarité entre la carte " $\pi_{C_{Test}}$ " et les cartes établies à la phase d'apprentissage " $\pi_{C_{Dur}}$ " et " $\pi_{C_{Mou}}$ ". Ce degré indique si les cartes sont identiques, similaires ou différentes. La mesure de similarité utilisée est "info specificity" [92]. La décision finale considérée est la classe correspondante à la carte ayant le plus haut degré de similarité avec " $\pi_{C_{Test}}$ ".

### III.2.3.2 Évaluation et discussions

La base de données acquise par le capteur ultrason EZ4 à 50 cm est utilisée pour évaluer cette deuxième stratégie. Les signaux utilisés pour créer cette base de données sont de petite taille environ 200 mesures (Cf. Figure III.10). La base de données est constituée de 1000 échantillons, dont 500 échantillons utilisés pour créer une base de données d'objets durs et 500 échantillons pour créer une base de données d'objets mous. Chacune de ces bases de données est partagée entre apprentissage et test avec les pourcentages respectifs 60% et 40% (Cf. Figure III.17).

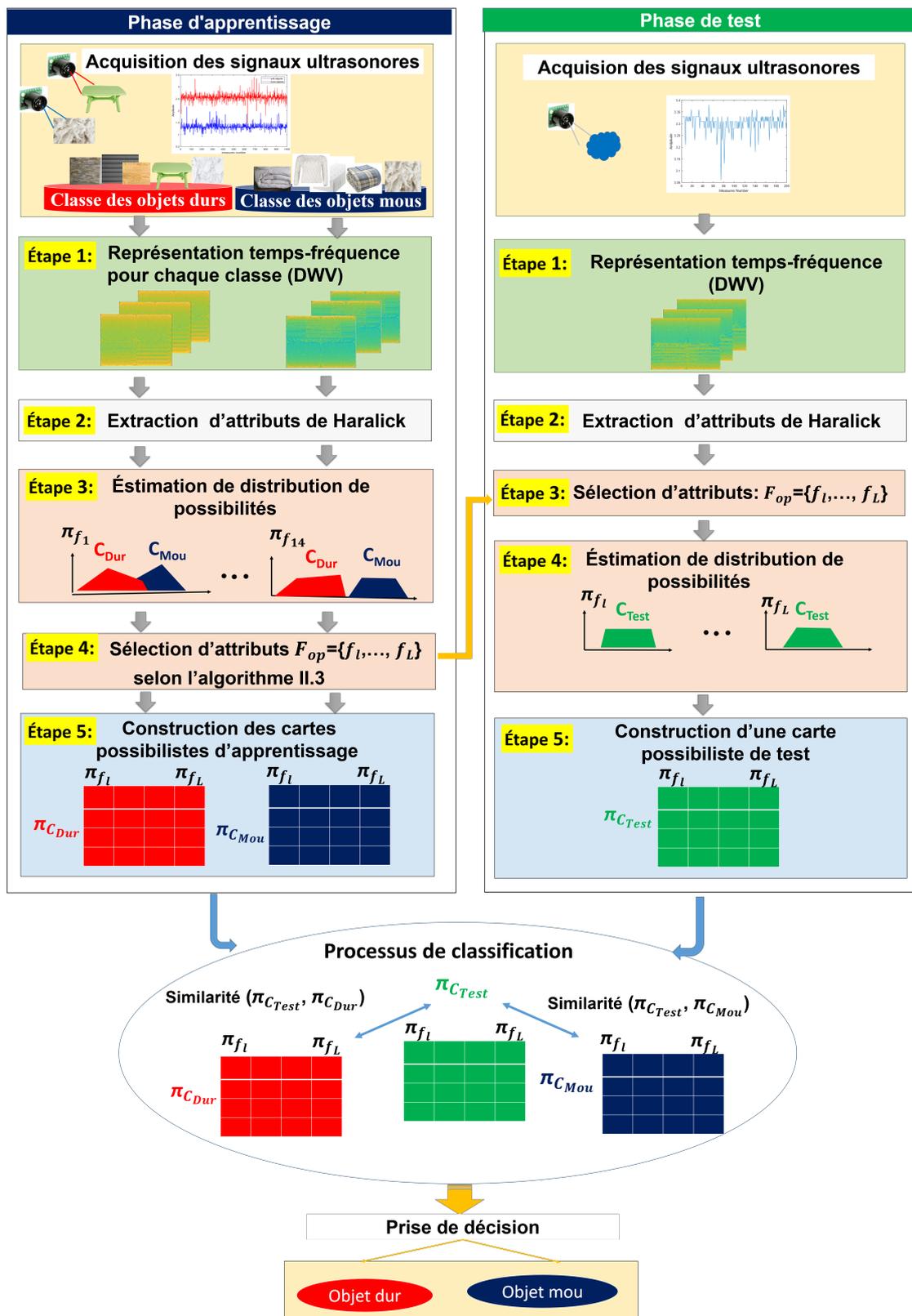


FIGURE III.16 – Approche possibiliste pour la reconnaissance de la rigidité de matériaux des objets détectés

— La base de données d'apprentissage comprend 300 échantillons utilisés pour la construction de la carte possibiliste " $\pi_{C_{Dur}}$ ", représentant la classe des objets durs, et de la carte possibiliste " $\pi_{C_{Mou}}$ " représentant la classe des objets mous.

- La base de données de test comprend 200 échantillons. À partir de chaque 20 échantillons, pris aléatoirement, une carte possibiliste de test " $\pi_{C_{Test}}$ " est construite. En total, pour chaque classe d'objet, nous avons construit une centaine de cartes possibilistes pour le test.

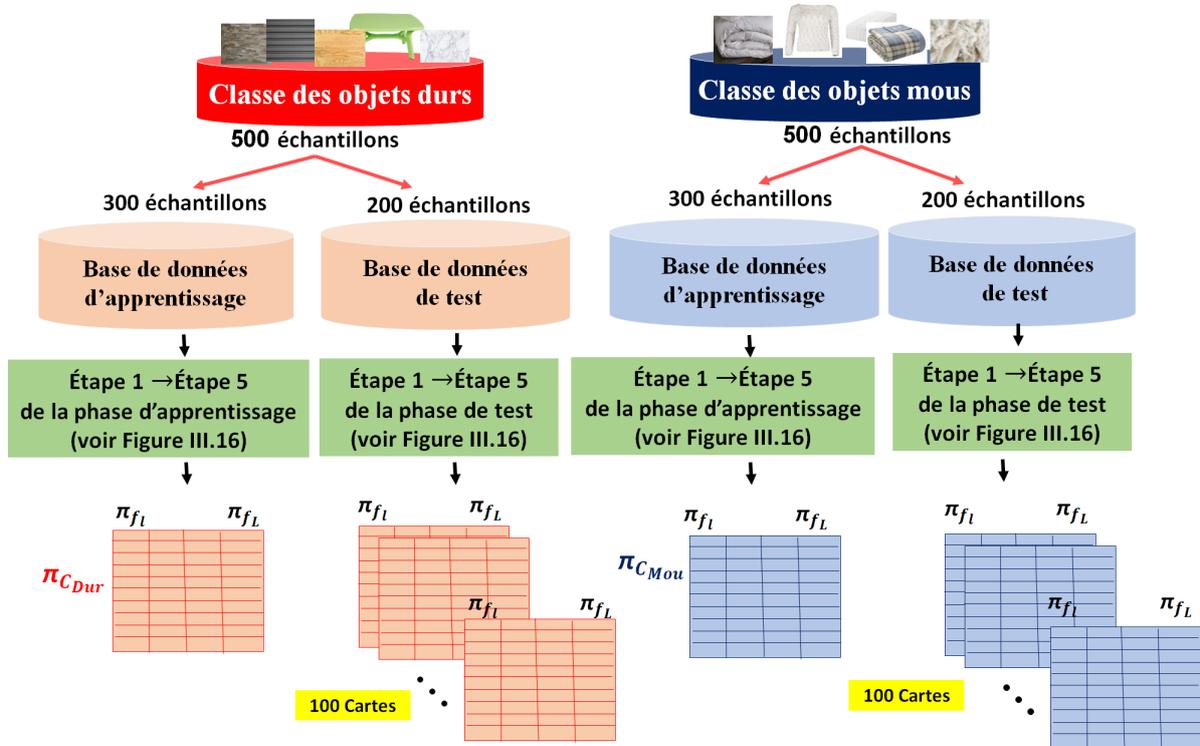


FIGURE III.17 – Structure de la base de données utilisée dans PM\_MRR

La construction des cartes " $\pi_{C_{Dur}}$ " et " $\pi_{C_{Mou}}$ " se fait sur la base des attributs optimaux sélectionnés par l'algorithme de sélection II.3 (App 3\_DP\_PMCC). Le Tableau III.3 présente les pouvoirs discriminants " $DP_{PMCC}^{f_k}$ " correspondants à chaque attribut  $f_k$ , avec  $k \in \{1, \dots, 14\}$ . Seuls les attributs dont le pouvoir discriminant est supérieur à 0.75 sont sélectionnés. Ainsi, l'ensemble d'attributs permettant la discrimination entre la classe des objets durs et la classe des objets mous est  $F_{op} = \{f_4, f_8, f_9, f_{11}\}$ . La Figure III.18 présente les distributions de possibilités correspondantes aux attributs sélectionnés.

Les résultats retrouvés par une sélection automatique d'attributs, assurée par l'algorithme proposé (Cf. Algorithme II.3), sont en parfait accord avec nos observations du fait que les deux attributs "Entropie" et "Variance", qui ont montré leur efficacité dans la discrimination des deux classes Dur/Mou par l'expérience, sont aussi apparues discriminants par l'algorithme II.3. L'attribut  $f_{11}$  qui s'écoule aussi de l'entropie s'avère aussi discriminant.

TABLE III.3 – Pouvoirs discriminants d'attributs correspondants à la discrimination de " $\pi_{C_{Dur}}$ " et " $\pi_{C_{Mou}}$ ".

Attribut	DP_PMCC	Attribut	DP_PMCC
$f_1$	0.72	$f_8$	<b><u>0.785</u></b>
$f_2$	0.66	$f_9$	<b><u>0.765</u></b>
$f_3$	0.635	$f_{10}$	0.65
$f_4$	<b><u>0.788</u></b>	$f_{11}$	<b><u>0.77</u></b>
$f_5$	0.533	$f_{12}$	0.631
$f_6$	0.664	$f_{13}$	0.635
$f_7$	0.675	$f_{14}$	0.627

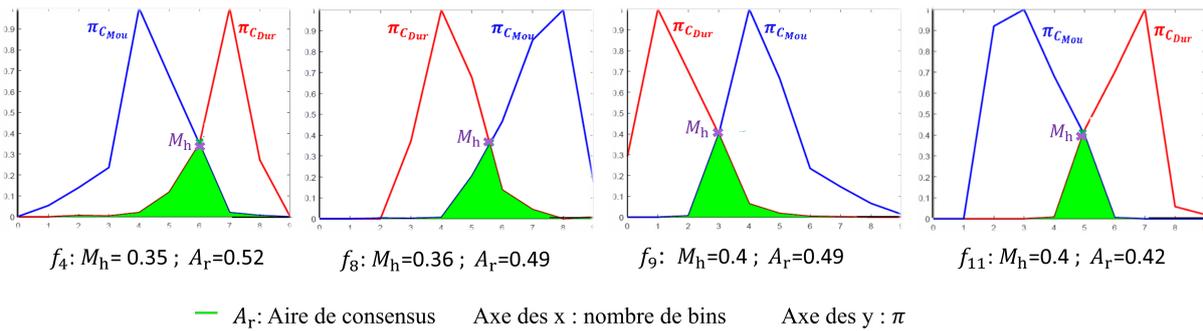


FIGURE III.18 – Distributions de possibilités correspondantes aux attributs sélectionnés

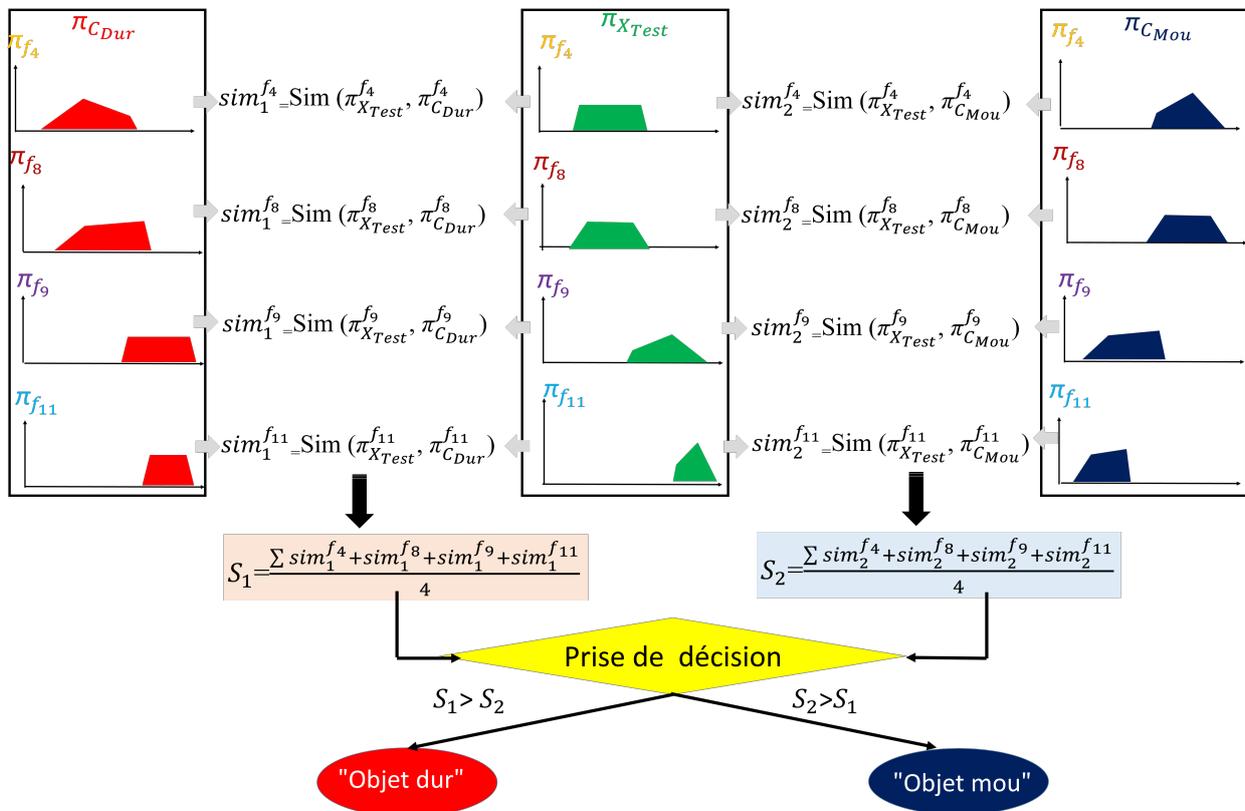


FIGURE III.19 – Processus de prise de décision basé sur la mesure de similarité

Une fois les deux cartes d'apprentissage et les cartes de test sont construites, nous passons au processus de classification. Comme le montre la Figure III.19, ce processus est basé sur

le degré de similarité entre la carte possibiliste de test et les cartes possibilistes construites en phase d'apprentissage. L'étiquette associée à la carte de test correspond à l'étiquette de la carte d'apprentissage ayant le plus haut degré de similarité.

À partir des résultats illustrés à la Figure III.20, nous pouvons tirer les conclusions suivantes :

- La modélisation des connaissances possibilistes offre un meilleur taux de précision que les classifieurs d'apprentissage automatique conventionnels que ce soit avant ou après la sélection d'attributs.
- Le taux de reconnaissance obtenu par un classifieur SVM a été réduit de 91% (en utilisant les 14 attributs de Haralick) à 89,5% en utilisant seulement trois attributs. La diminution du taux de reconnaissance tout en réduisant le nombre d'attributs est tout à fait compréhensible puisque l'incertitude n'a pas été tenu en compte.
- Pour l'approche possibiliste, bien que seuls quatre attributs soient utilisés, nous avons réussi à atteindre un taux de reconnaissance 93.5% supérieur au taux de reconnaissance obtenu à l'aide d'un classifieur classique SVM (91%) avec tous les attributs. Par conséquent, une modélisation possibiliste rend notre système plus fiable pour reconnaître la rigidité de matériaux des objets.

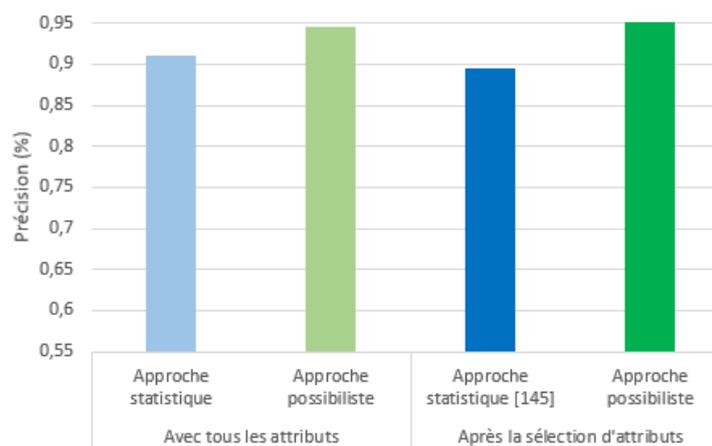


FIGURE III.20 – Taux de classification achevés avant et après sélection d'attributs en se basant sur une prise de décision vectorielle versus une prise de décision basée sur des modèles possibilistes

### III.3 Classification multi-classes : Reconnaissance de cinq matériaux

Dans cette section, nous nous concentrons sur la distinction des objets durs les uns des autres. Par conséquent, la reconnaissance de la rigidité matérielle est considérée comme un pro-

blème de classification multi-classes. Dans ce contexte, nous appliquons les classifieurs CBH(1) et CBH(2) proposés dans le Chapitre II. Comme le montre le Tableau III.4, l'univers fini de discours  $\Omega$  est composé de 5 classes  $\Omega = \{C_1, C_2, C_3, C_4, C_5\}$  et chaque classe  $C_i$  de  $\Omega$  est caractérisée par 14 attributs  $F = \{f_1, f_2, \dots, f_{14}\}$ .

TABLE III.4 – Contexte.

Notations	Descriptions
Source	Capteur ultrason
Nombre de classe : N	5
Nombre d'attributs : K	14
$\Omega = \{C_1, C_2, C_3, C_4, C_5\}$	$C_1$ : Marbre, $C_2$ : Plastique, $C_3$ : Béton, $C_4$ : Bois, $C_5$ : Métal
Nombre d'échantillons d'apprentissage	300(60x5)
Nombre d'échantillons de test	200(40x5)

### III.3.1 Utilisation de CBH(1) pour la reconnaissance de cinq matériaux : Simulation numérique

Le classifieur CBH(1) consiste à générer automatiquement, pour chaque itération  $t$ , la classe  $C_{op}^t$  la plus distincte de l'ensemble  $\Omega^t = \Omega^{t-1}/C_{op}^t$  selon un seul attribut optimal  $f_{op}^t$  qui favorise cette décomposition. Selon l'architecture proposée (Cf. Figure II.12), un problème de classification à 5 classes est converti en 4 problèmes de classification binaires.

La première décomposition binaire consiste à trouver simultanément l'attribut optimal  $f_{op}^1$  et la classe  $C_{op}^1$  la plus distincte de l'ensemble  $\Omega^1 = \Omega^0/C_{op}^1$ , avec  $\Omega^0 = \Omega$ , selon l'algorithme II.4. Le Tableau III.5 présente les pouvoirs discriminants " $DP\_SUMF_j^{f_k}$ ",  $j \in \{1, \dots, 5\}$ , correspondants à chaque couple possible  $X_j = (C_i, \Omega_i^1)$  avec  $\Omega_i^1 = \Omega^0/C_i$  selon chaque attribut  $f_k$ . La valeur maximale des pouvoirs discriminants correspond à " $DP\_SUMF_3^{f_{12}}$ ". Par conséquent,  $C_3$  est retenue comme la classe la plus distinguée de l'ensemble  $\Omega^1 = \Omega^0/C_3$  selon l'attribut optimal  $f_{12}$ . Les distributions de possibilités de l'ensemble  $\Omega^0$  selon  $f_{12}$  sont présentées à la Figure III.21.

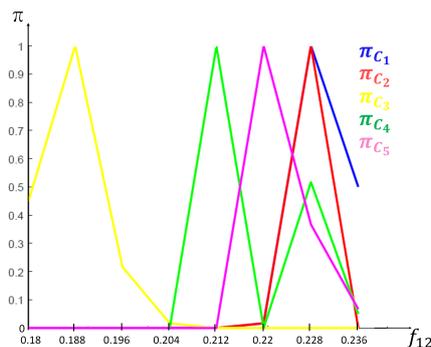

 FIGURE III.21 – Distributions de possibilités de l'ensemble  $\Omega^0$  selon  $f_{12}$

TABLE III.5 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 1<sup>ère</sup> itération du CBH(1).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>	<b>j=5</b>
$C_i$	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$\Omega_i^1 = \Omega^0 / C_i$	$\{C_2, C_3, C_4, C_5\}$	$\{C_1, C_3, C_4, C_5\}$	$\{C_1, C_2, C_4, C_5\}$	$\{C_1, C_2, C_3, C_5\}$	$\{C_1, C_2, C_3, C_4\}$
$f_1$	2.66	1.78	2.86	1.71	1.54
$f_2$	1.98	1.77	1.97	1.51	1.87
$f_3$	1.42	2.05	2.16	1.36	1.12
$f_4$	2.08	2.03	2.08	2.33	1.93
$f_5$	1.7	1.63	1.74	0.99	1.57
$f_6$	2.1	1.28	2.11	2.18	1.08
$f_7$	2.27	1.66	2.25	1.29	1.48
$f_8$	2.83	1.54	2.79	1.38	2.85
$f_9$	1.5	1.67	2.4	2.59	2.48
$f_{10}$	1.98	1.75	1.97	1.46	1.81
$f_{11}$	2.24	2.19	2.34	2.51	2.61
$f_{12}$	1.56	1.59	<b>2.94</b>	2.49	2.18
$f_{13}$	2.33	1.68	2.07	1.21	1.99
$f_{14}$	1.33	1.72	2.06	1.67	0.84

La deuxième décomposition binaire vise à extraire, simultanément, la classe la plus distinguée  $C_{op}^2$  de l'ensemble  $\Omega^2 = \Omega^1 / C_{op}^2$  selon l'attribut optimal  $f_{op}^2$ , en suivant le processus décrit dans l'algorithme II.4. Le Tableau III.6 présente les pouvoirs discriminants des "DP\_SUMF $_j^{f_k}$ ",  $j \in \{1, \dots, 4\}$ , correspondants à chaque couple possible  $X_j = (C_i, \Omega_i^2)$  avec  $\Omega_i^2 = \Omega^1 / C_i$  selon chaque attribut  $f_k$ . Comme la valeur maximale des pouvoirs discriminants correspond à "DP\_SUMF $_4^{f_{11}}$ ", alors la classe  $C_5$  est retenue comme la classe la plus distinguée de l'ensemble  $\Omega^2 = \Omega^1 / C_5$  selon l'attribut optimal  $f_{11}$ . Les distributions de possibilités de l'ensemble  $\Omega^1$  selon  $f_{11}$  sont présentées à la Figure III.22.

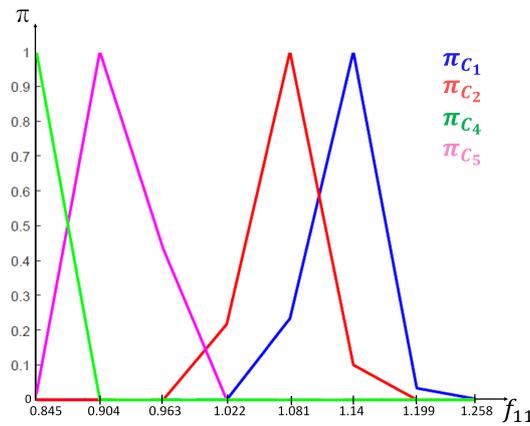
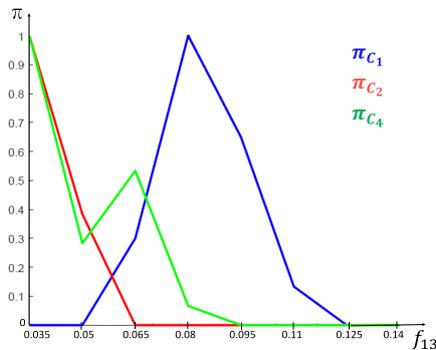

 FIGURE III.22 – Distributions de possibilités de l'ensemble  $\Omega^1$  selon  $f_{11}$

TABLE III.6 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 2<sup>ème</sup> itération du CBH(1).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>
$C_i$	$C_1$	$C_2$	$C_4$	$C_5$
$\Omega_i^2 = \Omega^1 / C_i$	$\{C_2, C_4, C_5\}$	$\{C_1, C_4, C_5\}$	$\{C_1, C_2, C_5\}$	$\{C_1, C_2, C_4\}$
$f_1$	2.5	1.85	1.71	1.5
$f_2$	2.72	1.79	1.37	1.93
$f_3$	1.28	2.09	1.24	0.82
$f_4$	2.87	1.98	2.17	1.98
$f_5$	2.55	1.52	0.75	1.62
$f_6$	2.88	0.86	2.1	0.71
$f_7$	2.75	1.53	0.86	1.3
$f_8$	2.86	1.55	1.14	2.83
$f_9$	1.51	1.32	2.26	2.29
$f_{10}$	2.84	1.84	1.3	1.87
$f_{11}$	2.15	2.1	2.44	<b><u>2.88</u></b>
$f_{12}$	1.49	1.57	2.22	2.1
$f_{13}$	2.49	1.67	0.79	1.96
$f_{14}$	0.79	1.74	1.65	0.62

Pour la troisième itération, suivant le même processus, on obtient que  $C_1$  est la classe la plus distincte de l'ensemble  $\Omega^3 = \Omega^2 / C_1$  selon l'attribut  $f_{13}$ . Les distributions de possibilités de l'ensemble  $\Omega^2$  selon  $f_{13}$  sont illustrées à la Figure III.23.


 FIGURE III.23 – Distributions de possibilités de l'ensemble  $\Omega^2$  selon  $f_{13}$ 

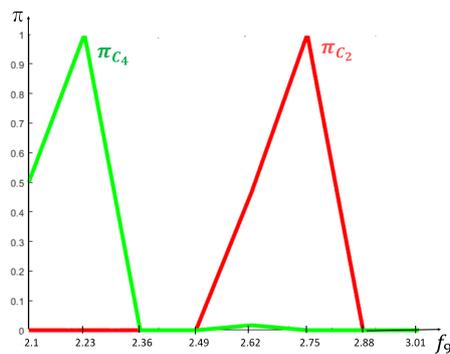
Enfin, les classes  $C_2$  et  $C_4$  restent d'être discriminées selon un attribut optimal. En appliquant l'algorithme II.4, on constate que ces deux classes sont bien distingués selon l'attribut  $f_9$ . La Figure III.24 présente les deux distributions de possibilités de  $\pi_{C_2}$  et  $\pi_{C_4}$  selon  $f_9$ . À la lumière de ces 4 classifications binaires, le CBH(1) conçu est représenté à la Figure III.25.

TABLE III.7 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 3<sup>ème</sup> itération du CBH(1).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>
$C_i$	$C_1$	$C_2$	$C_4$
$\Omega_i^3 = \Omega^2 / C_i$	$\{C_2, C_4\}$	$\{C_1, C_4\}$	$\{C_1, C_2\}$
$f_1$	1.22	1.26	2.10
$f_2$	2.74	2.34	1.20
$f_3$	2.32	1.86	0.80
$f_4$	2.82	2.47	2.16
$f_5$	2.68	2.44	0.97
$f_6$	2.82	2.51	2.31
$f_7$	2.60	2.15	1.46
$f_8$	2.10	2.36	1.34
$f_9$	2.25	1.42	2.25
$f_{10}$	2.70	2.22	1.2
$f_{11}$	2.61	2.30	3.1
$f_{12}$	2.81	1.78	0.85
$f_{13}$	<b>2.96</b>	2.43	1.61
$f_{14}$	2.67	1.95	0.88

 TABLE III.8 – Pouvoirs discriminants d'attributs correspondants à la discrimination de  $C_2$  et  $C_4$ .

Attribut	$DP\_SUMF^{f_k}$	Attribut	$DP\_SUMF^{f_k}$
$f_1$	1.40	$f_8$	0.74
$f_2$	0.53	$f_9$	<b>2.95</b>
$f_3$	0.80	$f_{10}$	0.63
$f_4$	1.41	$f_{11}$	2.74
$f_5$	0.45	$f_{12}$	0.03
$f_6$	1.98	$f_{13}$	0.77
$f_7$	1.20	$f_{14}$	0.093


 FIGURE III.24 – Distributions de possibilités correspondantes aux classes  $C_2$  et  $C_4$  selon  $f_9$

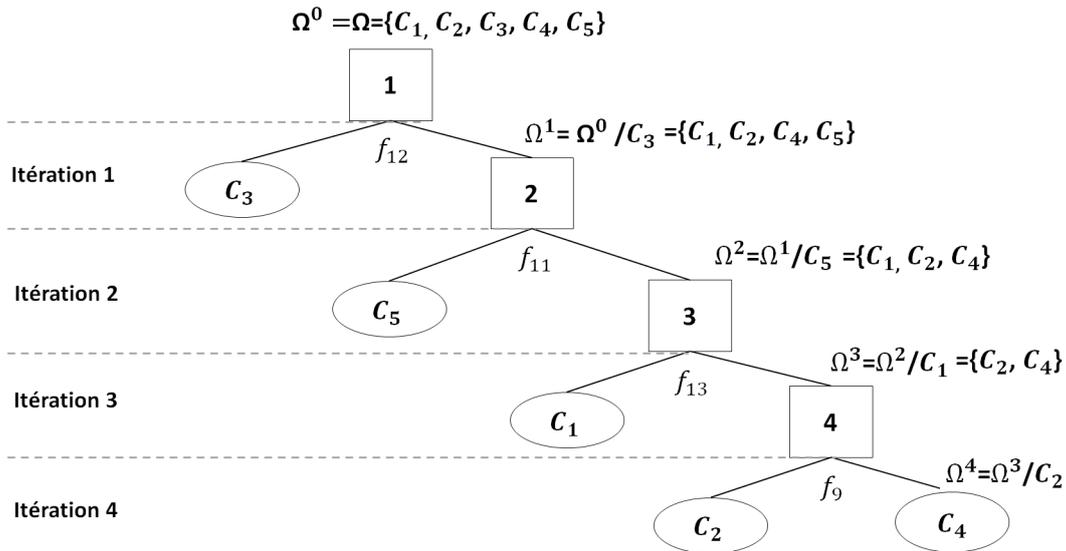


FIGURE III.25 – CBH(1) résultant

### III.3.2 Utilisation de CBH(2) pour la reconnaissance de cinq matériaux : Simulation numérique

Le classifieur CBH(2) consiste à générer automatiquement pour chaque itération  $t$ , la classe  $C_{op}^t$  la plus distincte de l'ensemble  $\Omega^t = \Omega^{t-1} / C_{op}^t$  selon un ensemble d'attributs les plus pertinents  $F_{op}^t$  qui favorisent cette décomposition. Selon l'architecture de CBH(2) (Cf. Figure II.15), un problème de classification à 5 classes est converti en 4 problèmes de classification binaires. Il est à noter qu'un seuil égal à 0.82 sur les pouvoirs discriminants a été retenu pour toutes les itérations de construction de CBH(2).

La première décomposition binaire consiste à trouver la classe  $C_{op}^1$  la plus distincte de l'ensemble  $\Omega^1 = \Omega^0 / C_{op}^1$ , avec  $\Omega^0 = \Omega$ , selon l'ensemble d'attributs  $F_{op}^1$  déterminé par l'algorithme II.3. Le Tableau III.9 présente les valeurs de " $DP\_PMCC_j^{f_k}$ ",  $j \in \{1, \dots, 5\}$ , correspondants à chaque couple possible  $X_j = (C_i, \Omega_i^1)$  avec  $\Omega_i^1 = \Omega^0 / C_i$  selon chaque attribut  $f_k$ . Le Tableau III.10 présente le score obtenu pour chaque décomposition  $X_j = (C_i, \Omega_i^1)$  selon l'ensemble d'attributs  $F_{op}^1$ . De ces deux tableaux, nous concluons que selon l'ensemble d'attributs  $F_{op}^1 = \{f_1, f_{12}\}$ , la classe  $C_3$  est la classe la plus distincte de l'ensemble  $\Omega^1 = \Omega^0 / C_3$ . La Figure III.26 illustre les distributions de possibilités correspondantes à  $C_{op}^1$  et  $\Omega^1$  selon tous les attributs de Haralick.

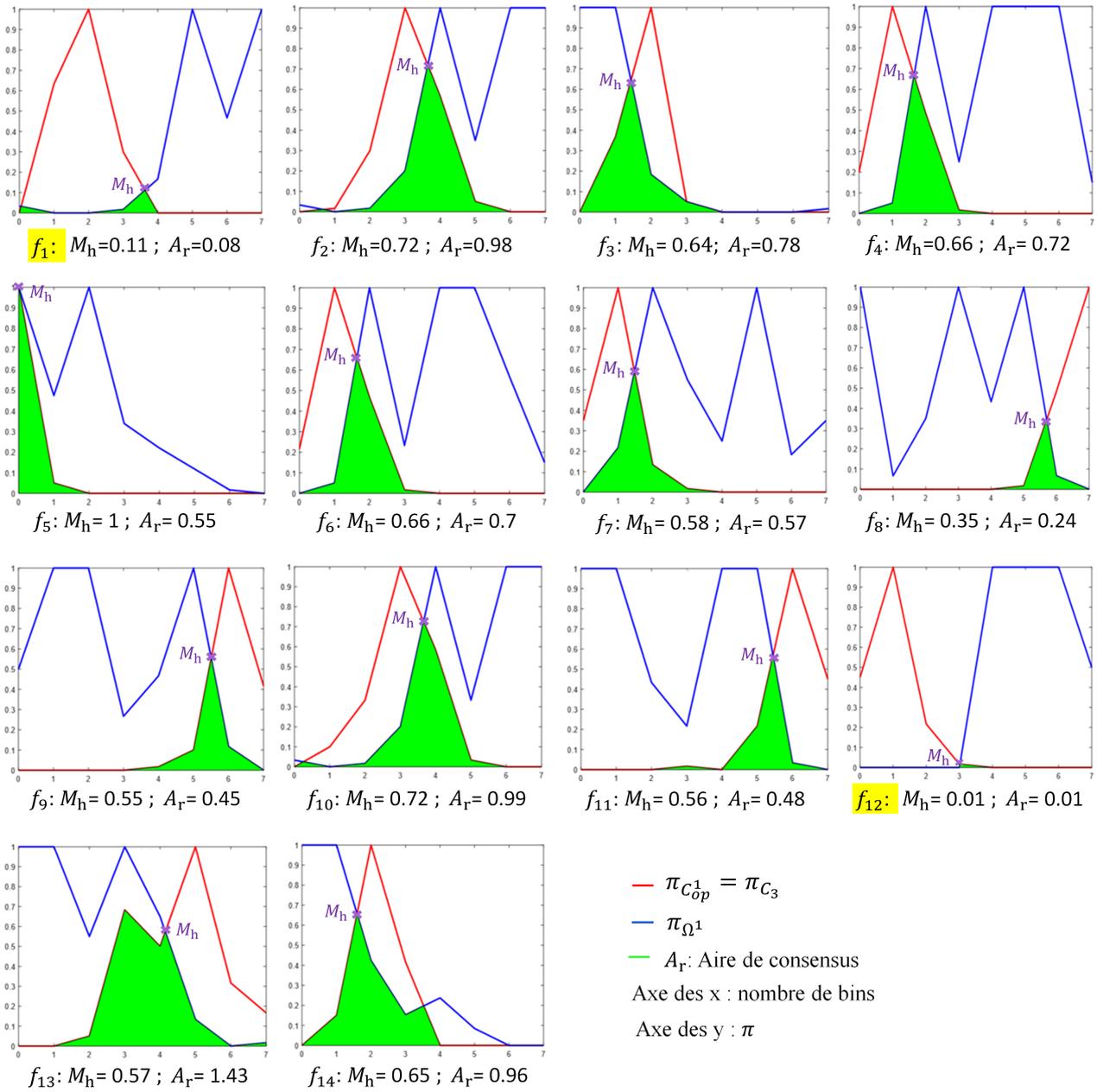
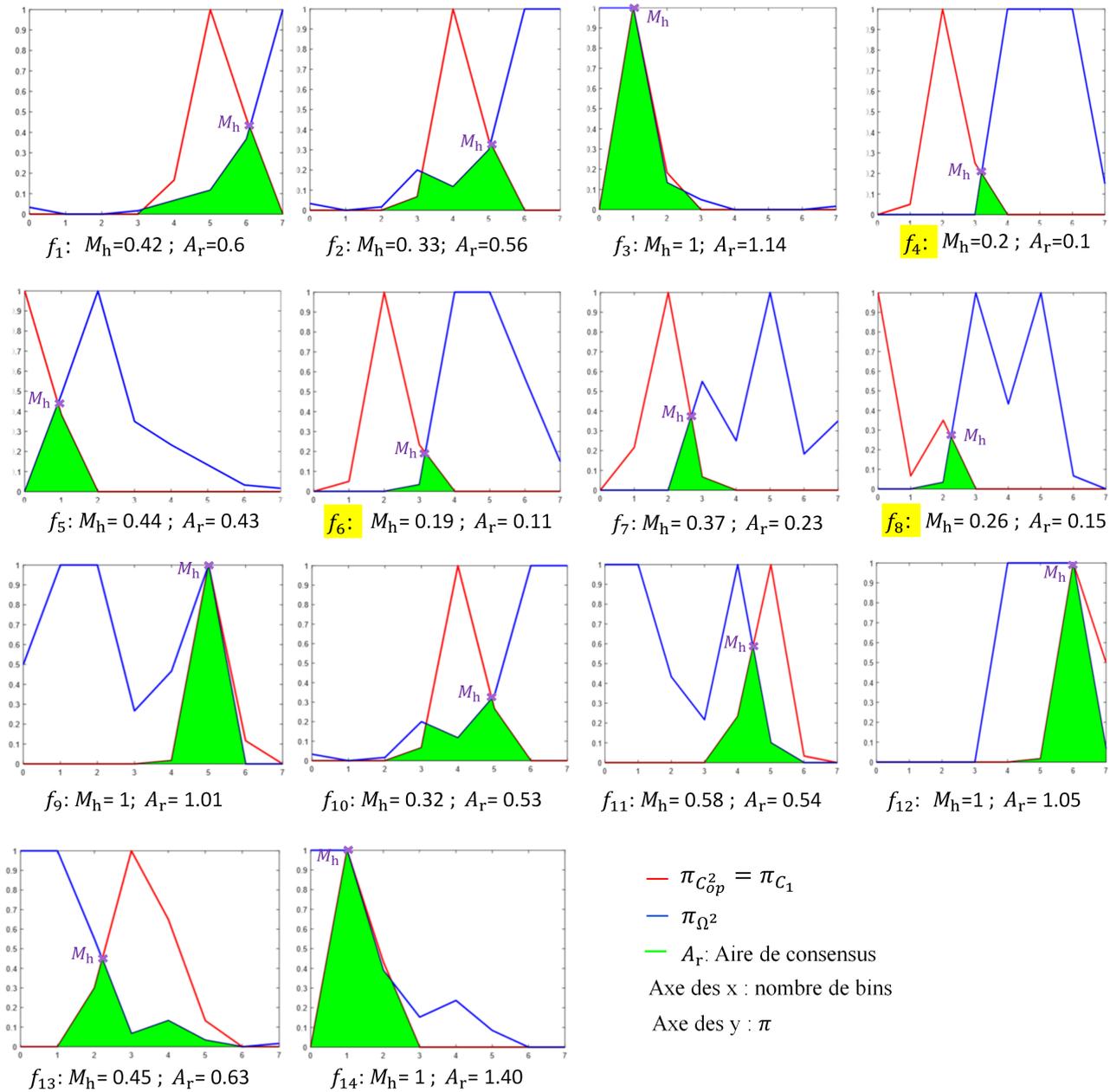


FIGURE III.26 – Distributions de possibilités correspondantes à  $C_{op}^1$  et  $\Omega^1$

La deuxième décomposition binaire vise à extraire, la classe  $C_{op}^2$  la plus distincte de l'ensemble  $\Omega^2 = \Omega^1 / C_{op}^2$  selon l'ensemble d'attributs  $F_{op}^2$ . Le Tableau III.11 présente les pouvoirs discriminants "DP\_PMCC $_j^{f_k}$ ",  $j \in \{1, \dots, 4\}$ , correspondants à chaque couple possible  $X_j = (C_i, \Omega_i^2)$ , avec  $\Omega_i^2 = \Omega^1 / C_i$  selon chaque attribut  $f_k$ . Le Tableau III.12 présente le score obtenu pour chaque décomposition  $X_j = (C_i, \Omega_i^2)$  selon l'ensemble d'attributs  $F_{op_i}^2$ . Sur la base de ces deux tableaux, la classe  $C_1$  est retenue comme la classe la plus distincte de l'ensemble  $\Omega_i^2 = \Omega^1 / C_1$  selon l'ensemble d'attributs  $F_{op}^2 = \{f_4, f_6, f_8\}$ . La Figure III.27 illustre les distributions de possibilités correspondantes à  $C_{op}^2$  et  $\Omega^2$ .


 FIGURE III.27 – Distributions de possibilités correspondantes à  $C_{op}^2$  et  $\Omega^2$ 

La troisième décomposition binaire vise à extraire, la classe  $C_{op}^3$  la plus distincte de l'ensemble  $\Omega^3 = \Omega^2 / C_{op}^3$  selon l'ensemble d'attributs  $F_{op}^3$ . Le Tableau III.13 présente les pouvoirs discriminants " $DP\_PMCC_j^{f_k}$ ",  $j \in \{1,2,3\}$ , correspondants à chaque couple possible  $X_j = (C_i, \Omega_i^3)$  avec  $\Omega_i^3 = \Omega^2 / C_i$  selon chaque attribut  $f_k$ . Le Tableau III.14 présente le score obtenu pour chaque décomposition. De ces deux tableaux, la classe  $C_2$  est retenue comme la classe la plus distincte de l'ensemble  $\Omega^3 = \Omega^2 / C_2$  selon l'ensemble d'attributs  $F_{op}^3 = \{f_9, f_{11}\}$ . La Figure III.28 illustre les distributions de possibilités correspondantes à  $C_{op}^3$  et  $\Omega^3$ .

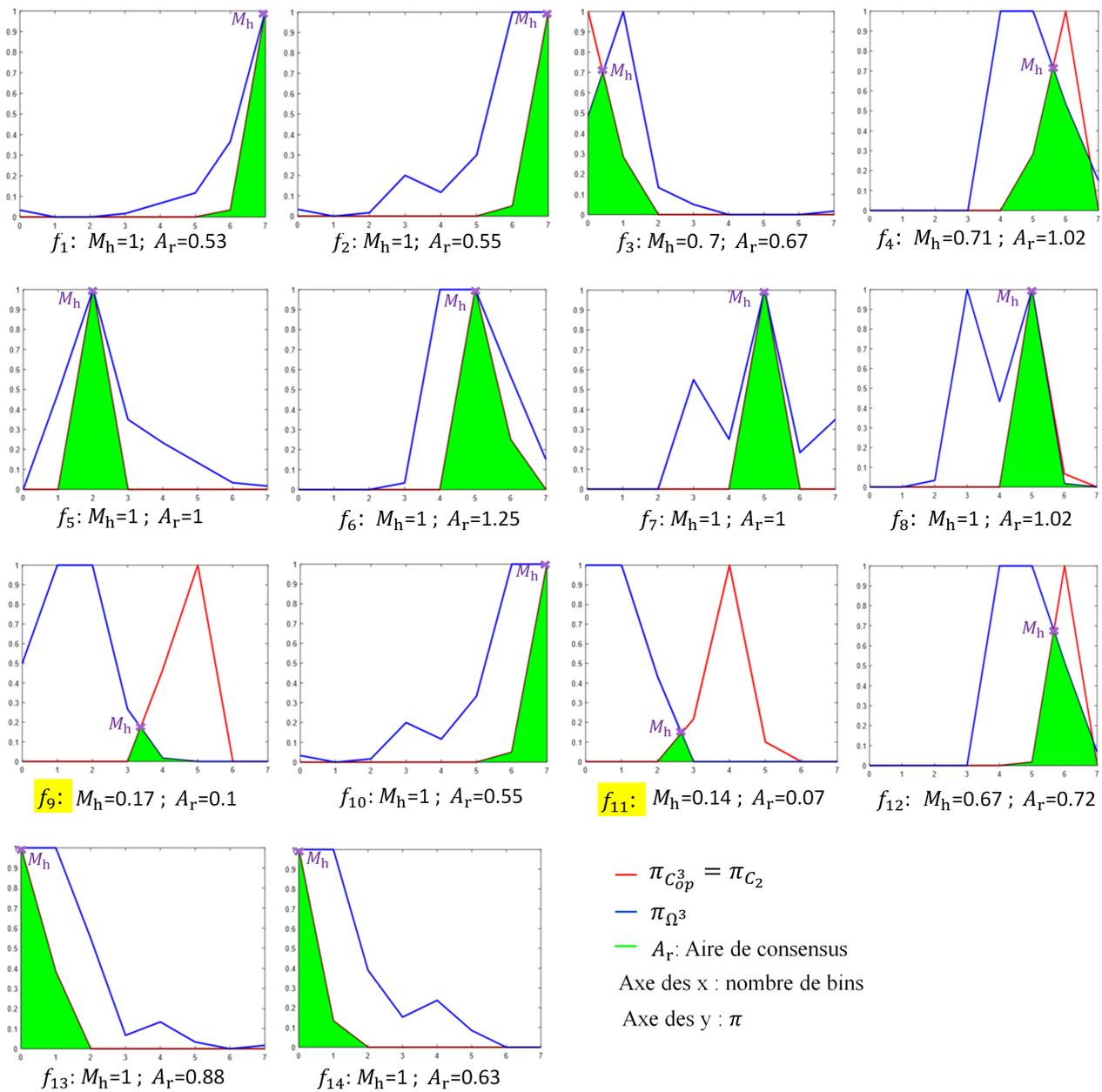


FIGURE III.28 – Distributions de possibilités correspondantes à  $C_{op}^3$  et  $\Omega^3$

Enfin, seules les classes  $C_4$  et  $C_5$  restent d'être discriminées. En effectuant le même raisonnement que ci-dessus, on constate que ces deux classes sont bien distinguées selon l'attribut  $f_8$ . La Figure III.29 présente les distributions de possibilités correspondantes à  $C_4$  et  $C_5$  selon chaque attribut de Haralick. À la lumière de ces 4 classifications binaires, le CBH(2) conçu est illustré à la Figure III.30.

TABLE III.9 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 1<sup>ère</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>	<b>j=5</b>
$C_i$	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$\Omega_i^1 = \Omega^0 / C_i$	$\{C_2, C_3, C_4, C_5\}$	$\{C_1, C_3, C_4, C_5\}$	$\{C_1, C_2, C_4, C_5\}$	$\{C_1, C_2, C_3, C_5\}$	$\{C_1, C_2, C_3, C_4\}$
$f_1$	0.755	0.462	<b>0.94</b>	0.438	0.423
$f_2$	0.571	0.46	0.57	0.423	0.515
$f_3$	0.415	0.602	0.624	0.411	0.399
$f_4$	0.613	0.6	0.618	0.683	0.565
$f_5$	0.437	0.43	0.46	0.398	0.428
$f_6$	0.62	0.41	0.62	0.658	0.398
$f_7$	0.673	0.43	0.67	0.410	0.42
$f_8$	<b>0.86</b>	0.424	0.808	0.414	<b>0.867</b>
$f_9$	0.42	0.43	0.693	0.727	0.716
$f_{10}$	0.573	0.46	0.57	0.419	0.51
$f_{11}$	0.668	0.667	0.685	0.726	0.732
$f_{12}$	0.425	0.428	<b>0.994</b>	0.703	0.635
$f_{13}$	0.62	0.437	0.613	0.402	0.58
$f_{14}$	0.4	0.455	0.606	0.431	0.359

TABLE III.10 – Score du chaque couple possible pour la 1<sup>ère</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>	<b>j=5</b>
$C_i$	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$\Omega_i^1 = \Omega^0 / C_i$	$\{C_2, C_3, C_4, C_5\}$	$\{C_1, C_3, C_4, C_5\}$	$\{C_1, C_2, C_4, C_5\}$	$\{C_1, C_2, C_3, C_5\}$	$\{C_1, C_2, C_3, C_4\}$
$F_{op_i}^1$	$\{f_8\}$	$\{\emptyset\}$	$\{f_1, f_{12}\}$	$\{\emptyset\}$	$\{f_8\}$
Score	0.86	$\emptyset$	<b>0.967</b>	$\emptyset$	0.867

TABLE III.11 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 2<sup>ème</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>
$C_i$	$C_1$	$C_2$	$C_4$	$C_5$
$\Omega_i^2 = \Omega^1 / C_i$	$\{C_2, C_4, C_5\}$	$\{C_1, C_4, C_5\}$	$\{C_1, C_2, C_5\}$	$\{C_1, C_2, C_4\}$
$f_1$	0.747	0.462	0.438	0.425
$f_2$	0.795	0.46	0.423	0.523
$f_3$	0.418	0.602	0.411	0.402
$f_4$	<b>0.893</b>	0.573	0.673	0.565
$f_5$	0.749	0.428	0.398	0.428
$f_6$	<b>0.897</b>	0.41	0.658	0.397
$f_7$	0.798	0.428	0.404	0.42
$f_8$	<b>0.86</b>	0.428	0.41	<b>0.854</b>
$f_9$	0.428	0.421	0.717	0.717
$f_{10}$	0.802	0.46	0.42	0.518
$f_{11}$	0.671	0.666	0.727	<b>0.883</b>
$f_{12}$	0.425	0.428	0.693	0.634
$f_{13}$	0.730	0.437	0.401	0.554
$f_{14}$	0.4	0.455	0.431	0.375

TABLE III.12 – Score du chaque couple possible pour la 2<sup>ème</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>	<b>j=4</b>
$C_i$	$C_1$	$C_2$	$C_4$	$C_5$
$\Omega_i^2 = \Omega^1 / C_i$	$\{C_2, C_4, C_5\}$	$\{C_1, C_4, C_5\}$	$\{C_1, C_2, C_5\}$	$\{C_1, C_2, C_4\}$
$F_{op_i}^2$	$\{f_4, f_6, f_8\}$	$\{\emptyset\}$	$\{\emptyset\}$	$\{f_8, f_{11}\}$
Score	<b>0.883</b>	$\emptyset$	$\emptyset$	0.876

 TABLE III.13 – Pouvoirs discriminants d'attributs correspondants aux différents couples possibles pour la 3<sup>ème</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>
$C_i$	$C_2$	$C_4$	$C_5$
$\Omega_i^3 = \Omega^2 / C_i$	$\{C_4, C_5\}$	$\{C_2, C_5\}$	$\{C_2, C_4\}$
$f_1$	0.462	0.438	0.438
$f_2$	0.46	0.423	0.555
$f_3$	0.602	0.411	0.413
$f_4$	0.572	0.685	0.565
$f_5$	0.428	0.428	0.428
$f_6$	0.41	0.666	0.398
$f_7$	0.428	0.42	0.42
$f_8$	0.427	0.41	<b>0.873</b>
$f_9$	<b>0.908</b>	0.722	0.712
$f_{10}$	0.46	0.42	0.546
$f_{11}$	<b>0.925</b>	0.727	<b>0.882</b>
$f_{12}$	0.613	0.693	0.634
$f_{13}$	0.437	0.401	0.566
$f_{14}$	0.455	0.431	0.54

 TABLE III.14 – Score du chaque couple possible pour la 3<sup>ème</sup> itération du CBH(2).

	<b>j=1</b>	<b>j=2</b>	<b>j=3</b>
$C_i$	$C_2$	$C_4$	$C_5$
$\Omega_i^3 = \Omega^2 / C_i$	$\{C_4, C_5\}$	$\{C_2, C_5\}$	$\{C_2, C_4\}$
$F_{op_i}^3$	$\{f_9, f_{11}\}$	$\{\emptyset\}$	$\{f_8, f_{11}\}$
Score	<b>0.916</b>	$\emptyset$	0.877

 TABLE III.15 – Pouvoirs discriminants d'attributs correspondants à la discrimination de  $C_4$  et  $C_5$ .

Attribut	DP_PMCC	Attribut	DP_PMCC
$f_1$	0.438	$f_8$	<b>0.823</b>
$f_2$	0.555	$f_9$	0.723
$f_3$	0.413	$f_{10}$	0.546
$f_4$	0.685	$f_{11}$	0.732
$f_5$	0.428	$f_{12}$	0.7
$f_6$	0.667	$f_{13}$	0.588
$f_7$	0.42	$f_{14}$	0.54

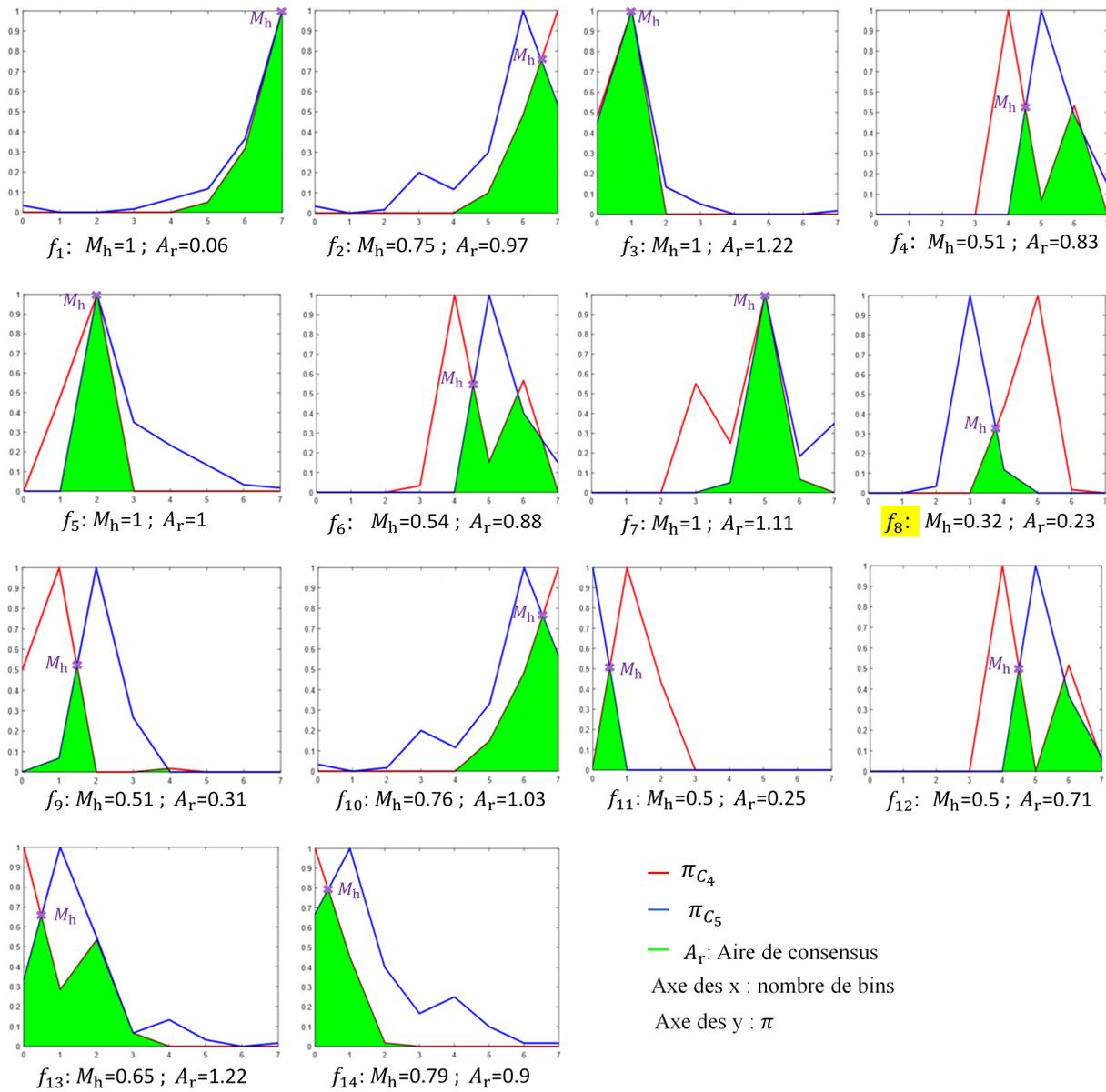


FIGURE III.29 – Distributions de possibilités correspondantes à  $C_4$  et  $C_5$

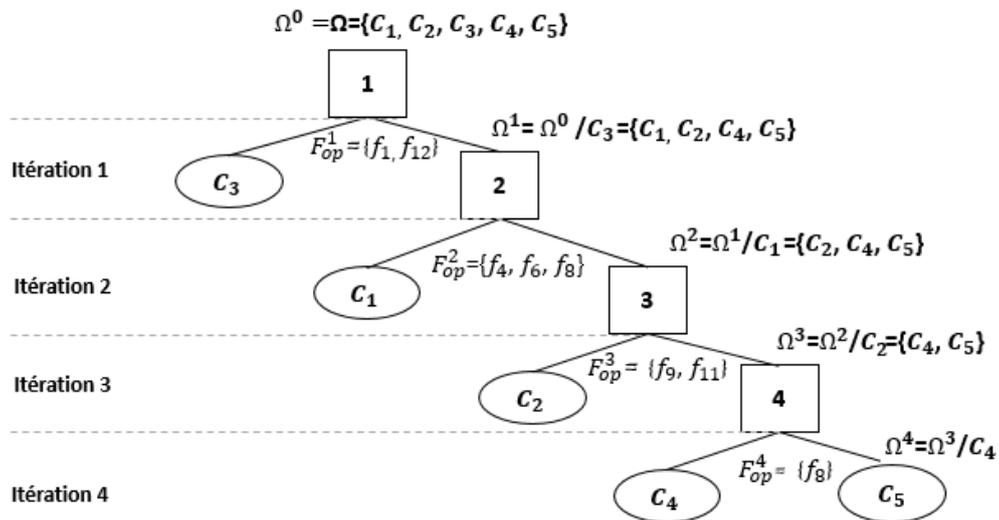


FIGURE III.30 – CBH(2) résultant

### III.3.3 Évaluation et discussions

Pour évaluer à nouveau les classifieurs CBH(1) et CBH(2), nous les appliquons à des données ultrasonores réelles en suivant le même raisonnement comparatif utilisé au Chapitre II. En fait, le CBH(1) et CBH(2) (selon les 6 mesures statistiques définies dans la Section II.3.4) sont comparés à d'autres classifieurs de nature différente. Du Tableau III.16 et la Figure III.31, on peut tirer les conclusions suivantes :

- Le CBH(2) performe tous les autres classifieurs.
- Le CBH(1) est le deuxième classifieur le plus performant, après CBH(2).
- Les classifieurs CBH(2) et CBH(1) performent l'arbre de décision conventionnel C4.5 ce qui prouve en outre que la méthode de décomposition OVA qu'on a adoptée est meilleure que celle d'un arbre de décision.
- Les classifieurs CBH(2) et CBH(1) performent le classifieur probabiliste NB. Ainsi, la modélisation possibiliste est plus adéquate qu'une modélisation probabiliste pour gérer des données incertaines de nature épistémiques.
- Les classifieurs CBH(1) et CBH(2) donnent un compromis satisfaisant entre le nombre d'attributs/ précision.

TABLE III.16 – Évaluation de CBH(1) et CBH(2) sur une base de données des signaux ultrasonores.

	J48	SVM	MLP	NB	Classifieur possibiliste conventionnel	Évaluation de CBH(1) par similarité	CBH(2)
VP	0,77	0,88	0,87	0,885	0,895	0,928	0,936
FP	0,058	0,03	0,033	0,028	0,027	0,018	0,16
Précision	0,829	0,908	0,902	0,91	0,93	0,947	0,951
Rappel	0,77	0,88	0,87	0,885	0,895	0,928	0,936
Mesure F	0,761	0,874	0,861	0,878	0,892	0,925	0,934
MCC	0,731	0,859	0,848	0,862	0,871	0,915	0,925

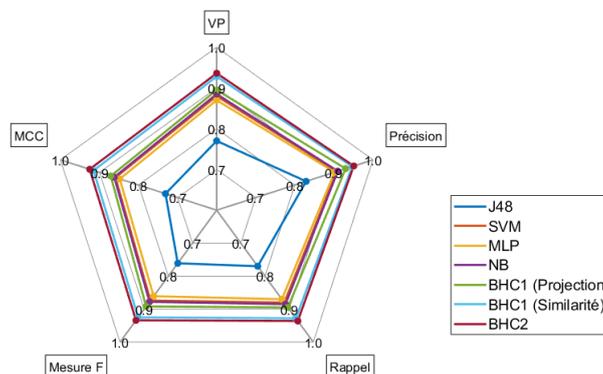


FIGURE III.31 – Performances des classifieurs sur des données ultrasonores

## III.4 Conclusion

Notre intérêt étant porté sur l'interprétation de scène observée par un capteur ultrason, on a consacré ce chapitre au traitement, à l'analyse, et à l'interprétation de l'écho ultrasonore afin d'en extraire des informations pertinentes sur l'objet qui l'a réfléchi. Tout d'abord, nous nous intéressons à la construction d'un signal de profil à partir d'une scène balayée horizontalement. Ce signal est obtenu sur la base du calcul de distance déterminée en fonction du temps d'arrivée de l'onde ultrasonore. Afin d'attribuer une étiquette de profondeur à chaque objet constituant la scène, une étape de segmentation du signal de profil est nécessaire.

Ensuite, nous essayons de tester la capacité de l'onde ultrasonore à identifier la rigidité de matériaux des objets détectés. Dans ce contexte, deux approches sont proposées : l'une est une méthode statistique, où la prise de décision est basée sur des vecteurs d'attributs, et l'autre est une méthode possibiliste, où la prise de décision est basée sur des modèles possibilistes. Il est à noter que l'approche possibiliste est plus fiable que l'approche statistique car elle prend en compte l'incertitude des données. Les deux approches proposées reposent sur la représentation temps-fréquence de l'écho ultrasonore. En fait, une telle représentation permet aux échos ultrasonores (signaux unidimensionnels), la possibilité d'être analysés dans un domaine bidimensionnel. La bi-dimensionnalité de la représentation temps-fréquence offre deux degrés de liberté, ce qui offre une grande flexibilité dans l'analyse de l'écho ultrasonore. Après avoir présenté l'écho ultrasonore sous forme de représentation temps-fréquence, nous extrayons l'ensemble des attributs de Haralick correspondant à chaque segment du signal de profil, le traitons et l'analysons afin d'attribuer une étiquette "Objet dur" ou "Objet mou".

Dans la dernière partie de ce chapitre, nous nous concentrons sur une inter-classification des objets durs en appliquant les classifieurs CBH(1) et CBH(2) proposés dans le chapitre précédent.

Dans le chapitre suivant, nous nous intéressons à interpréter une scène observée par un capteur de vision.

# Chapitre IV

## Interprétation de scène basée sur une caméra monoculaire RVB

### Sommaire

---

IV.1 État de l'art . . . . .	102
IV.2 La saillance visuelle . . . . .	103
IV.3 Approche d'interprétation de scène par une caméra monoculaire RVB . . . . .	106
IV.4 Conclusion . . . . .	115

---

## Introduction

La vision est sans aucun doute le sens humain le plus développé permettant aux êtres humains une excellente capacité à analyser et interpréter des scènes complexes en temps réel.

Ces dernières années, la vision par ordinateur est en plein essor et est utilisée dans différentes applications telles que de surveillance, de médecine, de robotique, etc. Dans telles applications, des données décrivant la scène sont acquises. Afin d'analyser la scène en utilisant ces données, il est très important de mettre en place un système d'interprétation de scène efficace. L'interprétation de scène est très liée à la perception visuelle du monde réel par des capteurs d'imagerie. En effet, la scène du monde réel (3D) est d'abord acquise par un ou plusieurs capteurs de vision puis projetée dans un nouvel espace de représentation des connaissances (images numérique (2D)). Par conséquent, l'interprétation de la scène est souvent rendue comme une interprétation des images représentatives de la scène. Le terme "interprétation de scène" fait référence au mécanisme de raisonnement et d'analyse qui vise à extraire la sémantique incluse dans les images. Considérant que l'image est le support de mesure physique approximative de la scène, cinq niveaux d'interprétation sont déterminés [6] : niveau de sous-pixel, niveau de pixel, niveau de région, niveau d'objet et niveau de scène.

Il est à noter que le résultat du système d'interprétation de scène dépend du but à atteindre, des connaissances à priori liées à la scène, du modèle de connaissance, du mécanisme d'interprétation et voire du capteur de vision utilisé.

De nos jours, bien que les ordinateurs ont une puissance de calcul plus élevée que les humains, ils n'arrivent pas à produire de robustes systèmes de vision. Ceci est principalement dû à deux fossés : fossé sensoriel et fossé sémantique (Cf. Figure IV.1) :

- Le fossé sensoriel intervient lors de l'acquisition de la scène. Il est défini comme *"le fossé entre les objets dans le monde réel et l'information contenue dans une description (informatique) dérivée de l'enregistrement de cette scène"* [147].
- Le fossé sémantique est très délicat à gérer. Il est défini comme *"le manque de concordance entre les informations qu'on peut extraire des données visuelles et l'interprétation qu'on ces mêmes données pour un utilisateur dans une situation déterminée"* [147]. En effet, les connaissances *a priori* sont indispensables pour extraire la sémantique incluse dans l'image.

L'être humain parvient à remédier ces deux lacunes parce qu'il possède une quantité massive de connaissances *a priori*, une capacité d'apprentissage, une capacité d'analogie et d'approximation, etc. Au contraire, il est encore difficile pour une machine sans de tels avantages de résoudre ce problème. Cela explique pourquoi les humains et les machines ont des interprétations différentes de la même scène.

FIGURE IV.1 – Le fossé sensoriel et le fossé sémantique (modification de la Fig.1.1 de [6])  
(illustration retirée afin de respecter les droits d'auteurs)

Dans les domaines de la vision par ordinateur et de l'intelligence artificielle, la construction d'un système d'interprétation et de simplification de scènes est actuellement un axe très prometteur, notamment pour les applications de navigation autonome. Dans ce contexte, les systèmes actuellement utilisés, pour l'interprétation de l'environnement, apportent de nombreuses informations qui ne sont parfois pas nécessaires pour assurer une navigation autonome. Par conséquent, dans ce chapitre, inspiré par l'attention sélective humaine, nous proposons une stratégie de perception sélective qui permet de simplifier au maximum la scène observée tout en assurant une navigation autonome.

## IV.1 État de l'art

Les systèmes de vision peuvent fournir une mine d'informations sur l'environnement et ils constituent une alternative prometteuse, compte tenu de la disponibilité des capteurs de vision bon marché actuels et des processeurs hautes performances. Dans [148, 149], une caméra stéréo est utilisée pour localiser et construire une carte 3D de l'environnement. Il est vrai que le système de vision binoculaire peut facilement obtenir des informations de position 3D, mais il lui faut du temps pour faire correspondre les caractéristiques des deux images acquises par les deux caméras. Ainsi, pour diverses raisons à savoir le temps de traitement réduit, la légèreté, le faible coût, etc., de nombreux travaux sont basés sur des systèmes de vision monoculaires. Une vision monoculaire peut être réalisée par une caméra ordinaire ou une caméra omnidirectionnelle. Dans [150, 151], une caméra omnidirectionnelle a été utilisée pour la construction des cartes d'environnement. L'inconvénient majeur de ce capteur de vision est qu'une grande distorsion réduit sa précision. Les caméras ordinaires sont plus répandues que les caméras omnidirectionnelles. Elles sont utilisées dans diverses applications, telles que l'estimation de la position relative des robots [152], l'estimation de la position 3D des objets au sol [153], la localisation et cartographie simultanée [154], la détermination de la position d'un point au sol [155], la détection du plan du sol [156], le dimensionnement des objets [157], la détection des objets

[158, 159], etc. Généralement, les méthodes de détection du sol échouent lorsque la texture du sol varie ou que l'apparence des obstacles est très similaire à celle du sol. Autre méthodes basées sur la reconstruction 3D en utilisant le flux optique ont été proposé pour la détection d'obstacles bas difficiles à différencier du sol [160, 161]. Cependant, ces méthodes nécessitent des fonctions puissantes pour suivre les régions d'obstacles.

Basé sur une caméra monoculaire ordinaire, des travaux de détection de portes dans les environnements intérieurs ont été développés. Dans [162, 163], la forme rectangulaire a été considérée comme descripteur pour distinguer la porte d'un mur, alors que dans [162, 164], les coins de porte ont été utilisés pour la détection des portes. Kim *et al.* [165] ont montré que la porte est de la forme  $\Pi$ , qui est définie par le haut et les deux montants de la porte. Dans [166], Salinas a proposé une technique basée sur la transformée de Hough pour extraire des segments, puis les analyser par logique floue. La porte est détectée en identifiant la ligne horizontale correspondant à son sommet et les deux lignes verticales correspondant à ses montants. Chen *et al.* [162] ont proposé une méthode supervisée basée sur l'algorithme Adaboost, qui permet d'identifier une porte fermée à partir d'un ensemble de descripteurs (contour, couleur, texture), et d'un descripteur spécifique correspondant à l'écart entre le bas de la porte et le sol. En fait, quand la porte est fermée, cet écart est généralement caractérisé par une couleur sombre. Dans [167], une méthode de détection des portes de couloir a été proposée. Dans cette méthode, la première étape consiste à identifier le sol et le mur afin que le mur puisse être considéré comme une zone de recherche contenant des portes. La séparation sol / mur est basée sur l'emplacement du point de fuite, et les lignes détectées. Seki *et al.* [168] ont proposé une nouvelle méthode basée sur l'extraction d'une ou de deux ellipses, projetées depuis le projecteur du robot, pour détecter la limite entre le sol et le mur. Les auteurs ont découvert qu'une lumière conique projetée sur la zone où le sol et le mur se croisent, créant deux ellipses qui se chevauchent dans l'image. Ainsi, ils ont considéré que la ligne séparant ces deux ellipses est la frontière sol-mur.

L'approche d'interprétation de scène, que nous proposons, est principalement basée sur 3 étapes à savoir : 1) la détection des objets saillants de la scène, 2) la détermination du barycentre de chaque objet détecté et 3) la localisation de l'objet.

## IV.2 La saillance visuelle

L'attention est le lien clé entre la perception et la cognition. Le système visuel humain peut identifier rapidement et efficacement les régions les plus importantes de son environnement en se concentrant sur les régions d'intérêt. Par conséquent, la détection de saillance a été largement

utilisée par les chercheurs en neurosciences cognitives, en psychologie et en particulier en vision par ordinateur, telles que le reciblage d'images [169], segmentation d'objet [170], etc.

Dans la littérature, Ma et Zhang ont mentionné que la concentration de l'attention peut être déterminée par deux mécanismes, à savoir par stimulus (approche ascendante) ou par objectif (approche descendante) [171]. Les approches ascendantes sont étroitement liées aux caractéristiques de l'image telles que la couleur, le contour, le contraste, le gradient, etc. Contrairement aux méthodes descendantes qui sont lentes, les méthodes ascendantes sont rapides, imprévisibles et directes.

Dans ce qui suit, nous nous concentrons uniquement sur les méthodes ascendantes. Il est à noter que ces méthodes transforment l'image d'entrée en une carte de saillance, où chaque pixel / superpixel se voit attribuer une valeur ou une probabilité de saillance.

### IV.2.1 Modèles de prédiction de la fixation oculaire

L'un des premiers modèles de prédiction de fixation oculaire est développé par Itti *et al.* [172]. Ce modèle a suscité la première vague d'intérêt dans de multiples disciplines, notamment la psychologie cognitive, la vision par ordinateur, etc. Il s'agit d'une implémentation d'une approche ascendante basée sur les mécanismes centraux (comme l'architecture de l'attention computationnelle de Koch et Ullman [173], le modèle de recherche guidée de Wolfe *et al.* [174] et la théorie de l'intégration des caractéristiques de Treisman et Gelade [175]). Harel *et al.* [176] ont proposé un modèle de saillance basé sur l'emplacement du pixel et sa valeur de poids, qui est proportionnelle à la dissemblance déterminée en fonction de direction et de changement d'intensité. Achanta *et al.* [177] ont proposé un modèle qui consiste à générer une carte de saillance à partir de l'aberration chromatique de l'image entière après avoir utilisé le flou dans la région environnante. Le contraste est également utilisé pour identifier les points saillants d'apparence complexe [172, 178] à travers des modèles graphiques [176], des modèles de probabilité bayésiens [179], un histogramme de co-occurrence de couleur [180], etc. D'autres modèles de prédiction de fixation oculaire sont basés sur l'analyse spectrale [181, 182]. Ils sont équivalents aux opérateurs de gradient locaux. Par conséquent, les grandes zones saillantes ne peuvent pas être bien détectées.

### IV.2.2 Modèles de détection des régions saillantes

Bien que les cartes de saillances générées par les modèles de fixation soient très clairsemées, le but des modèles de détection des régions saillantes est de générer une carte dont les limites des régions saillantes sont bien définies (Cf. Figure IV.2).

À l'instar des modèles de prédiction de fixation oculaire, de nombreux modèles de détection des régions saillantes reposent sur l'hypothèse que le contraste entre l'objet et ses régions environnantes est élevé [177, 183, 184, 185, 186]. Lou *et al.* [187] ont proposé un modèle pour détecter les régions saillantes à partir du contraste des couleurs dominantes, qui combine des descripteurs visuels de bas et de moyen niveaux. Inspiré de Lou [187], Cheng *et al.* [188] ont développé un modèle de contraste "HC" basé sur l'histogramme et un deuxième modèle de contraste "RC" basé sur les régions. Perazzi *et al.* [184] ont ignoré les détails inutiles en décomposant l'image en des régions compactes et visuellement uniformes, ce qui crée les régions saillantes. Jiang *et al.* [189] ont utilisé le contraste de couleur global et local pour détecter les régions saillantes. Dans [190, 191], les auteurs ont incorporé différentes mesures de contraste couramment utilisées, à savoir le contraste global, le contour central, le contour des coins, afin de générer leur carte de saillance. Qian *et al.* [192] ont proposé un nouveau concept de contraste optimal basé sur le calcul de l'entropie pour détecter le meilleur contraste des régions saillantes. Afin de mettre en évidence de manière uniforme les régions d'objet, ces modèles utilisent la compacité spatiale, l'uniformité de la région, etc.

En plus du contraste, de nombreuses méthodes récentes [183, 185] ont été développées sur la base des bordures d'image. Néanmoins, ces méthodes considèrent que toutes les bordures de l'image appartiennent à l'arrière-plan. Ceci est très fragile et peut échouer même si l'objet ne touche que légèrement la bordure. Afin de résoudre ce problème, Zhu *et al.* [193] ont proposé une nouvelle méthode de saillance basée sur la mesure de la connectivité aux frontières [193]. Cette mesure ne considère pas que la bordure d'image [183, 185], ou qu'un patch d'image connecté à la bordure d'image [194] comme l'arrière-plan, mais elle considère qu'un patch d'image n'est un arrière-plan que si la région à laquelle il appartient est très liée à la bordure d'image. La séparation des objets de l'arrière-plan est basée sur une mesure de contraste améliorée, appelée contraste pondéré en arrière-plan. Compte tenu des caractéristiques spatiales des régions d'arrière-plan, la méthode proposée par Zhu [193] est robuste et stable aux changements de contenu d'image, et peut effectuer une détection d'arrière-plan de haute précision. Pour cette raison, nous avons choisi d'utiliser ce modèle d'interprétation de scène.

En utilisant des techniques d'apprentissage automatique, certains modèles [195, 196, 197] ont obtenu d'excellentes performances. Ces modèles peuvent gérer des situations difficiles qui ne peuvent être résolues par des méthodes sans apprentissage. Néanmoins, ils ont besoin d'une formation hors ligne coûteuse basée sur des images entièrement annotées. Beaucoup de ces modèles sont basés sur les réseaux de neurones convolutifs CNN utilisant des fonctions complexes

au niveau des régions. Ainsi, ces modèles sont généralement beaucoup plus lents que les modèles non appris.

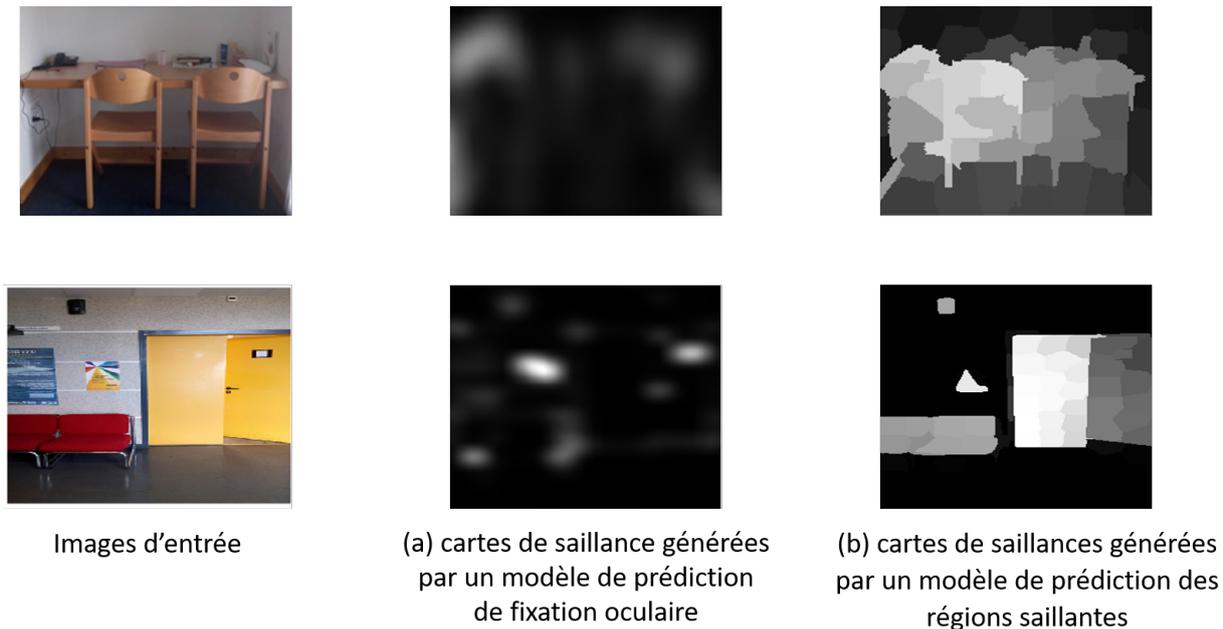


FIGURE IV.2 – Génération de cartes de saillance en appliquant (a) un modèle de prédiction de fixation oculaire et (b) un modèle de prédiction des régions saillantes

## IV.3 Approche d'interprétation de scène par une caméra monoculaire RVB

La Figure IV.3 présente notre approche pour l'interprétation d'une scène du monde réel acquise par une caméra monoculaire RVB. Dans ce qui suit, nous discutons plus en détail chaque étape de l'approche proposée.

### IV.3.1 Création d'une image panoramique

Prenons une scène réelle balayée horizontalement par une caméra monoculaire. Ainsi, la scène d'espace (3D) sera convertie en vidéo (2D). Afin de développer une stratégie simple, le problème d'interprétation de la scène du monde réel est simplement transformé en problème d'interprétation de l'image panoramique de la vidéo de la scène acquise. La création d'image panoramique est réalisée par l'éditeur de composition d'images (Image Composite Editor ICE). Il s'agit d'une application open source créée par Microsoft pour assembler des images ou utiliser des vidéos pour composer des panoramas.

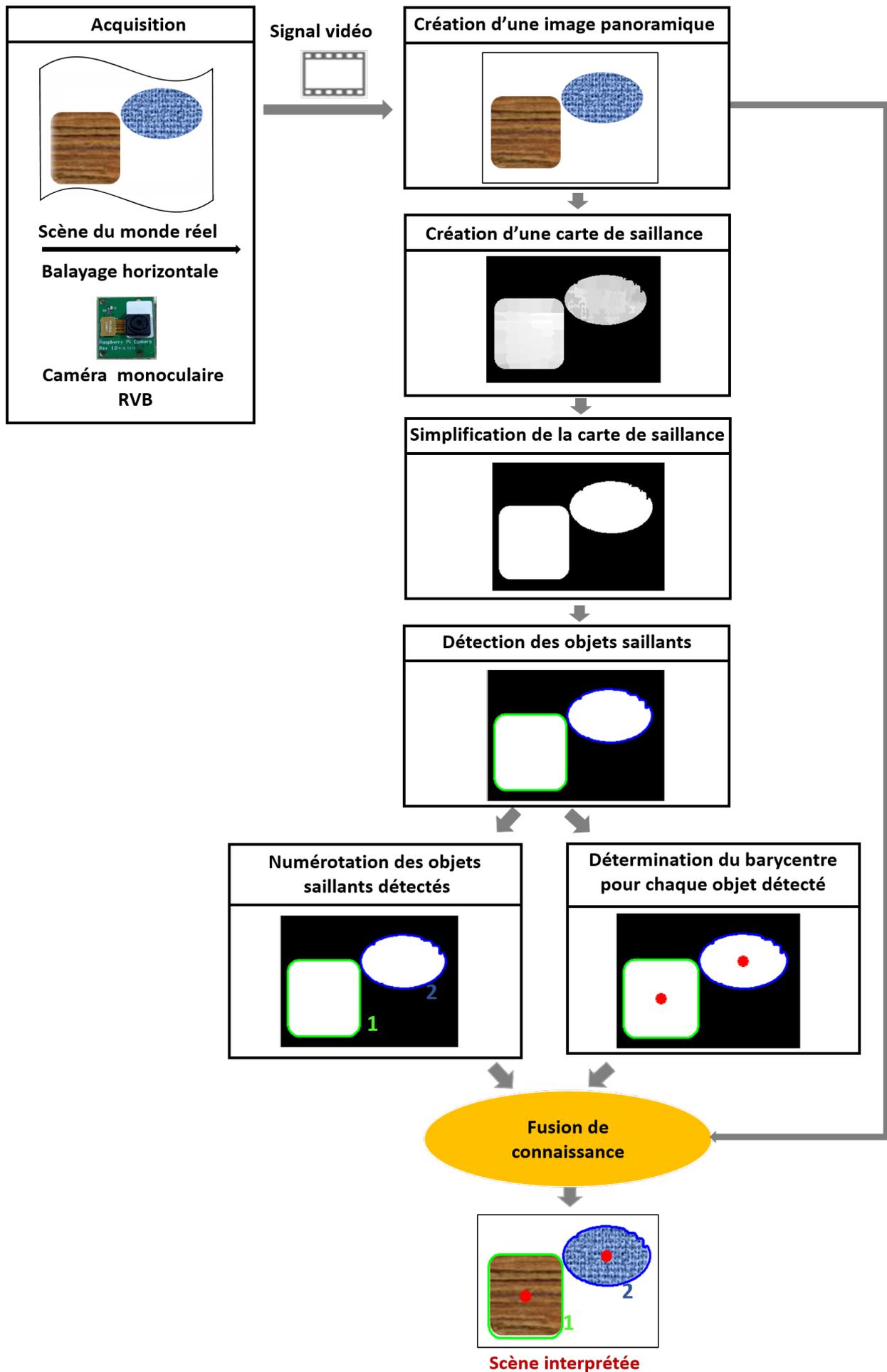


FIGURE IV.3 – Approche d'interprétation de scène à partir d'une caméra monoculaire RVB



(a) Logo



(b) Interface

FIGURE IV.4 – Présentation de ICE

La création du panorama se fait en 4 étapes :

Étape1 (Import) : Importer une vidéo

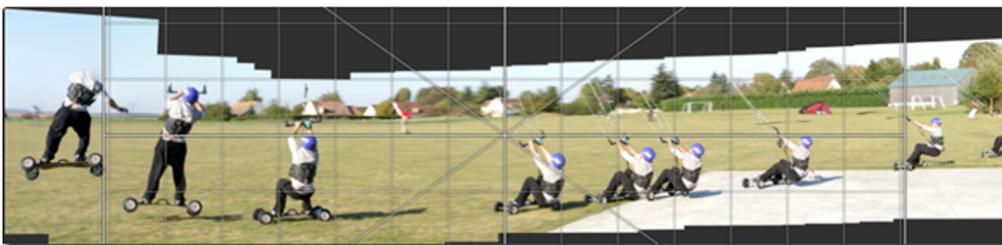
Étape2 (Stitch) :

- Segmenter la vidéo
- Aligner les images
- Assembler les images

Étape3 (Crop) :

- Extrapoler les pixels manquants d'un assemblage
- Recadrer le panorama

Étape4 (Export) : Enregistrer le panorama sous une format d'image.



(a)



(b)

FIGURE IV.5 – Exemples de panoramas (a) avant et (b) après l'extrapolation

### IV.3.2 Génération d'une carte de saillance (1<sup>ère</sup> simplification de scène)

Notre objectif étant de rendre l'interprétation de la scène la plus simple possible, nous avons choisi de ne traiter que les objets d'intérêt au lieu d'analyser tous les objets présents dans l'image panoramique. À cette fin, nous convertissons l'image panoramique en une carte de saillance ne représentant que les régions saillantes. La génération de notre carte de saillance est basée sur le modèle proposé par Zhu *et al.* [193]. Ce modèle est basé sur l'hypothèse que les régions d'objet et les régions d'arrière-plan ont des agencements spatiaux différents, c'est-à-dire que la connexion entre la région d'objet et le bord de l'image est moins importante que la connexion entre la région d'arrière-plan et le bord de l'image. Comme le montre la Figure IV.6, l'image synthétique contient 4 régions, dont la région jaune présente la région saillante, car c'est la région la plus grande, la plus compacte et ne touchant que légèrement le bord de l'image. Bien que la région verte ne touche que partiellement le bord de l'image, ce n'est pas une région saillante en raison de sa taille réduite. En d'autres termes, elle ressemble plus à un objet partiellement rogné. La région blanche et la région bleue touchent clairement les bords de l'image, ce sont donc évidemment des régions d'arrière-plans.

FIGURE IV.6 – Exemple illustratif de connexions aux limite (illustration retirée afin de respecter les droits d'auteurs)

Afin de quantifier le degré de connexion entre la région  $R$  et la frontière de l'image, Zhu *et al.* [193] ont proposé une mesure appelée connectivité aux limites. Cette mesure est définie comme le rapport entre le périmètre de la région et la frontière  $R$  au périmètre globale de la région :

$$BndCon(R) = \frac{|\{p|p \in R, p \in Bnd\}|}{\sqrt{|\{p|p \in R\}|}} \quad (IV.1)$$

Où :

$Bnd$  : l'ensemble des patchs des bords de l'image

$p$  : un patch d'image

La segmentation d'image est un problème difficile à résoudre, ce qui rend l'Équation IV.1 difficile à calculer. Cependant, l'utilisation de techniques de segmentation dure implique un problème de sélection d'algorithmes / paramètres, ainsi que des artefacts de discontinuité indésirables au bords de la région. Par conséquent, Zhu *et al.* [193] ont souligné qu'il n'est pas nécessaire de procéder à une segmentation précise et dure de l'image. Au lieu de cela, ils ont

recouru à la méthode SLIC [198] pour segmenter l'image d'entrée en un ensemble de superpixels. Ainsi, dans un esprit similaire à l'Équation IV.1, la connectivité aux limites est définie comme :

$$BndCon(R) = \frac{Len_{bnd}(p)}{\sqrt{Area(p)}} \quad (IV.2)$$

Où :

$Len_{bnd}$  : la longueur du superpixel

$p$  : le long de la frontière

Area : la zone de recouvrement du superpixel  $p$

Afin d'optimiser la carte de saillance, certains chercheurs ont utilisé des méthodes d'optimisation basées sur CRF / MRF [185, 199]. Cependant, ces méthodes sont compliquées et généralement lentes. D'autres chercheurs ont optimisé de manière heuristique la carte de saillance par somme pondérée ou multiplication [184, 185]. Dans [193], les auteurs ont développé une fonction de coût (Équation IV.3) qui atteint aussitôt l'objectif de détection des régions d'objets saillants. La fonction est formée de trois termes sous forme linéaire. Le premier terme encourage que la saillance de la région d'arrière-plan soit faible. Le deuxième terme oblige que la saillance de la région d'objet soit élevée. Le dernier terme représente une fonction de lissage pour assurer d'avoir une carte de saillance uniforme sur les régions plates.

$$\sum_{i=1}^N \omega_i^{bg} s_i^2 + \sum_{i=1}^N \omega_i^{fg} (s_i - 1)^2 + \sum_{i,j} \omega_{ij} (s_i - s_j)^2 \quad (IV.3)$$

Où :

$s_i$  : la valeur de saillance du superpixel  $p_i$

$\omega_i^{bg}$  : probabilité que le superpixel  $p_i$  appartienne à la région d'arrière-plan

$\omega_i^{fg}$  : probabilité que le superpixel  $p_i$  appartienne à la région d'objet

$\omega_{ij}$  : poids pour chaque paire de superpixels adjacente (i,j)

### IV.3.3 Simplification de carte de saillance (2<sup>ème</sup> simplification de scène)

Afin de simplifier au maximum le processus d'interprétation de scène, nous avons appliqué des opérateurs morphologiques comme une 2<sup>ème</sup> simplification de scène. Pour ce faire, une étape de binarisation des cartes de saillance est requise.

### IV.3.3.1 Binarisation

Après avoir généré la carte de saillance, on la convertit en une image binaire à l'aide de la méthode d'Otsu [200]. Elle s'agit d'une technique de seuillage la plus basique et la plus rapide pour binariser une image.

La carte de saillance peut être considérée comme une fonction bidimensionnelle, contenant  $N$  pixels, dont les valeurs de niveaux de gris de 1 à  $L$ . En considérant que  $f_i$  est le nombre de pixels du  $i^{\text{ème}}$  niveau de gris, la probabilité de ce niveau de gris dans la carte de saillance est égale :

$$P_i = \frac{f_i}{N} \quad (\text{IV.4})$$

Afin de trouver un seuil approprié entre les pixels appartenant à la région d'objet et les pixels appartenant à la région d'arrière-plan, les pixels sont répartis en deux classes :  $C_1$  comprend les niveaux de gris de  $[1, \dots, t]$  et  $C_2$  comprend les niveaux de gris de  $[t + 1, \dots, L]$ . Les distributions de probabilité correspondantes à  $C_1$  et  $C_2$  sont, respectivement, calculées comme suit :

$$\begin{aligned} \omega_1 &= \sum_{i=1}^t P_i \\ \omega_2 &= \sum_{i=t+1}^L P_i \end{aligned} \quad (\text{IV.5})$$

De plus, les valeurs moyennes des classes  $C_1$  et  $C_2$  sont respectivement calculées comme suit :

$$\begin{aligned} m_1 &= \frac{1}{\omega_1} \sum_{i=1}^t iP_i \\ m_2 &= \frac{1}{\omega_2} \sum_{i=t+1}^L iP_i \end{aligned} \quad (\text{IV.6})$$

Sur la base de la méthode Otsu, la variance intra-classe de la carte de saillance est définie par l'Équation IV.7. Le seuil optimal est sélectionné de façon que la variance intra-classe  $v_B^2$  soit minimisée.

$$v_B^2 = \omega_1(m_1 - m)^2 - \omega_2(m_2 - m)^2 \quad (\text{IV.7})$$

Où :

$v_B^2$  : la variance intra-classe

$m$  : l'intensité moyenne pour la carte de saillance entière.

Une fois la carte de saillance est binarisée, nous passons à la simplification de scène en appliquant les opérateurs morphologiques.

### IV.3.3.2 Opérateurs morphologiques

Les opérations morphologiques sont une extension de la théorie des ensembles de Minkowski [201] et sont largement utilisées comme techniques de traitement d'image. Une transformation morphologique génère une image avec la même taille que l'image d'entrée mais avec des caractéristiques différentes. Les opérations morphologiques sont plus adaptées aux images binaires qu'aux images monochromes (niveaux de gris). Dans cette section, nous nous intéressons aux opérateurs morphologiques appliqués aux images binaires, qui peuvent se résumer comme suit :

**L'érosion** : généralement utilisée pour réduire les objets de l'image en élargissant la largeur des régions minimales et en réduisant les pics. L'érosion d'une image  $f$  par un élément structurant  $B$  est représentée par :

$$[\varepsilon_B(f)](x) = \bigwedge_{b \in B} f(x + b) \quad (\text{IV.8})$$

**La dilatation** : utilisée pour étirer ou rétrécir l'image d'entrée en élargissant la largeur des régions maximales. La dilatation d'une image  $f$  par un élément structurant  $B$  est définie par :

$$[\delta_B(f)](x) = \bigvee_{b \in B} f(x + b) \quad (\text{IV.9})$$

**L'ouverture** : est simplement réalisée en appliquant une érosion suivie d'une dilatation en utilisant le même élément structurant  $B$  pour les deux opérations. L'idée derrière l'ouverture morphologique est d'éliminer les objets de taille plus petite que l'élément structurant en appliquant une opération d'érosion, puis de restaurer la forme des objets restants par une opération de dilatation.

$$\gamma_B(f) = \delta_B[\varepsilon_B(f)] \quad (\text{IV.10})$$

**La fermeture** : est une opération d'ouverture effectuée dans l'ordre inverse. Contrairement à l'ouverture, son but est de restaurer l'image dilatée.

$$\phi_B(f) = \varepsilon_B[\delta_B(f)] \quad (\text{IV.11})$$

L'ouverture et la fermeture par reconstruction sont plus utiles que les opérateurs d'ouverture et de fermeture classiques. Ils permettent d'éliminer les zones indésirables sans affecter de manière significative le reste des objets de l'image. De plus, ils aident à restaurer les structures qui ont été partiellement endommagées par la dilatation ou l'érosion. Cette caractéristique découle de la façon dont la transformation est construite par une transformation géodésique de dilatation et d'érosion. L'ouverture et la fermeture par reconstruction sont basées sur deux images : une image est traitée et définie comme l'image de marque, notée  $f$ , et la seconde image est considérée comme l'image de masque, notée  $g$ . L'ouverture et la fermeture par reconstruction

de  $f$  par rapport à  $g$  sont respectivement définies par l'Équation IV.12 et IV.13.

$$\delta_g^l(f) = \delta^l(f) \wedge g \quad (\text{IV.12})$$

$$\varepsilon_g^l(f) = \varepsilon^l(f) \vee g \quad (\text{IV.13})$$

où :

$\wedge$  : signifie prendre le minimum point par point

$\vee$  : signifie prendre le maximum point par point

#### IV.3.4 Détection des objets saillants

Une fois que la carte de saillance est binarisée et simplifiée par l'opérateur ouverture par reconstruction, l'étape suivante consiste à détecter et numéroter chaque région d'objet dans la carte de saillance puis de déterminer son barycentre, qui est ensuite utilisé dans l'attribution une étiquette de position selon l'algorithme I.1.

#### IV.3.5 Résultats et discussions

Nos expérimentations consistent à évaluer notre approche proposée, qui vise à simplifier au maximum l'interprétation de la scène tout en assurant une navigation autonome. Dans notre méthode, deux niveaux de simplification sont portés à savoir sur : (i) la génération de la carte de saillance (ii) l'application d'opérateurs morphologiques. Le modèle de saillance est évalué en fonction de la précision et de la simplicité de la carte de saillance qu'il génère. Deux types de modèles de saillance sont appliqués. Pour les modèles de détection des régions saillantes, on a par exemple SEG [202], AC [177], RBD [193], etc.) et des modèles qui prédisent la fixation oculaire (comme SIM [203], SR [204], COV [178], etc.) sont utilisés pour réaliser ces tests. De la Figure IV.7, on peut tirer les conclusions suivantes :

- Généralement, les modèles de détection d'objets saillants génèrent des cartes de saillance plus performantes que les modèles de prédiction de fixation.
- La précision de la carte de saillance générée par le modèle RBD surpasse les autres cartes générées par les différents modèles de détection d'objets.
- La précision de la carte de saillance générée par le modèle de fixation SIM est la pire.

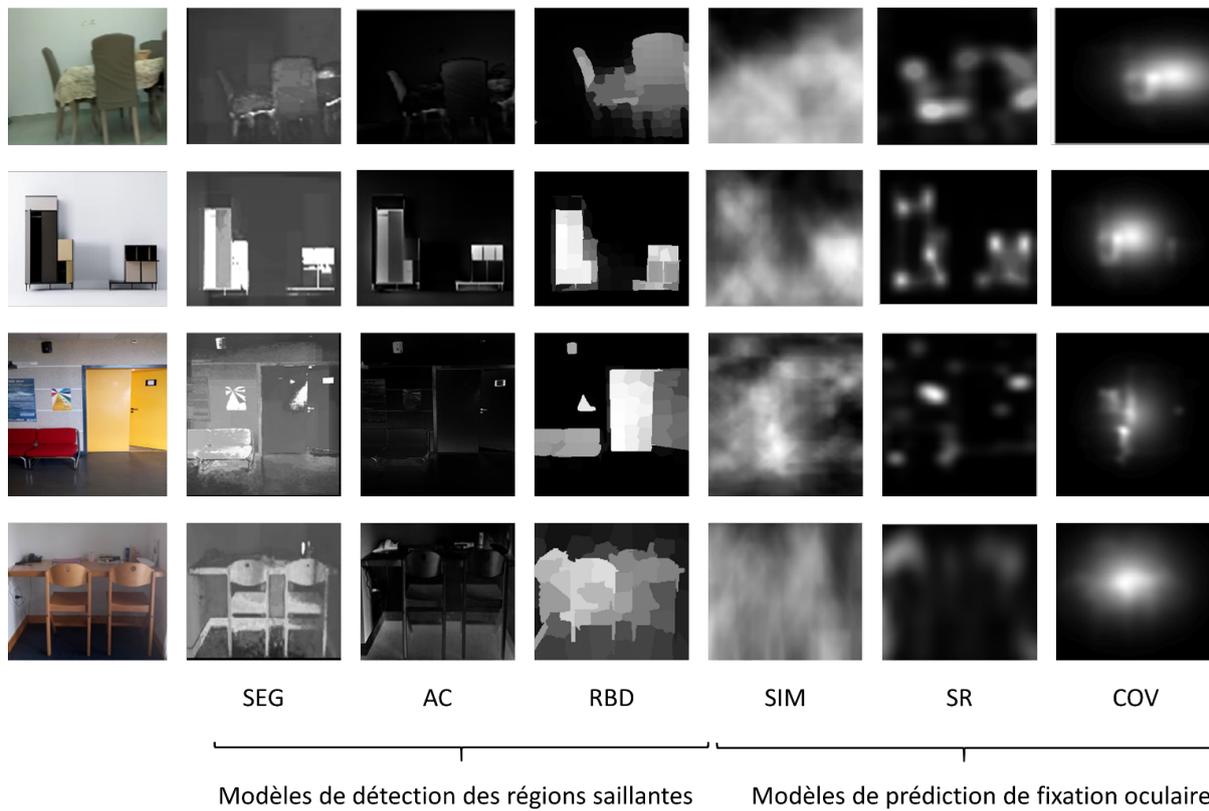


FIGURE IV.7 – Cartes de saillance générées appliquant différents (a) modèles de détection d'objets saillants et (b) modèles de prédiction de fixation oculaire

Le temps de traitement est aussi important que la précision de la carte de saillance. Les résultats obtenus à partir d'une dizaine d'images, sont présentée dans le Tableau IV.1, montrent :

- Le modèle SR est le modèle le plus rapide (environ 0,04 ms).
- Le temps de traitement du modèle AC est proche de RBD, mais sa carte de saillance n'est pas très utile pour le traitement.
- Le modèle COV est le plus lent modèle.

TABLE IV.1 – Comparaison du temps de traitement requis pour chaque modèle.

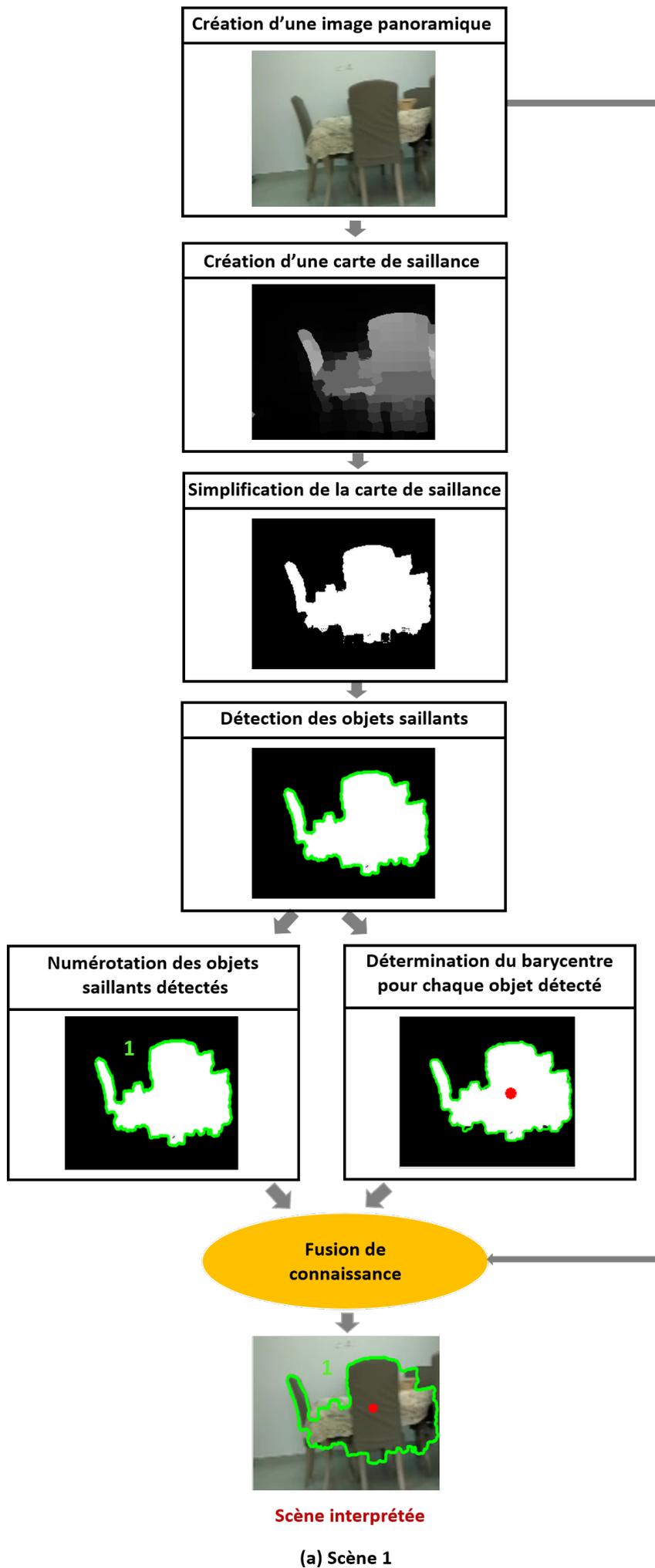
Model	Salient object models			Fixation prediction models		
	SEG	AC	RBD	SIM	SR	COV
Code Type	Matlab C++	Matlab C++	Matlab	Matlab	Matlab	Matlab
<i>Time(ms)</i>	11.4	0.125	0.130	1.12	0.04	25.3

La Figure IV.8 présente les résultats de notre méthode proposée appliquée à différentes scènes. Les images de sortie ont moins de détails que les images d'entrée en raison de l'application de la simplification à deux niveaux. Bien que la quantité d'informations soit réduite, notre offre une navigation en toute sécurité. En effet, les informations nécessaires pour une telle navigation sont incorporées généralement dans les objets d'intérêt présentés dans la scène.

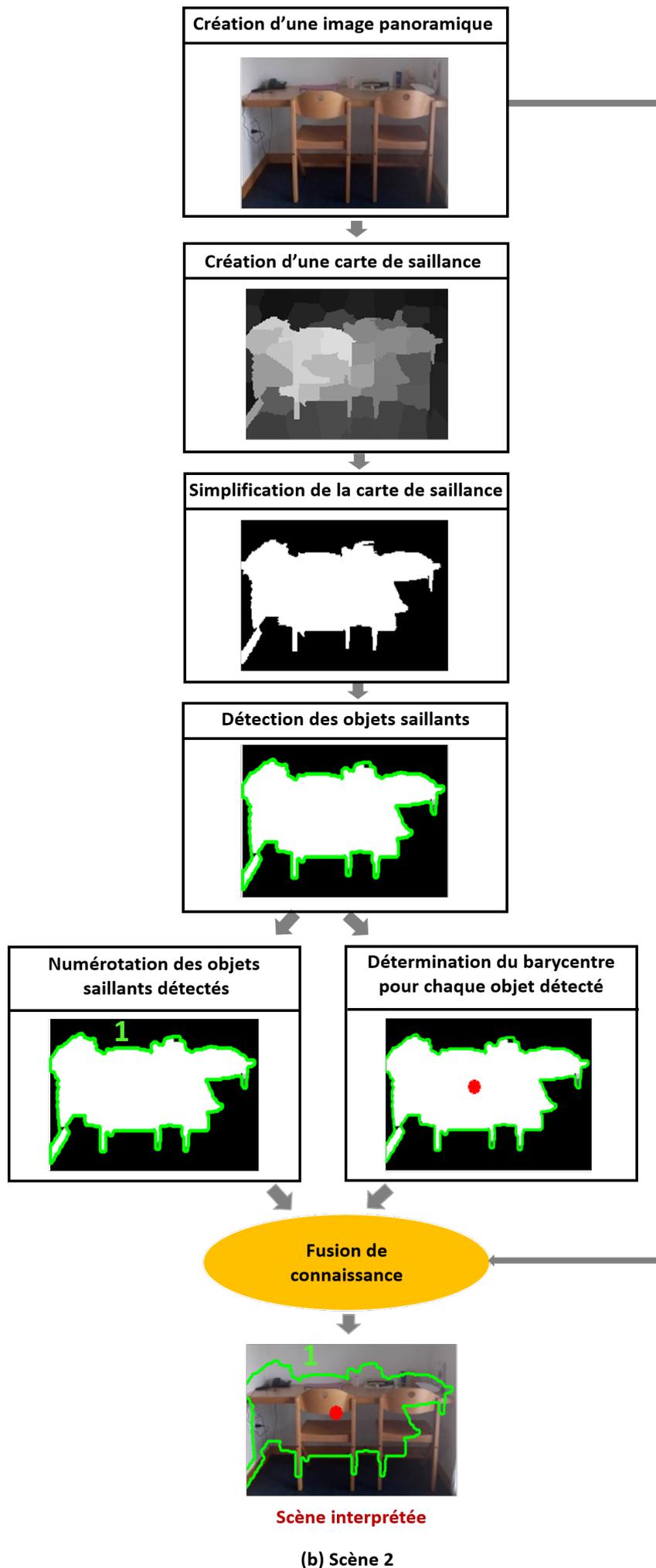
## IV.4 Conclusion

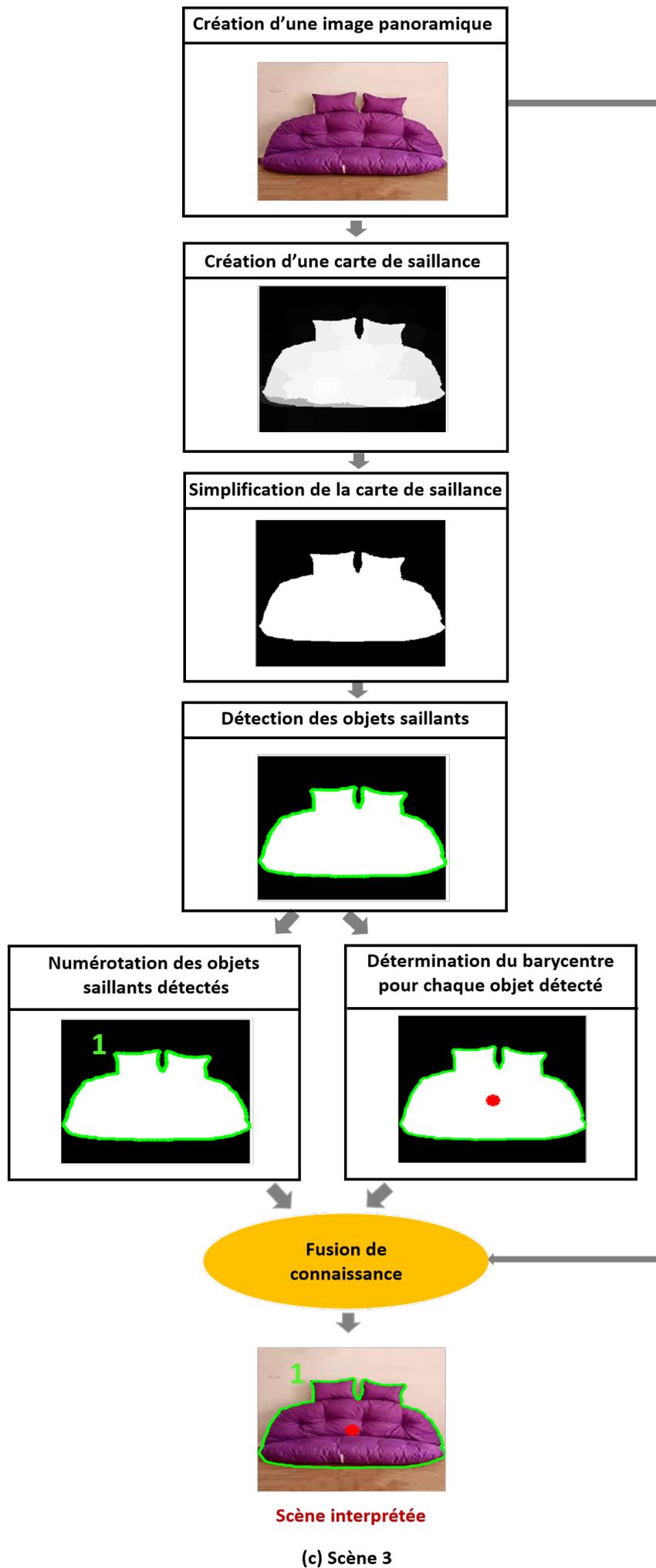
Dans ce chapitre, une approche d'interprétation basée sur une simplification de scène observée par une caméra monoculaire RVB a été proposée. La première simplification de scène est effectuée en générant une carte de saillance en utilisant le modèle de détection d'objet saillant RBD. Ainsi, seuls les objets d'intérêt sont présentés. La deuxième simplification consiste à simplifier davantage la carte de saillance en utilisant l'opérateur morphologique ouverture par reconstruction. Enfin, le barycentre de chaque objet détecté est calculé pour contribuer par la suite à attribuer une étiquette de positionnement (Cf. Algorithme I.1)

Dans le chapitre suivant, nous proposons une stratégie d'interprétation de scène perçue par un système multi-capteurs équipé de capteurs ultrasons et d'une caméra monoculaire RVB. Ainsi, le système d'interprétation de scène est basé à la fois sur les connaissances obtenues par l'analyse et le traitement des données ultrasonores et des données de caméra.



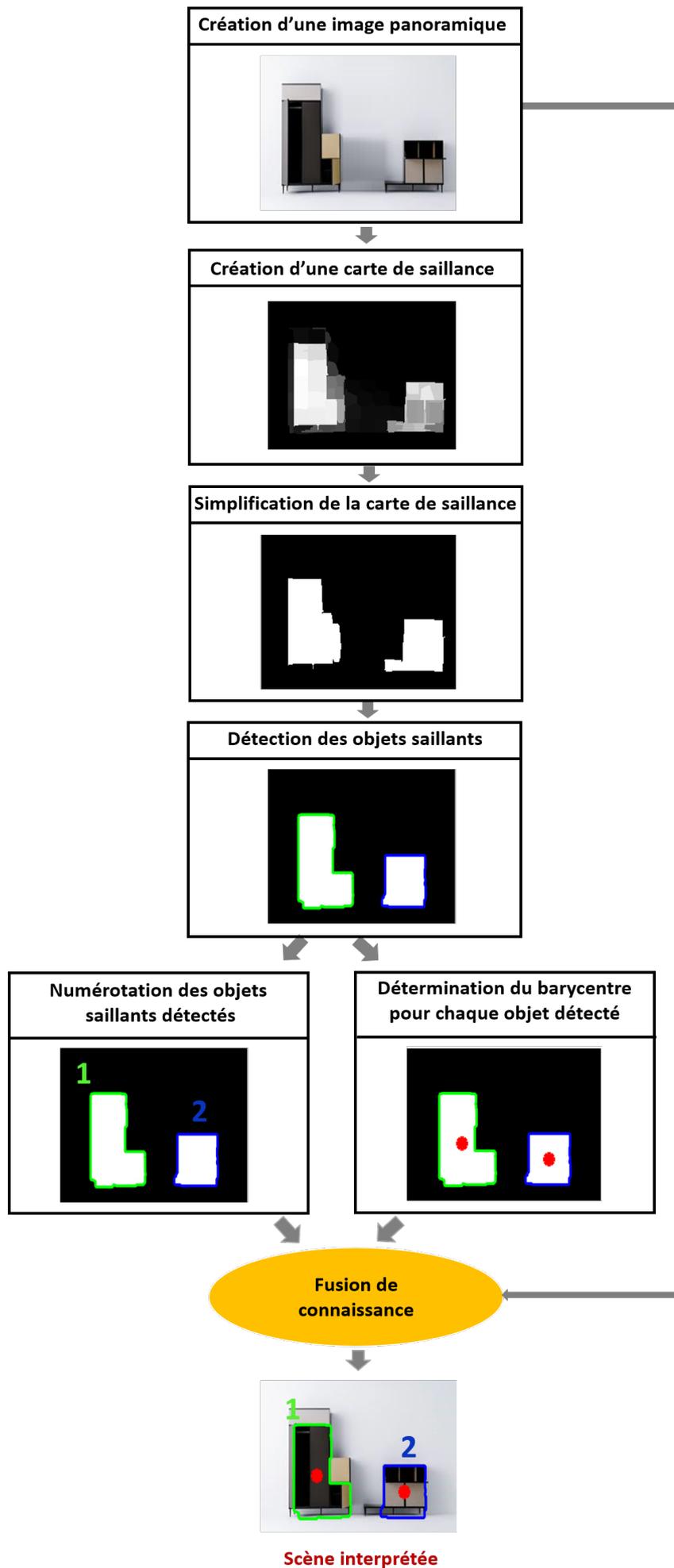
(a) Scène 1

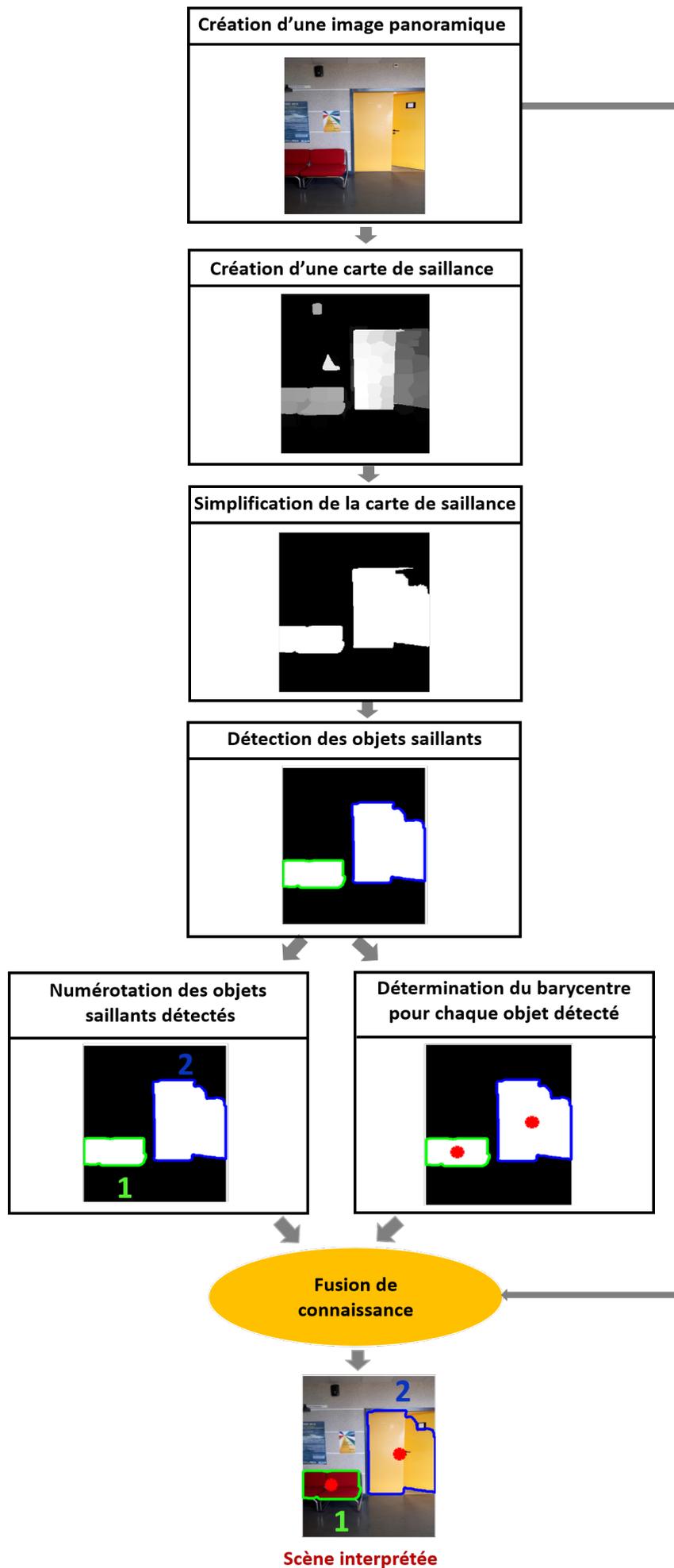


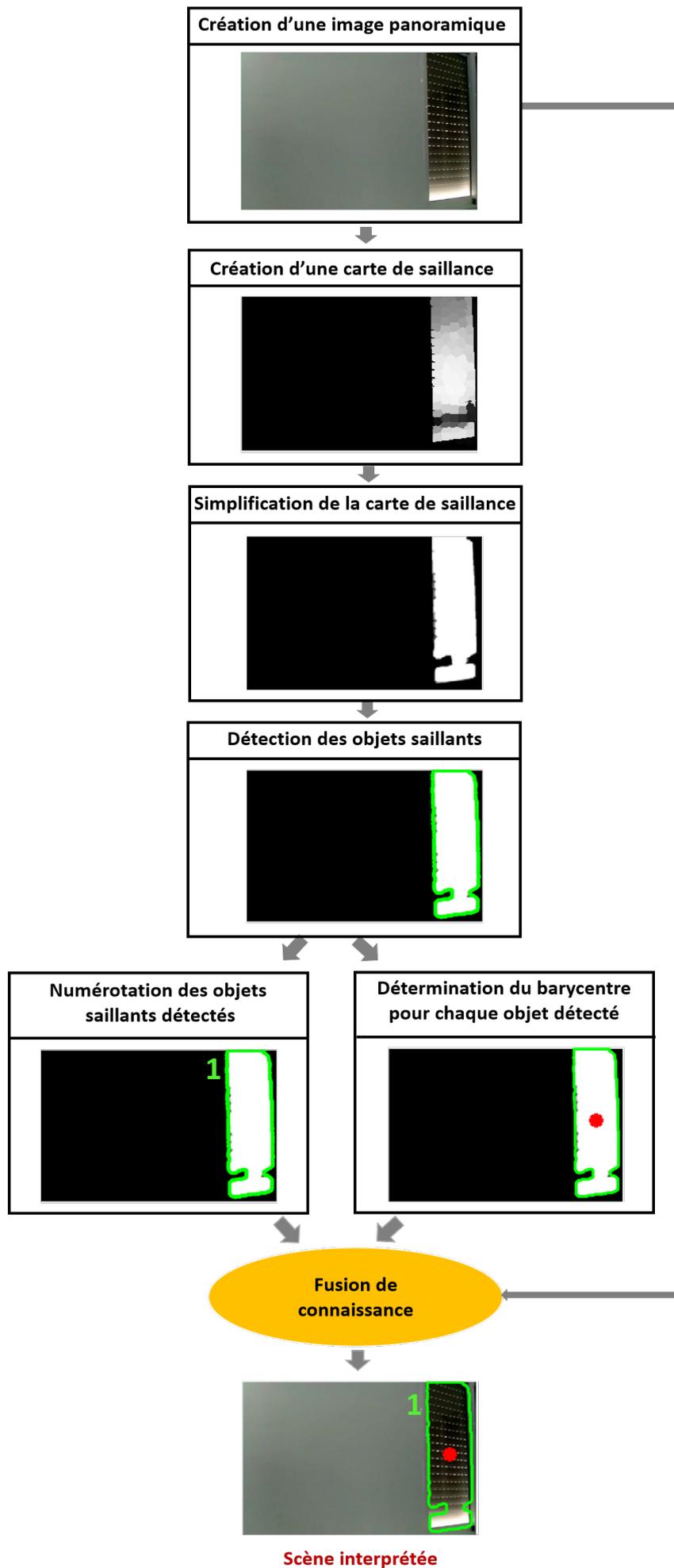


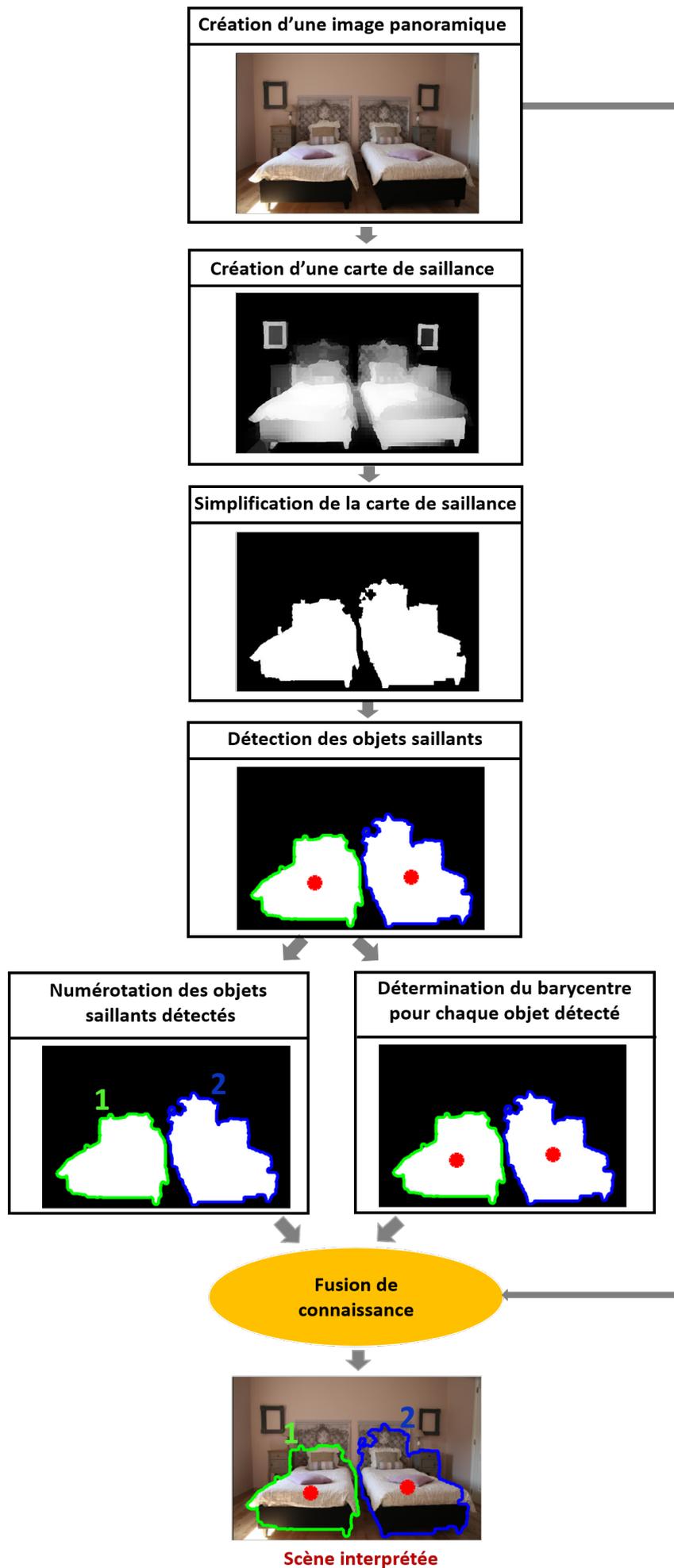
Scène interprétée

(c) Scène 3

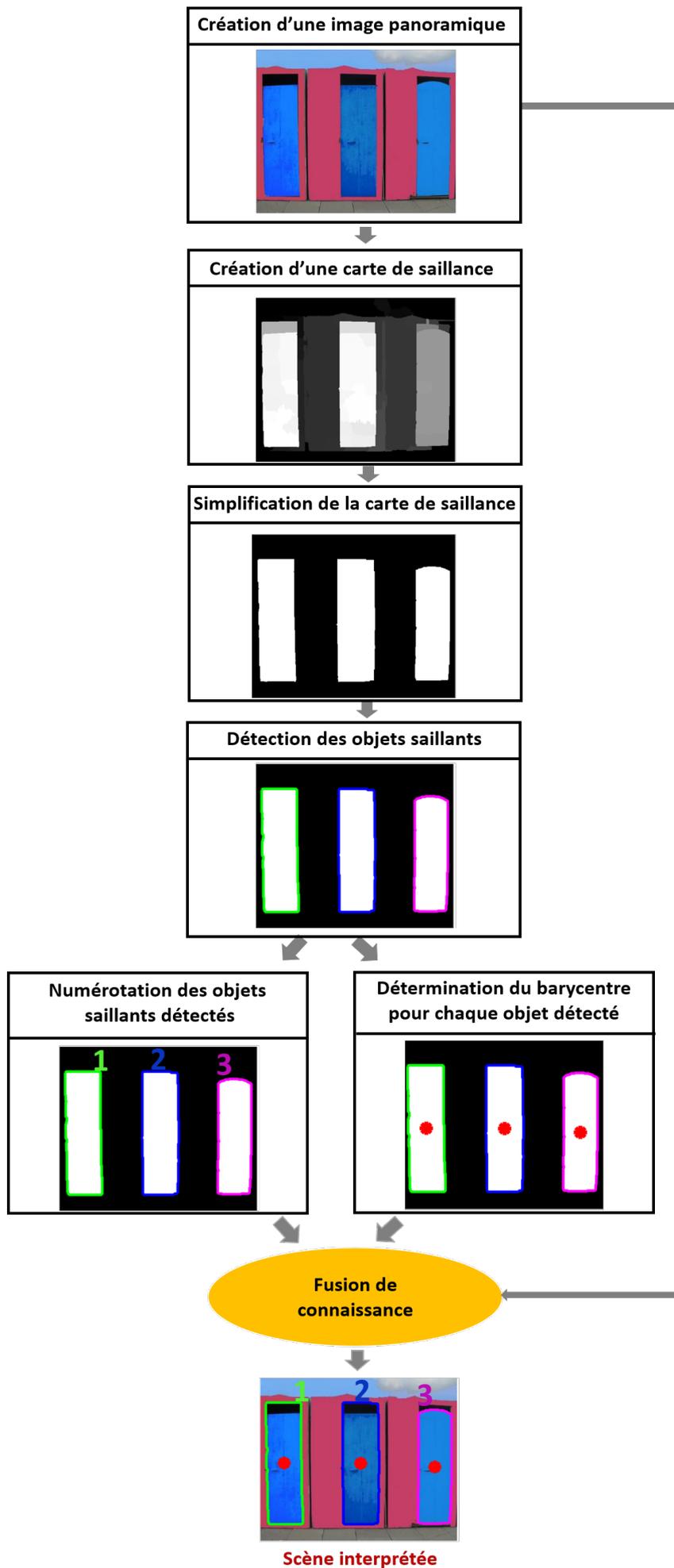








(g) Scène 7



(h) Scène 8

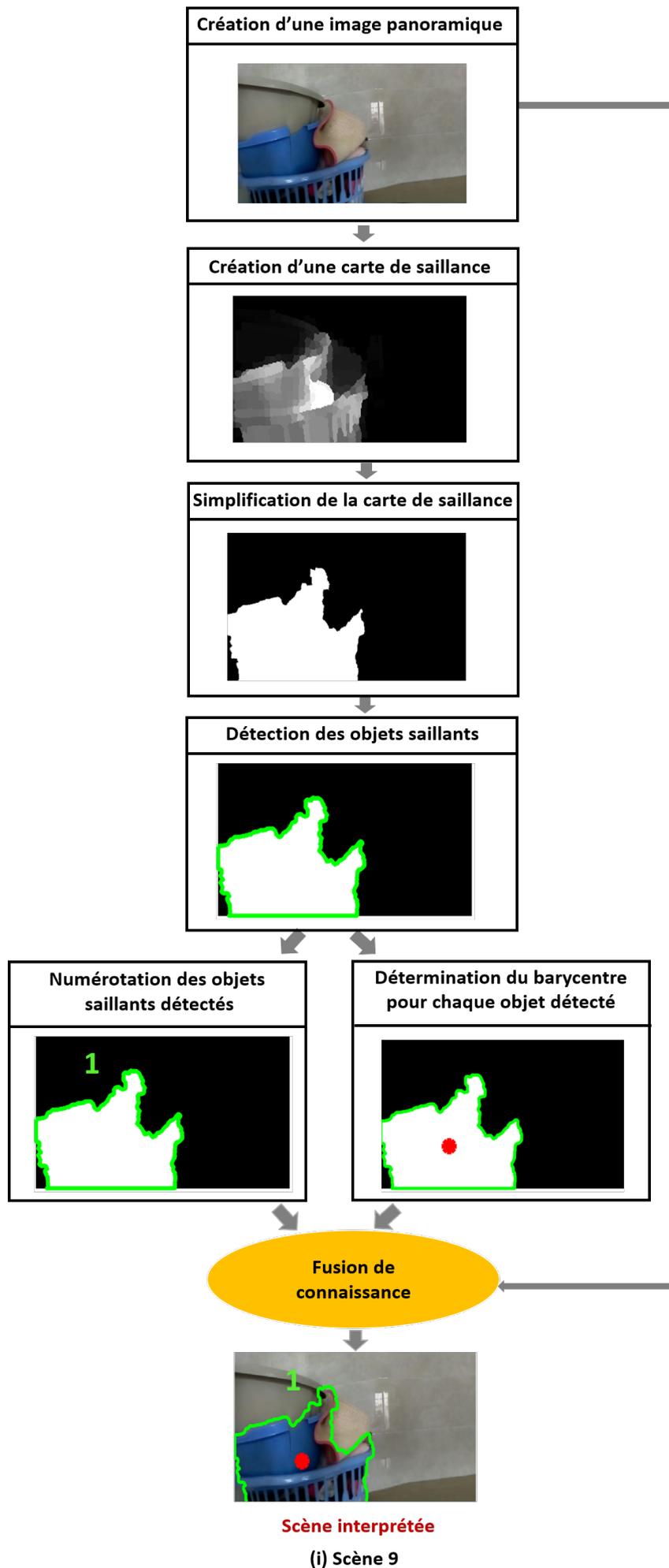


FIGURE IV.8 – Application de la stratégie proposée sur un exemple de scénarii

# Chapitre V

## Perception de scène par un système multi-capteurs, application à l'aide de navigation des déficients visuels

### Sommaire

---

V.1	État de l'art sur les systèmes d'aide à la navigation des déficients visuels . . . . .	126
V.2	Présentation du système d'aide à la navigation ("NA_System" : Navigation Aid System) . . . . .	128
V.3	Interprétation de scène par fusion de connaissances . . . . .	134
V.4	Conclusion . . . . .	141

---

## Introduction

De nos jours, selon l'Organisation Mondiale de la Santé (OMS), il existe plus de 2,2 milliards de personnes non-voyantes et malvoyantes dans le monde [205]. Une personne non-voyante est une personne qui a perdu la vue en raison d'un accident, d'une cécité à la naissance, ou d'une maladie oculaire (telle que le trachome, le diabète, la cataracte congénitale, le traumatisme oculaire, et la dégénérescence maculaire liée à l'âge [205]). En revanche, une personne malvoyante est celle ayant une réduction de l'acuité visuelle de loin et de près, mais conservant une partie de leur enveloppe de vision [206].

Compte tenu de leurs problèmes de communication et de déplacement, de nombreuses personnes non-voyantes et malvoyantes se sentent socialement marginalisées. Tout en les intégrant dans la société et en assurant leur sécurité, l'utilisation par des outils d'aide peut être d'un grand intérêt pour eux. Certaines personnes sont satisfaites des outils classiques bon marché, tels que la canne blanche, le chien-guide, etc. Ces outils, leurs permettent, d'une part, de distinguer des personnes en bonne santé pour avertir autrui de leur handicap, et d'autre part, de leur éviter quelques dangers. D'autres déficients visuels s'engagent à surmonter leur déficience en utilisant des aides électroniques à la navigation plus sophistiqués pour combler leurs lacunes de perception.

Dans ce chapitre, un état de l'art sur les systèmes d'aide à la navigation des personnes à déficience visuelle est présenté. Ensuite, une étude détaillée du prototype "NA\_System" est présentée. Dans ce qui suit, nous présentons notre approche de recalage et de fusion de différentes données acquises par NA\_System. Enfin, nous appliquons le modèle générique de l'interprétation de scène à un ensemble de scènes.

## V.1 État de l'art sur les systèmes d'aide à la navigation des déficients visuels

La vision est l'un des sens humains les plus importants, et son absence provoque un déséquilibre dans la vie des personnes qui en est privée. La perte de vision s'accompagne généralement d'une perte d'indépendance. Les malvoyants, en particulier les non-voyants, ne peuvent pas profiter des divers services, y compris l'interaction sociale, les opportunités éducatives et professionnelles, et la mobilité qui a été définie par Foulke [207] comme "*the ability to travel safely, comfortably, gracefully and independently through the environment*".

La mobilité dans un environnement intérieur est plus critique que la mobilité dans un

environnement extérieur [208]. Les systèmes de navigation extérieurs sont généralement équipés d'un système de positionnement global (GPS) pour localiser les aveugles. Cependant, dans un environnement intérieur, les signaux GPS ne peuvent pas être reçus [209]. Ainsi, la plupart des chercheurs utilisent le GPS pour les systèmes de navigation extérieurs, alors que pour les systèmes de navigation intérieurs, ils utilisent différentes technologies, telles que des capteurs infrarouges, des capteurs ultrasons, des caméras, etc.

Plusieurs systèmes d'aide à la navigation ont été proposés dans la littérature pour avertir les personnes à déficience visuelle par des vibrations ou des sons audibles (Cf. Figure V.1). Certains de ces systèmes reposent sur l'utilisation de capteurs de profondeur, tels que des capteurs lasers [211], des capteurs infrarouges [212, 213] et des capteurs ultrasons [214, 215, 216, 217, 218, 219]. D'autres systèmes utilisent une seule caméra ou une caméra stéréo [213, 220]. Dans [221], une enquête approfondie sur l'applicabilité et la pertinence de ces systèmes a été élaborée. Par exemple, il a été montré que ces dispositifs peuvent rencontrer de nombreux problèmes :

- Il couvrent une grande partie des oreilles, bloquent la langue, etc. Par conséquent, les malvoyants ne peuvent jamais être à l'aise. De plus, ce ne sont pas assez faciles d'être manipulés par les enfants.
- L'utilisation de ces appareils nécessite beaucoup de formation, ce qui est particulièrement difficile pour les enfants.
- Ces dispositifs peuvent même fonctionner de manière insatisfaisante.
- Le coût de ces systèmes est trop élevé pour être payés par une personne ordinaire.

FIGURE V.1 – Quelques systèmes d'aides à la navigation des aveugles [210] (illustration retirée afin de respecter les droits d'auteurs)

D'autres systèmes d'aide à la navigation ont été proposés dans la littérature pour contrôler la navigation des personnes aveugles et malvoyantes à travers des systèmes de messagerie auditive. En fait, la sortie audio est plus riche en contenu que la vibration, car elle permet aux utilisateurs d'obtenir des instructions claires et simples de direction ou de description de scène. Certains chercheurs utilisent des capteurs ultrasons [222, 223, 224] pour détecter les objets, ainsi que d'autres [225, 226], utilisent des capteurs infrarouges pour détecter les escaliers. Suivant le principe l'Active Badge [227], une grille de 16 émetteurs infrarouges [228] a été construite pour déterminer la distance entre l'utilisateur et les objets. Dans [229], un système d'assistance basé sur une caméra stéréo a été proposé pour déterminer la profondeur des objets dans la scène. Bai *et al.* [230] ont proposé des lunettes intelligentes équipées d'une caméra de profondeur et

d'un capteur ultrason pour estimer la profondeur [230]. De tels systèmes informent uniquement l'utilisateur de la distance à l'obstacle, sans l'identifier ni le décrire. Cette simple information de distance ne peut pas offrir suffisamment de confort et de sécurité aux malvoyants.

D'autres systèmes basés uniquement sur des capteurs de vision pour aider les malvoyants ont été proposés [231, 232]. Ces systèmes indiquent uniquement le nom de l'objet détecté à l'utilisateur, sans lui fournir d'aperçu de la distance qui le sépare de l'objet.

Plusieurs systèmes équipés par une caméra monoculaire et des capteurs ultrasons ont été proposés [233, 234, 235]. Dans ces systèmes les capteurs ultrasons sont utilisés uniquement pour déterminer la distance des objets détectés. La détection d'objets est basée sur le classifieur d'images YOLO (You Only Look Once) [233], ou sur Tensorflow [234, 235] qui sont tous les deux basés sur un réseau de neurone conventionnel. Il est à noter que pour un tel système, sans un bon GPU, le système devient très lent. De plus, ces classifieurs d'images permettent la détection de tous les objets présents dans la scène. Cependant, certains objets ne nécessitent pas d'être avertis pour les malvoyants car ils ne fournissent aucune information primordiale pour une navigation autonome.

D'autres systèmes sont basés sur des caméras ToF pour aider les personnes à déficience visuelle à naviguer dans leur environnement intérieur [236, 237, 238]. Cependant, ces systèmes sont lourds, nécessitent beaucoup d'énergie pour fonctionner, et sont encombrants en termes de processeur et de mémoire.

TABLE V.1 – Systèmes d'aide à la navigation des personnes à déficience visuelle [210]  
(illustration retirée afin de respecter les droits d'auteurs)

## **V.2 Présentation du système d'aide à la navigation ("NA\_System" : Navigation Aid System)**

Les systèmes d'aide à la navigation disponibles sur le marché avertissent convenablement l'utilisateur de la présence d'obstacles. Ceci est une information primordiale pour les personnes malvoyantes. Le "NA\_System" conçu peut également parfaitement remplir cette fonction, mais prend aussi en compte le fait que les utilisateurs désirent d'être informés de ce qu'est l'obstacle, et à quoi ressemble la scène. En effet, grâce à l'utilisation des capteurs ultrasons accompagné par une caméra, le "NA\_System" assure :

- une détection précise des obstacles
- une détermination de distance séparant ces objets à l'utilisateur

- une reconnaissance de ces obstacles (couleur, forme, rigidité matérielle et dimension)
- une identification robuste de l'état du sol (plat, escalier montant, escalier descendant).

Les systèmes d'aides à la navigation disponibles sur le marché informent les utilisateurs par le biais de sensations tactiles, des vibrations ou bips sonores. Par conséquent, une formation d'utilisation est essentielle, ce qui n'est pas à la portée de la cognition ou même de la portée matérielle de tous les utilisateurs. Pour le *NA\_System*, le retour est riche en information, facile à comprendre et acceptable par les utilisateurs. En fait, il s'agit d'un retour vocal, dans la langue sélectionnée par l'utilisateur lors de la réception du dispositif, précédé par une légère vibration (avertissant l'utilisateur qu'un message vocal est sur le point d'être reçu.). Par conséquent, aucune formation n'est requise [7].

La détection et la reconnaissance d'obstacles sont deux tâches de base pour les systèmes d'aide à la navigation. Les systèmes de détection d'obstacles disponibles sur le marché permettent la détection d'objets situés au même niveau que le sol, mais ils ne sont pas aptes pour détecter des objets dont le niveau est plus bas que celui du sol, à savoir des marches descendantes, des escaliers descendants ou encore des fossés. Ainsi, le fait de ne pas reconnaître ces obstacles constitue une menace réelle pour la sécurité des déficients visuels. Dans notre invention, contrairement à d'autres dispositifs, nous accordons une attention particulière aux obstacles qui constituent une menace sérieuse pour la sécurité des personnes à déficience visuelle [7].

## V.2.1 Historique de *NA\_System*

### V.2.1.1 Premier dispositif

Le premier dispositif conçu est une canne blanche électronique équipée de multiples capteurs, tels que : deux capteurs ultrasons pour détecter les objets en hauteur et identifier l'état du sol, une caméra, un microprocesseur, une carte SD, un module Bluetooth, et une batterie. Comme le montre la Figure V.2, la présence du chariot au bout de la canne la rend lourde et donc inconfortable pour les déficients visuels à se déplacer. Par conséquent, un deuxième dispositif a été proposé pour surmonter cette limitation.



FIGURE V.2 – Premier dispositif

### V.2.1.2 Deuxième dispositif

Le deuxième dispositif s'agit également d'une canne blanche électronique, plus légère et plus pratique que la première canne. En fait, comme le montre la Figure V.3, le support est remplacé par une petite pointe en caoutchouc. De plus, la tige utilisée pour ce deuxième dispositif est pliable.



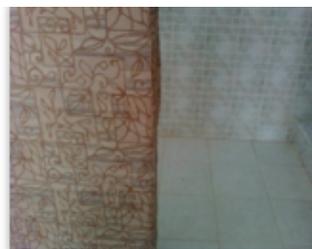
FIGURE V.3 – Deuxième dispositif

### V.2.1.3 Limitations des deux dispositifs

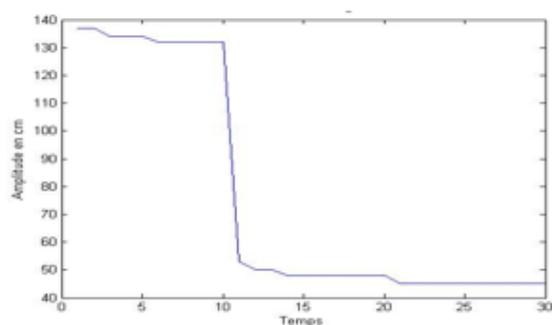
Les deux dispositifs mentionnés ci-dessus ne sont pas capables de détecter la direction dans laquelle l'utilisateur balaye la scène, ce qui conduit à une mauvaise interprétation de la scène. Comme le montre la Figure V.4, pour la même scène, le signal de profil construit suite d'un balayage de gauche à droite est différent de celui construit suite d'un balayage de droite à gauche. En fait, cela peut poser des questions sur la position de l'utilisateur par rapport aux objets présentés dans la scène.

Il convient de noter également que lors de la marche ou du balayage d'une scène par l'un de ces dispositifs, les données acquises sont trop perturbées. Ainsi, un troisième dispositif nommé NA\_System équipé d'un gyroscope et d'un accéléromètre a été conçu pour détecter ces per-

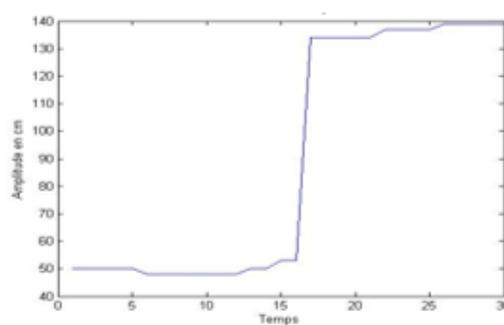
turbations. Ce troisième dispositif est également équipé des servomoteurs, qui permettent de s'assurer que la position des capteurs est constamment mise à jour en temps réel pour compenser ces perturbations. Par conséquent, le NA\_System se caractérise par un système d'acquisition stable capable d'acquérir des données de meilleure qualité que les deux premiers dispositifs. Dans ce qui suit, nous présentons ce troisième dispositif plus en détail.



a- Exemple de scène



b- Signal de profil construit suite à un scan de droite à gauche



c- Signal de profil construit suite à un scan de gauche à droite

FIGURE V.4 – Influence de la direction de balayage sur la construction du signal de profil

## V.2.2 Description du NA\_System

Le système d'aide à la navigation "NA\_System" est un dispositif multi-capteurs conçu par notre équipe en collaboration avec la société NOVELTI pour aider les personnes malvoyantes et non-voyantes à se déplacer dans leur environnement et à mieux surmonter leurs handicaps. La conception du "NA\_System" est illustrée à la Figure V.5.

Le "NA\_System" remet en question les limites indiquées par les produits existant sur le marché, et d'autre part, il répond aux besoins et attentes recueillis auprès d'utilisateurs après une recherche de besoin menée à cet effet.

Le "NA\_System" est de poids et dimension très appropriés pour être fixé sur une canne blanche ordinaire ou porté comme un pendentif. En effet, les plaintes de style des futurs utilisateurs sont en fait divisées en deux opinions basées sur l'âge : les plus jeunes cherchent une assistance dans l'ombre afin de ressembler le plus possible à des personnes en bonne santé. En conséquence, le "NA\_System" est fourni sous forme d'un boîtier qui est facile à installer sur une ceinture et qui ressemble à un téléphone portable. Bien que pour les personnes moins jeunes,

la canne blanche n'est pas seulement un moyen d'explorer et d'éviter les obstacles, mais aussi un symbole pour encourager les autres à céder le passage et même à les aider davantage. Par conséquent, ils veulent conserver cette apparence symbolique. En conséquence, le NA\_System est fourni sous forme d'une boîte, qui peut être facilement fixée à leur canne blanche [7].

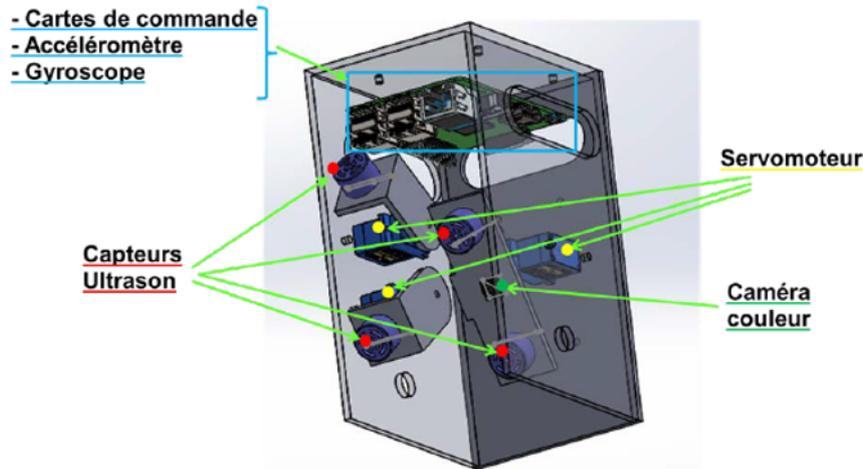


FIGURE V.5 – Vue en 3D du dispositif du le système d'aide à la navigation conçu [7]

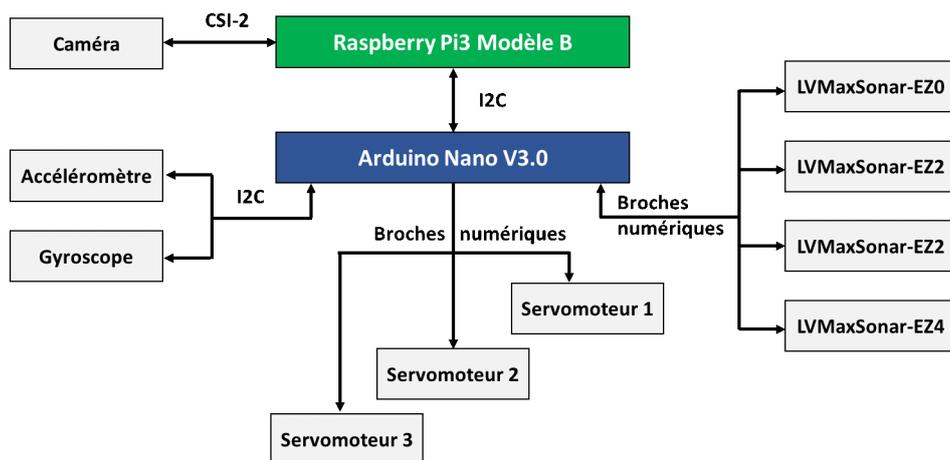


FIGURE V.6 – Schéma fonctionnel du "NA\_System" conçu

### V.2.2.1 Capteurs utilisés

Le "NA\_System" est équipé de :

- Deux capteurs ultrasons dont la position peut être ajustée en fonction de la taille de l'utilisateur. Afin de détecter les obstacles en hauteur, l'un des deux capteurs est orienté vers le haut, et l'autre vers le bas pour identifier l'état du sol (plat, escalier montant et escalier descendant) ;
- Une caméra Raspberry, connectée à la carte Raspberry Pi 3 via le protocole CSI-2, est en liaison encastrement avec deux autres capteurs ultrasons. L'ensemble vise à acquérir la

- scène présente devant l'utilisateur de façon à la reconstruire en profondeur et permettre au NA\_system de détecter et de reconnaître les différents objets de la scène ;
- Carte Raspberry Pi 3 Model B est utilisée comme une carte mère dans le dispositif "NA\_System". Cette carte dispose d'un processeur plus puissant qu'une carte de deuxième génération. Le Tableau V.2 présente une comparaison entre une carte Raspberry de troisième génération et une autre carte de deuxième génération ;
  - Carte Arduino Nano V3.0 est utilisée comme une carte esclave pour la manipulation des capteurs ultrasons. Le Tableau V.3 présente les différentes spécifications de cette carte ;
  - Trois servomoteurs sont utilisés pour maintenir la position de divers capteurs afin de maintenir la stabilité de leurs acquisitions ;
  - Un gyroscope et un accéléromètre pour compenser les changements de position des capteurs lors de l'utilisation du "NA\_System" ;
  - Carte mémoire SD montée sur la carte Raspberry pour enregistrer les données acquises ;
  - Vibreur pour avertir l'utilisateur avant de recevoir le message vocal ;
  - Une oreillette Bluetooth pour la transmission des informations aux utilisateurs ;
  - Batterie rechargeable.

FIGURE V.7 – Carte Raspberry Pi 3 Model B (illustration retirée afin de respecter les droits d'auteurs)

TABLE V.2 – Comparaison entre Raspberry Pi 2 et Raspberry Pi 3. (illustration retirée afin de respecter les droits d'auteurs)

TABLE V.3 – Spécification de la carte Arduino Nano V3.0 [252]. (illustration retirée afin de respecter les droits d'auteurs)

### V.2.2.2 Espace de travail

Lors de l'acquisition de la scène, la carte raspberry Pi3 communique avec le django web "SmartCane" grâce au protocole SSH (Cf. Figure V.8). L'interface de ce site web est illustrée à la Figure V.9.

```

pi@SmartCane: ~
DELL@DELL-PC ~
$ ssh pi@172.24.1.1
raspi@172.24.1.1's password:

The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Sun Nov 13 06:17:09 2016
pi@SmartCane:~$

```

FIGURE V.8 – Connexion à SmartCane

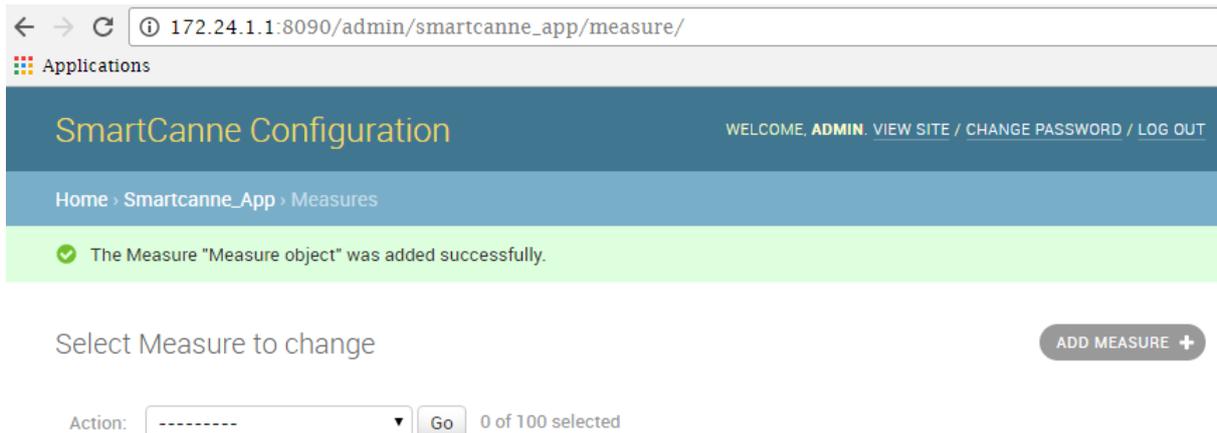


FIGURE V.9 – Interface du django Web "SmartCane"

## V.3 Interprétation de scène par fusion de connaissances

### V.3.1 Adaptation du modèle générique d'interprétation de scène

Comme mentionné ci-dessus, NA\_System est un dispositif multi-capteurs. Dans ce qui suit, nous supposons que la scène présentée devant l'appareil est acquise que par le module équipé d'une caméra monoculaire RVB et de deux capteurs ultrasons situés en dessous et au-dessus de la caméra (voir Figure V.10).

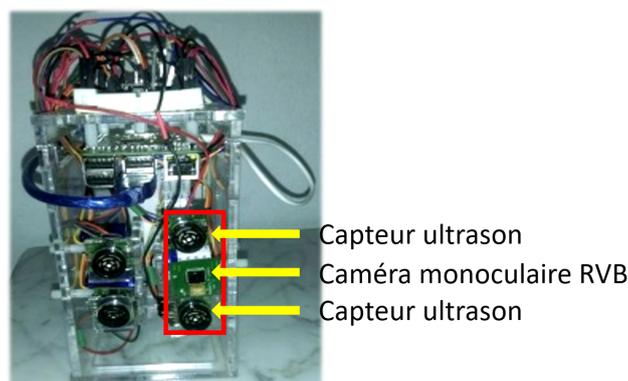


FIGURE V.10 – Module de capteurs considéré pour la perception de scène

La stratégie d'interprétation de scène proposée est basée sur un modèle simpliste (voir Figure

I.12) pour décrire la scène environnante, partageant le champ intercepté par les capteurs en trois zones principales, à savoir : face, gauche et droite (Cf. Algorithme I.1). Notre stratégie consiste d'abord à balayer la scène horizontalement. Ensuite, les différents types de données collectées par le module en question sont d'abord traités, puis analysés séparément. En d'autres termes, pour les données ultrasonores, la méthode de détection et d'évitement d'obstacle proposée dans la Section III.1 est utilisée afin d'attribuer une étiquette de profondeur à chaque objet détecté. De plus, l'approche de reconnaissance de rigidité matérielle proposée dans la Section III.2.3 convient également pour attribuer une étiquette "Objet dur" ou "Objet mou" à chaque objet détecté. Concernant les données de la caméra, la stratégie décrite au Chapitre IV est appliquée afin de numérotter les objets présentés dans la scène et d'identifier leur barycentre. Enfin, les informations relatives aux capteurs ultrasons (signaux de profils) et les informations relatives à la caméra (image panoramique) sont recalées. Suite au recalage, et selon le modèle de scène proposé, ces informations sont attribuées conjointement aux objets de la scène. La Figure V.11 présente une adaptation du modèle générique illustré à la Figure I.14 en fonction de notre application.

### V.3.2 Recalage des informations

Le recalage des informations relatives aux capteurs ultrasons et des informations relatives à la caméra nécessitent la détermination sur quel pixel de l'image le capteur ultrason pointe. Pour cela, certaines connaissances *a priori* doivent être prises en compte, à savoir : la distance entre le centre du capteur ultrason et le centre de la caméra (Cf. Figure V.12), la vitesse de balayage, le sens de balayage, le champ de vision [253] (Cf. Figure V.13), etc.

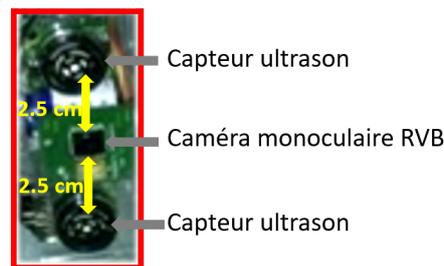


FIGURE V.12 – Positionnement des capteurs dans NA\_System

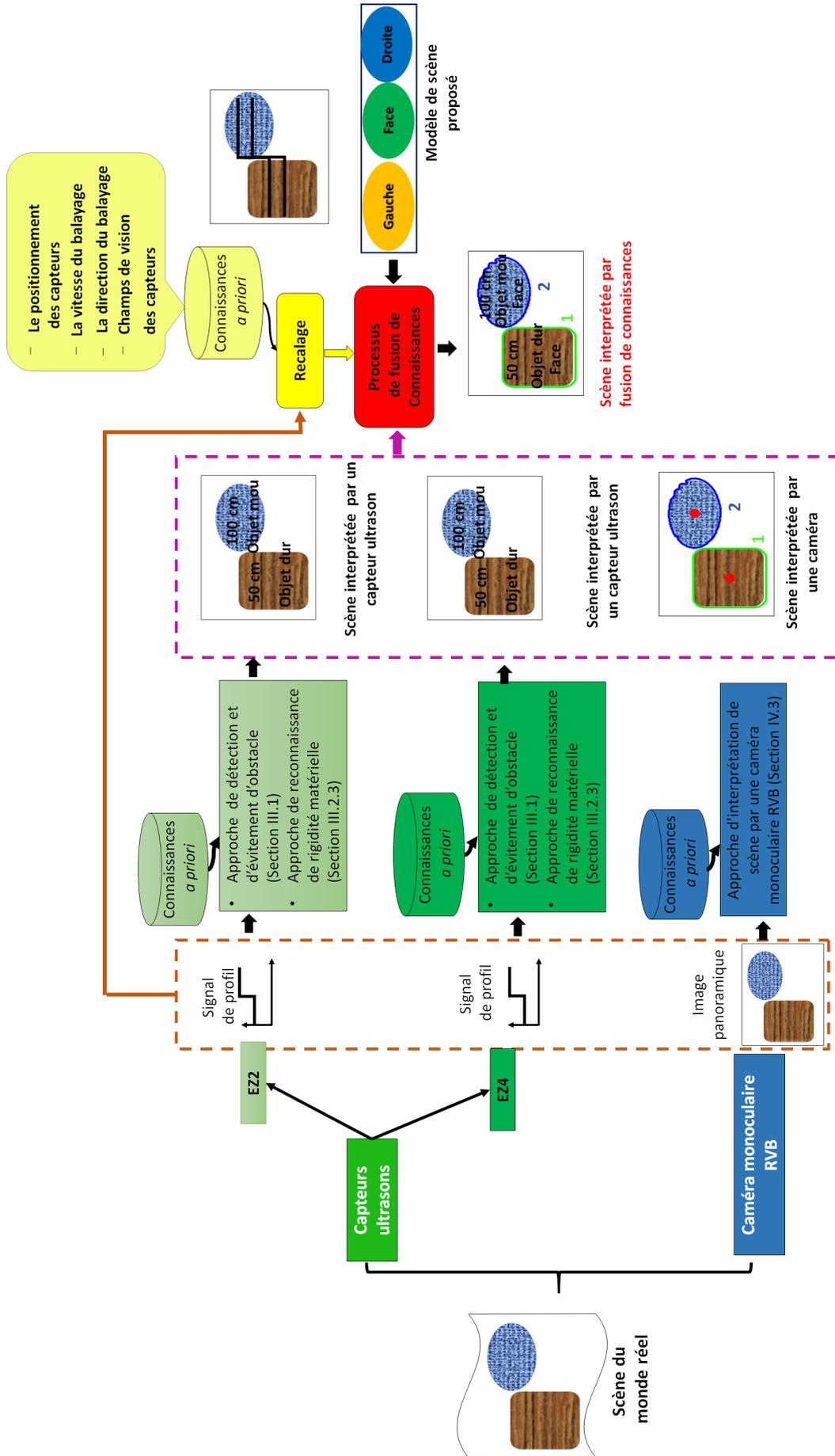


FIGURE V.11 – Système d'interprétation de scène perçue par des capteurs ultrasons et une caméra monoculaire RVB

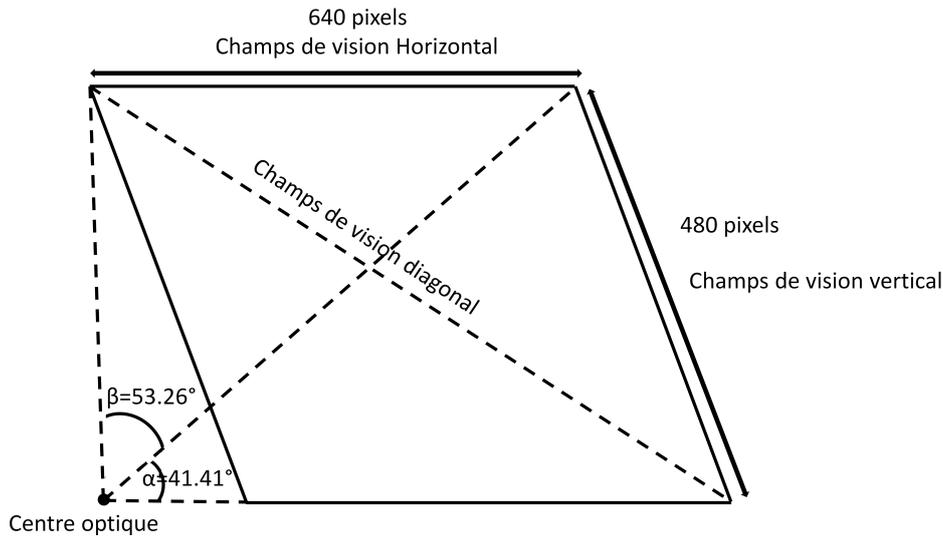


FIGURE V.13 – Différents champs de vision pour la caméra intégrée dans NA\_System

### V.3.2.1 Détermination du pixel pointé par le capteur ultrason dans l'image

Comme le montre la Figure V.12, le module considéré est équipé par deux capteurs ultrasons alignés verticalement avec la caméra. Dans la suite, nous considérons que les centres de ces trois capteurs sont alignés dans l'image. Après avoir mené plusieurs expériences, nous avons observé que plus l'objet est éloigné du NA\_System, plus le capteur ultrason pointe vers le centre de l'image. Ainsi, nous avons opter de trouver une fonction permettant de déterminer le pixel auquel le capteur ultrason pointe dans l'image, par rapport au centre de l'image, en fonction de la distance séparant le NA\_System et l'objet détecté. La Figure V.14 illustre notre motivation pour définir cette fonction. De cette figure, on remarque que :

$$\alpha = d * \tan\left(\frac{41.41}{2}\right) \quad (\text{V.1})$$

Où :

d : distance séparant l'objet du NA\_System

En appliquant la règle de trois, le nombre de pixels  $x$  entre le centre du capteur ultrason et le centre de la caméra peut être déterminé selon l'Équation V.2 :

$$x = E\left(\frac{240 * 2.5}{\alpha}\right) \quad (\text{V.2})$$

En remplaçant  $\alpha$  par l'Équation V.1, le nombre de pixel  $x$  est défini :

$$x = E\left(\frac{1587.435}{d}\right) \quad (\text{V.3})$$

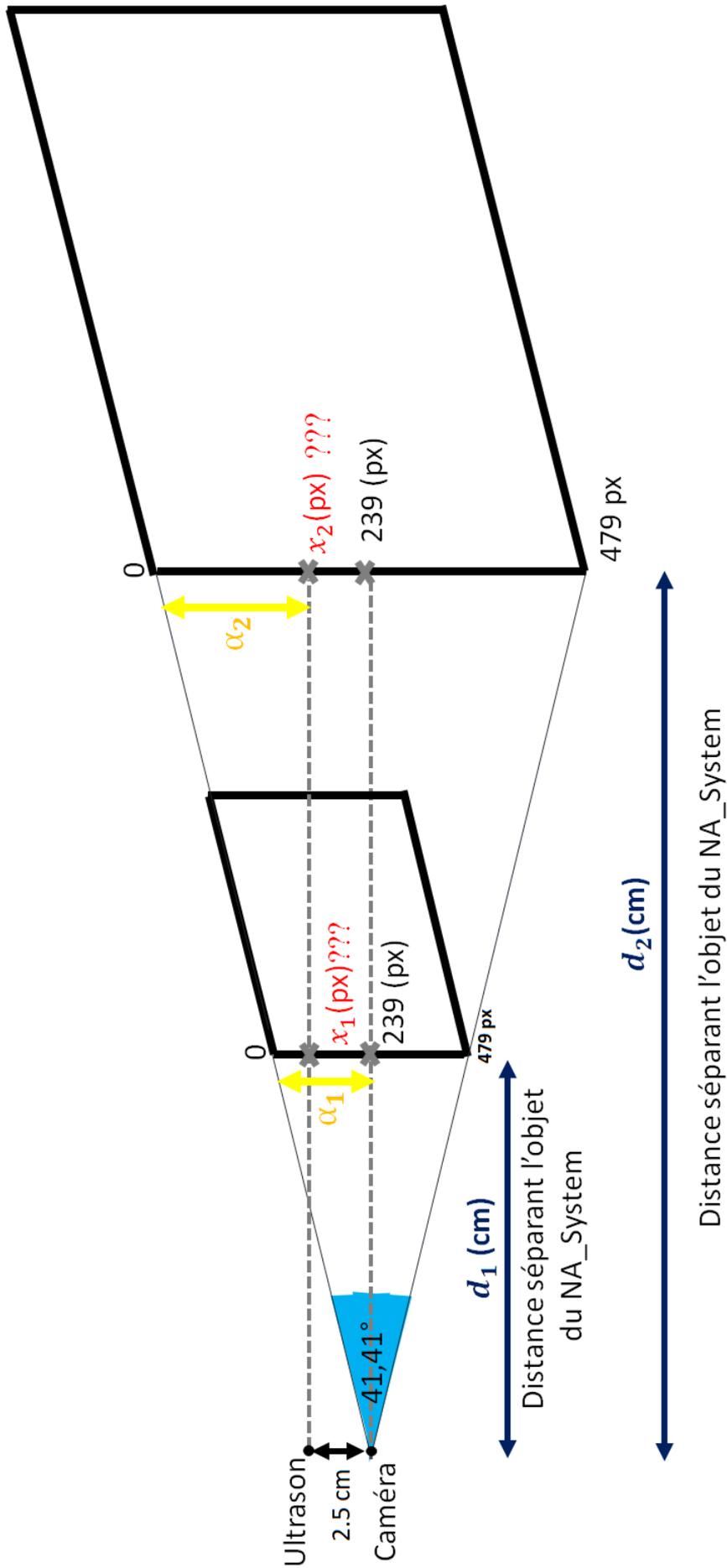


FIGURE V.14 – Détermination du nombre de pixel  $x$  entre le centre du capteur ultrason et le centre de la caméra en fonction de la distance séparant le système de l'objet

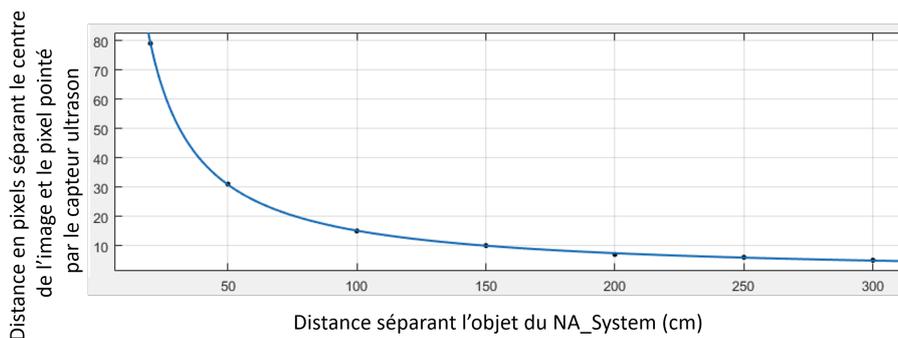


FIGURE V.15 – Détermination du pixel pointé par le capteur ultrason dans l'image par rapport au centre de l'image

### V.3.2.2 Recalage du signal de profil et l'image panoramique

Le principe de recalage du signal de profil et de l'image panoramique est illustré à la Figure V.16. Soit une scène balayée horizontalement est perçue par NA\_System. Ainsi, deux signaux de profil sont construits à partir des données ultrasonores et une vidéo est capturée par la caméra. Supposons que la vidéo se compose de 4 images (Cf. Figure V.16) et supposons que le balayage est effectué à une vitesse constante. Par conséquent, les images sont assemblées à la même distance entre leurs centres. Ensuite, une étape consistant à extrapoler les pixels manquants d'assemblage et à rogner l'image panoramique à une hauteur égale à 480 pixels est requise. En effet, la fonction définie par l'Équation V.3 permet la détermination du pixel vers lequel pointe le capteur ultrason dans l'image, dont la hauteur est égale à 480 pixels. Ensuite, nous prenons la première et la dernière mesure de distance du signal de profil et nous calculons à quels pixels de l'image elles correspondent. Enfin, le signal de profil est positionné entre ces deux pixels de l'image.

Il est à noter que le signal de profil obtenu par le capteur ultrason situé au-dessus de la caméra semble apparaître sous l'effet miroir sur l'image panoramique. En fait, cela est dû au fait que le nombre de pixels entre le centre de l'image et le pixel pointé par le capteur ultrason est inversement proportionnel à la distance entre le NA\_System et les objets présentés dans la scène.

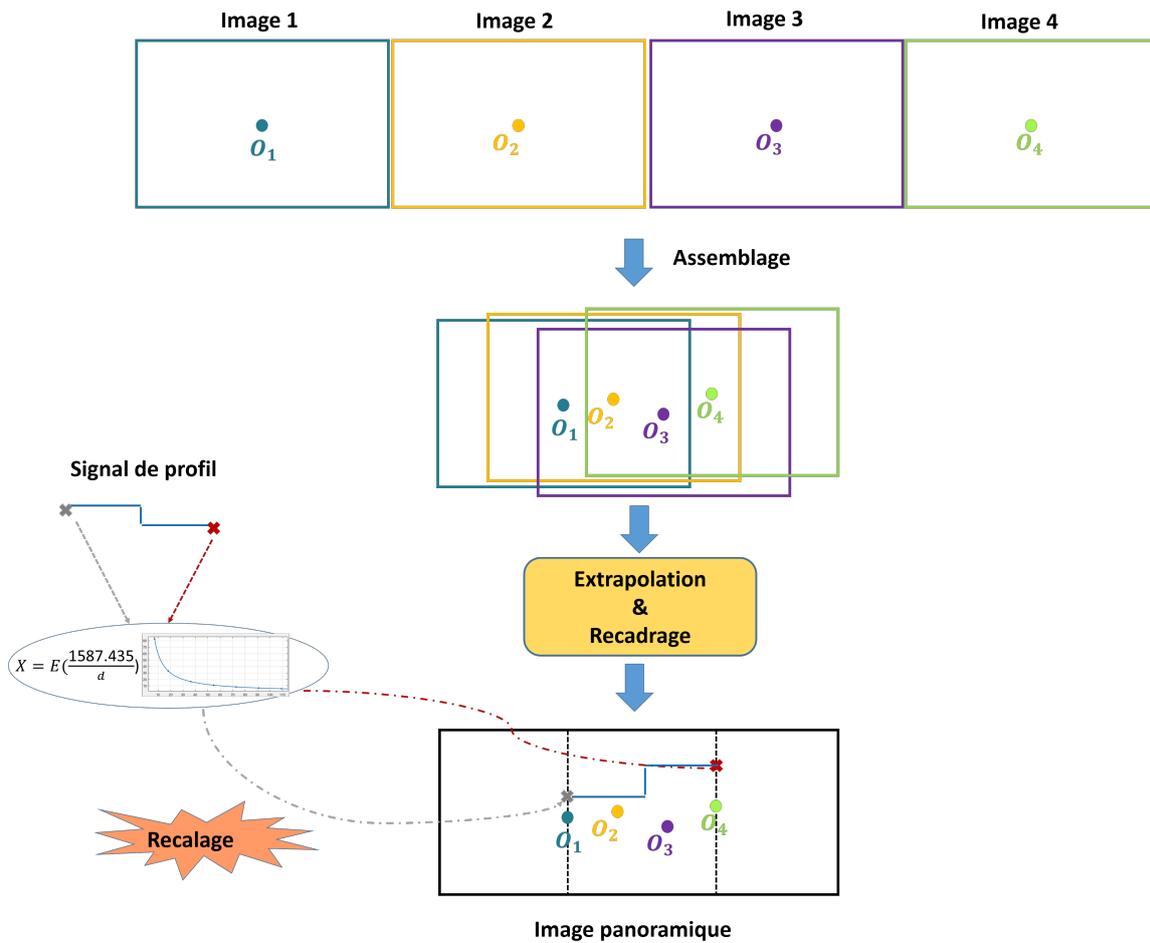


FIGURE V.16 – Principe de recalage le signal de profil avec l’image panoramique

### V.3.3 Résultats et discussions

Nos expérimentations consistent à évaluer notre modèle d’interprétation de scène, qui se base sur l’application de la stratégie d’interprétation de scène basée sur des données ultrasonores (proposée dans le Chapitre III) et l’application de la stratégie d’interprétation de scène basée sur un capteur de vision (proposée dans le Chapitre IV), et d’autre part, sur le recalage des différentes données ainsi la fusion de différentes connaissances.

Il est à noter que les différentes scènes représentées ci-dessous sont prises à une même durée d’acquisition et à la même vitesse de balayage.

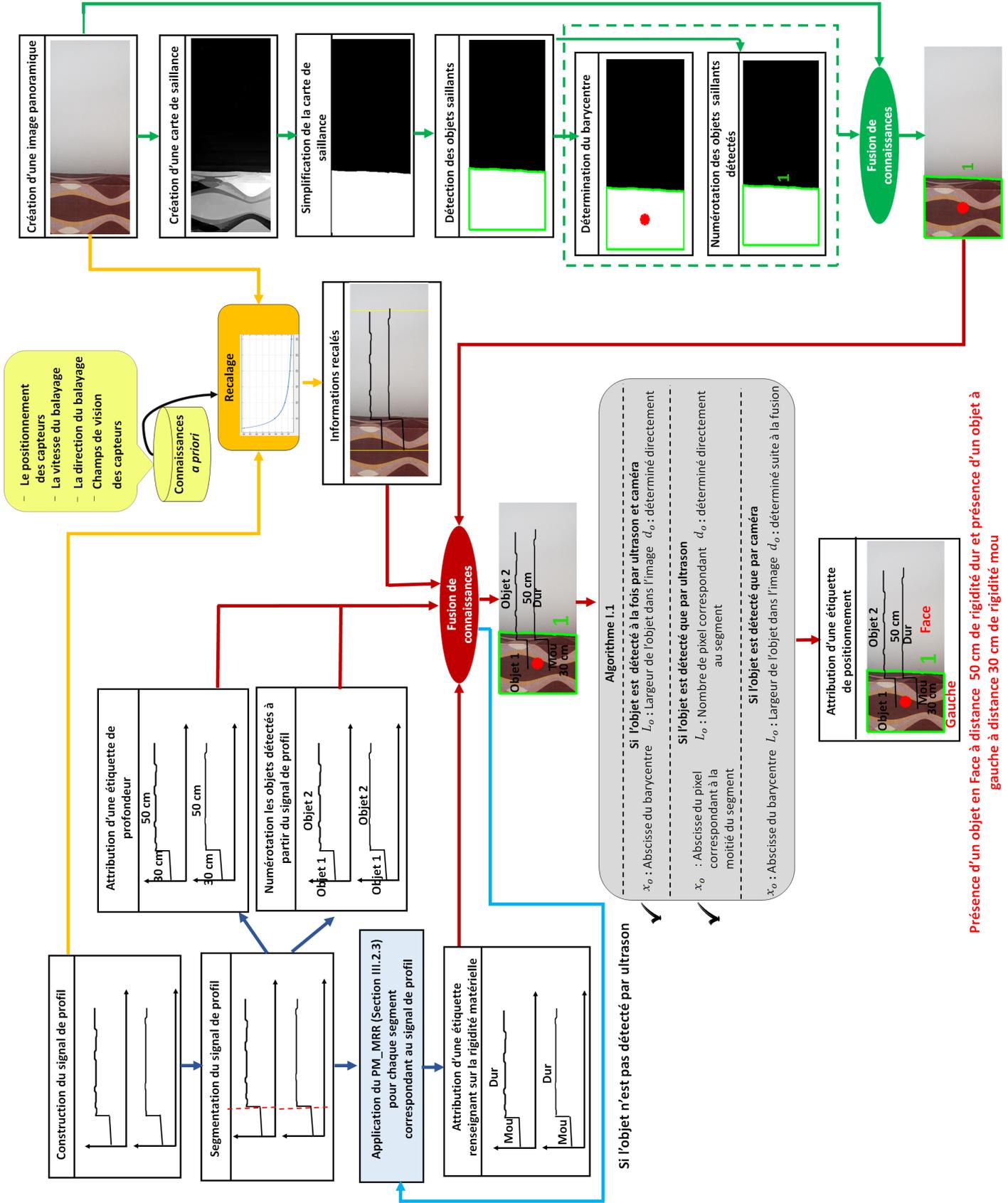
- Pour la première scène, un seul objet est détecté en appliquant la stratégie d’interprétation de scène basée sur une caméra. Cependant, en appliquant la stratégie de scène basée sur un capteur ultrason, deux objets sont détectés l’un à une distance de 30 cm de nature mou et l’autre à une distance 50 cm de nature dur. Après fusion de ces différentes connaissances, et en recalant l’image panoramique avec les signaux de profils, une étiquette de positionnement est attribuée à chaque objet en appliquant l’Algorithme

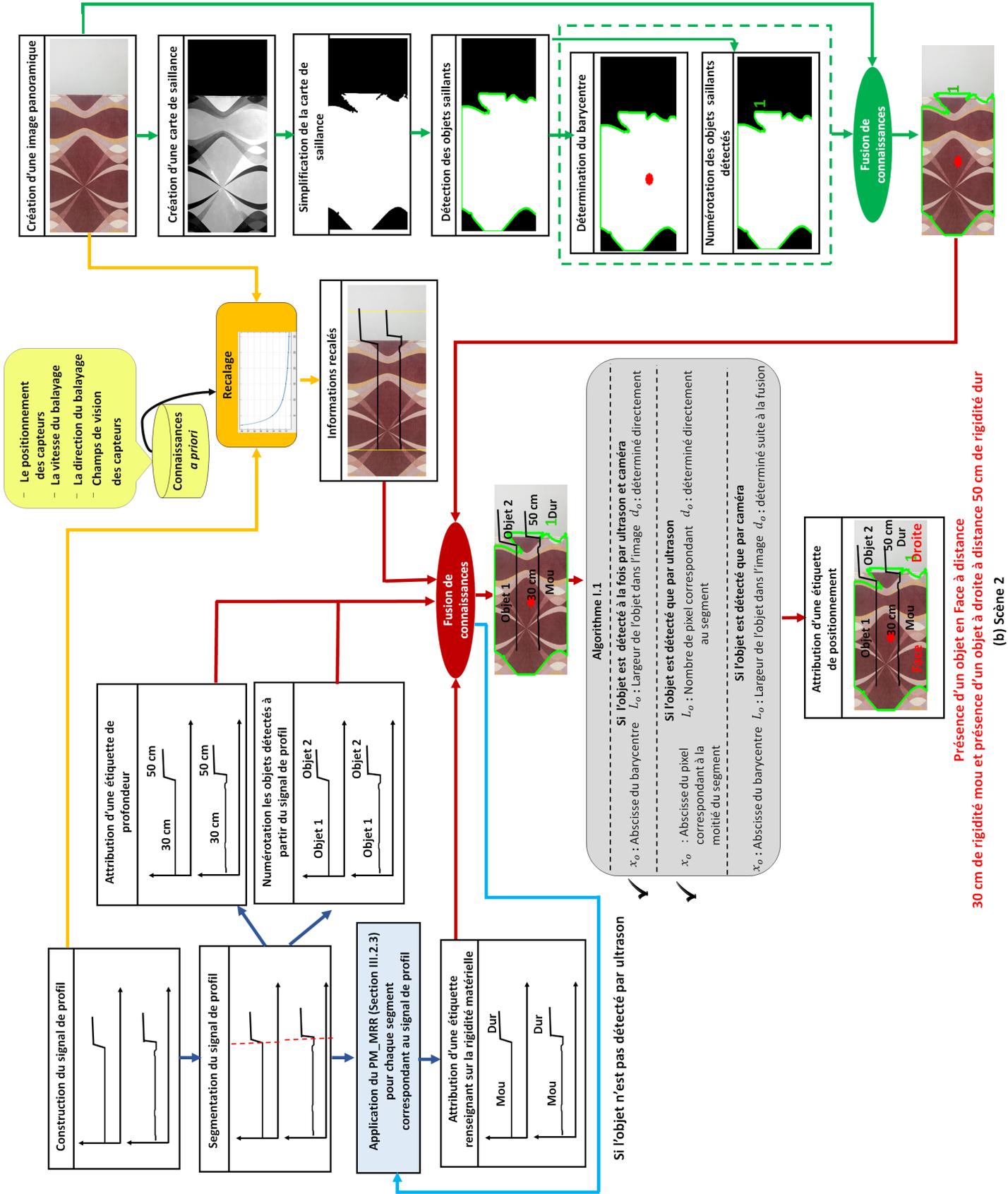
I.1. Comme l'objet à gauche est détecté à la fois par la caméra et le capteur ultrason, sa localisation est alors basée sur son barycentre (déterminé par la stratégie d'interprétation basée sur une caméra) et sur sa distance déterminée par le capteur ultrason. Cependant, l'objet en face n'est détecté que par le capteur ultrason. Ainsi, on a pris le pixel auquel correspondait la moitié du segment du signal de profil correspondant à cet objet comme abscisse de son barycentre afin de le localiser.

- L'interprétation de la deuxième scène est similaire à la première scène. En fait, elles présentent les mêmes objets aux mêmes distances, mais dans des positions différentes.
- La troisième scène présente un seul objet détecté à la fois par la caméra et par le capteur ultrason. Après avoir fusionner les différentes connaissances, on constate qu'il s'agit d'un objet en face de rigidité dur à une distance de 50 cm.
- Pour la quatrième scène, un seul objet est détecté en appliquant la stratégie d'interprétation de scène basée sur une caméra. Comme, il n'y a pas de pic au niveau du signal de profil, nous avons l'intention que ce soit un objet unique dans la scène. Cependant en recalant les signaux de profils avec l'image panoramique, nous constatons qu'une partie du signal de profil dépasse les limites de l'objet détecté par la stratégie d'interprétation basée sur la caméra. Par conséquent, nous constatons qu'il s'agit de deux objets situés à la même distance. Il est à noter que l'objet détecté par la caméra est localisé en "Face" en fonction de son abscisse de son barycentre, et que l'autre objet, également en "Face", est localisé en prenant le pixel auquel correspondait à la moitié du segment du signal de profil comme l'abscisse de son barycentre.
- Pour la cinquième scène, deux objets sont détectés en appliquant la stratégie d'interprétation de scène basée sur une caméra. Cependant, en appliquant la stratégie de scène basée sur un capteur ultrason, trois objets sont détectés : le premier objet est à 30 cm de nature mou, le deuxième objet à 50 cm de nature dur et le troisième objet à 20 cm de nature mou.

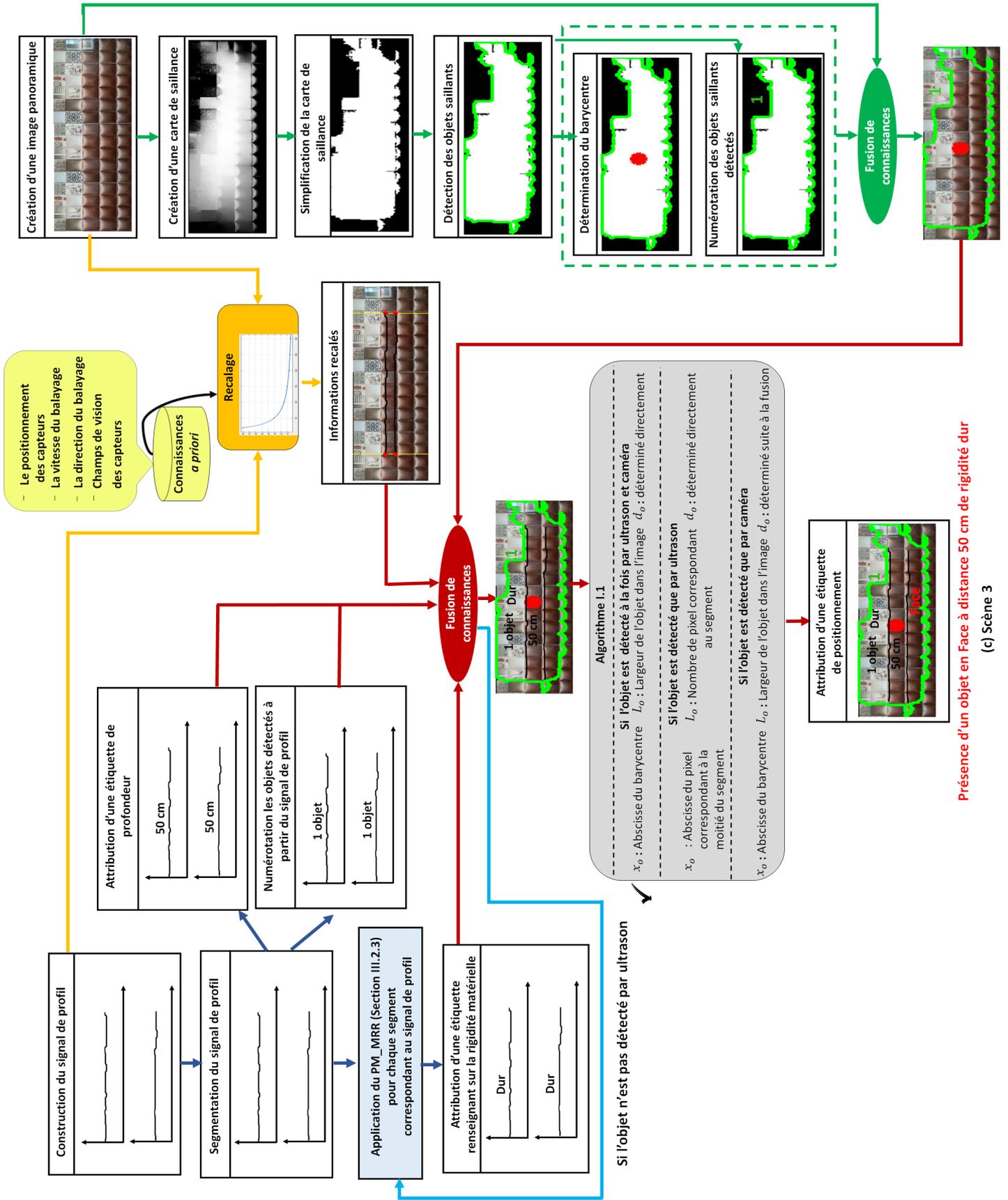
## V.4 Conclusion

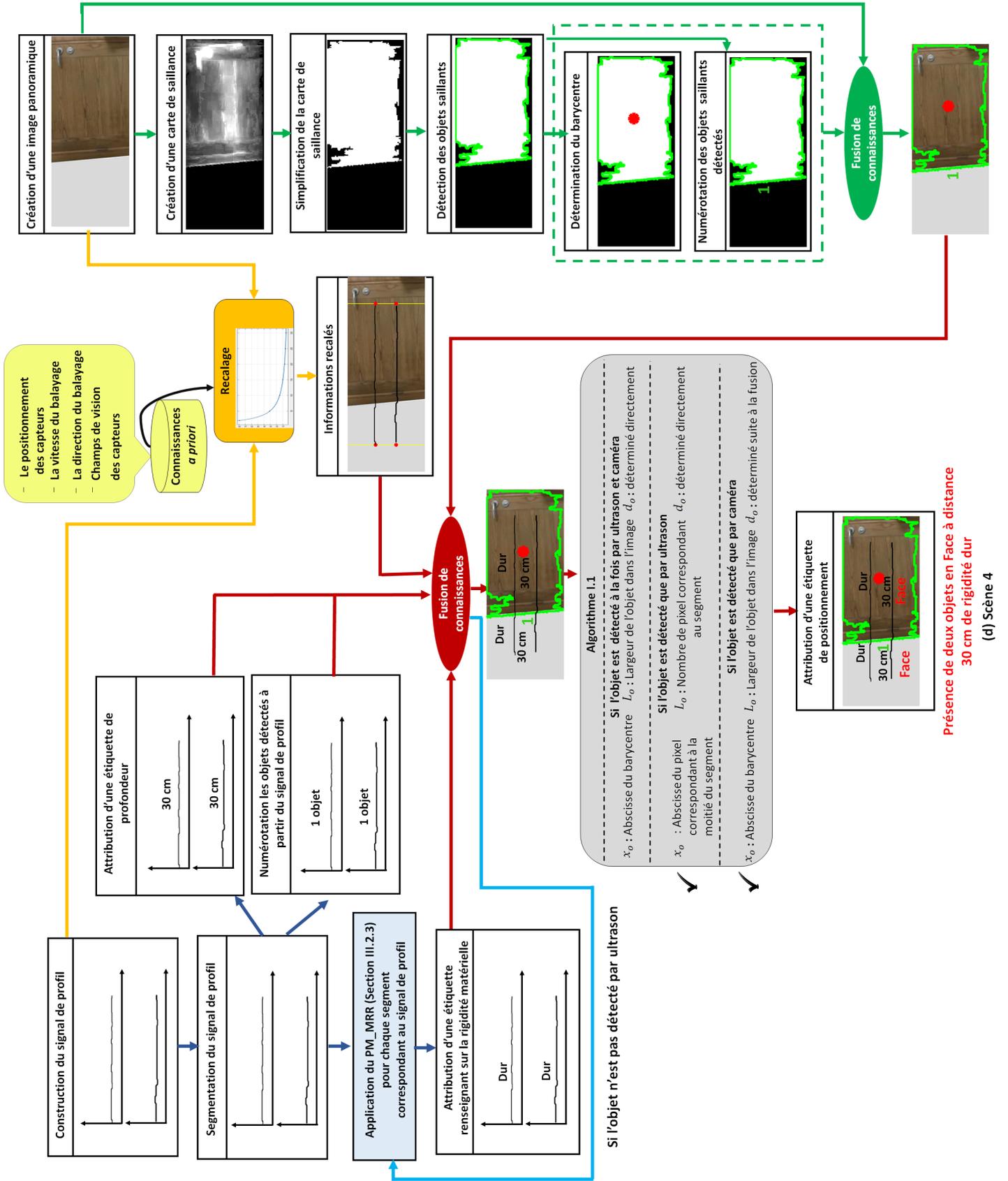
Dans ce chapitre, nous avons présenté un état de l'art sur les systèmes d'aides à la navigation pour les personnes à déficience visuelle. De plus, nous avons présenté l'historique et les composants électroniques du prototype d'acquisition "NA\_System" utilisé. Ensuite, nous avons prouvé que l'application de la stratégie d'interprétation de scène par un capteur ultrason (proposée dans le chapitre III) et l'application de la stratégie proposée pour l'interprétation





Présence d'un objet en Face à distance 30 cm de rigidité mou et présence d'un objet à droite à distance 50 cm de rigidité dur (b) Scène 2





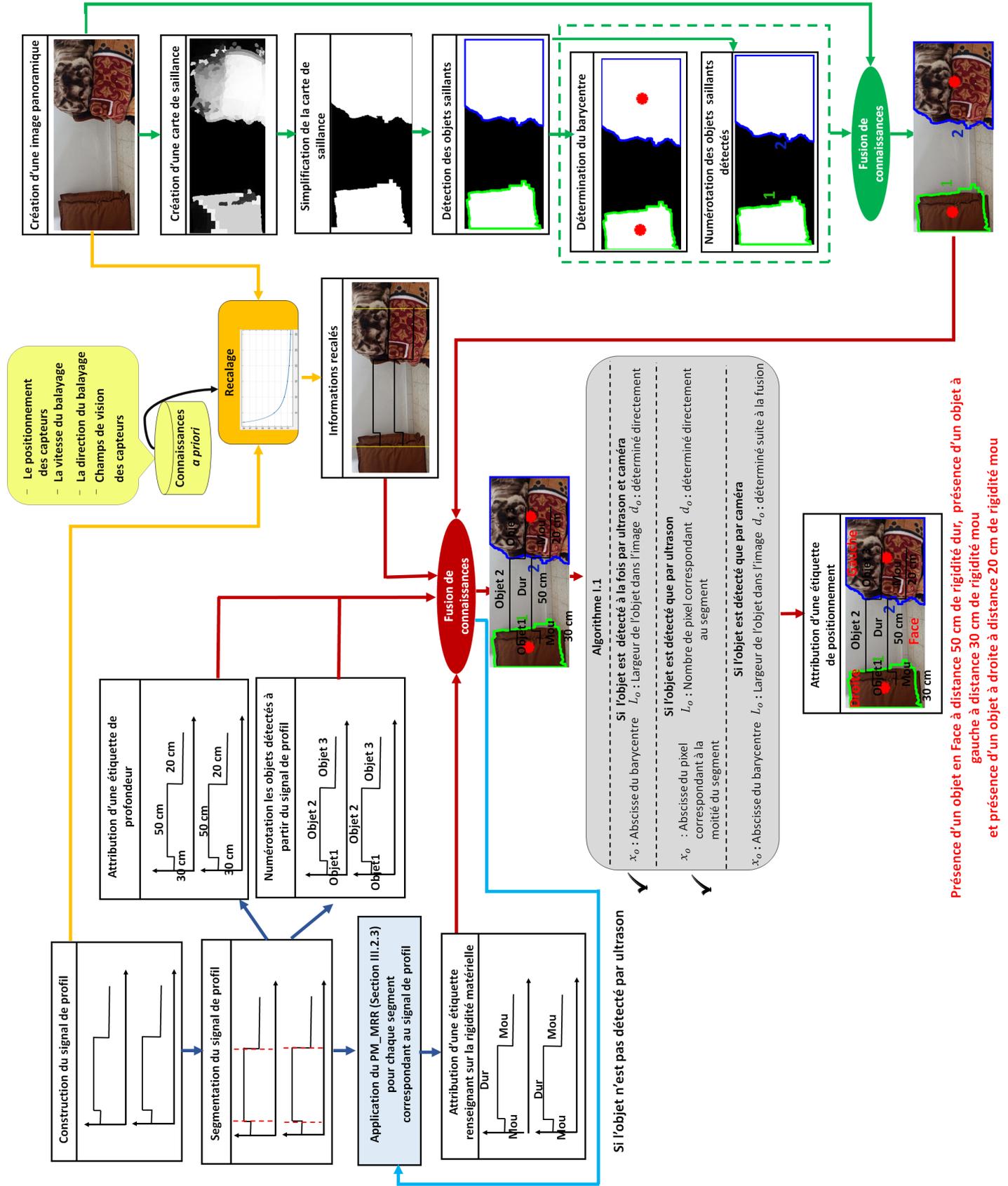


FIGURE V.17 – Application du modèle d'interprétation proposé sur un exemple de scénario

---

de scène par une caméra (proposée dans le chapitre IV), fusionnant ainsi leurs données hétérogènes, permet aux personnes à déficience visuelle de se déplacer dans leur environnement intérieur avec une simple modélisation de scène.

# Chapitre VI

## Conclusions et perspectives

Dans ce chapitre, nous synthétisons les travaux abordés dans ce mémoire, résumant les principaux résultats et contributions de nos travaux. Ensuite, nous présentons les perspectives ouvertes sur ce travail.

### VI.1 Synthèse des travaux présentés dans ce mémoire

Cette thèse s'articule autour de deux axes de recherche : l'un consiste en une étude et conception d'un système d'interprétation de scène d'environnement intérieur, observée par un système multi-capteurs réunissant un capteur ultrason et une caméra RVB, et l'autre inclut l'analyse et la modélisation d'informations imparfaites récupérées lors du processus d'acquisition.

#### VI.1.1 Motivations et problématique

À l'origine, l'utilisation des systèmes de navigation intérieure a été augmenté de façon exponentielle. Plusieurs systèmes ont été proposés pour suivre et localiser des objets équipant des environnements intérieurs. L'utilisation de plusieurs capteurs et processus de traitement tend à augmenter la complexité de ces systèmes, en particulier dans les modèles de scènes dont ils utilisent pour l'interprétation.

Afin de réduire le traitement des données provenant de sources multiples dans le système d'aide à la navigation, nous proposons un nouveau système d'interprétation de scène intérieure observée par deux types de capteurs. Le système d'interprétation proposé permet une interprétation de scène selon un modèle simpliste, partageant le champ intercepté par les capteurs en trois zones principales, à savoir face, gauche et droite. Bien qu'il soit simple, une aide à la navigation basée sur le système d'interprétation proposé est assurée. De plus, nous avons également développé certaines fonctionnalités du système d'aide à la navigation, notamment

fournir à l'utilisateur des informations sur la distance qui le sépare des objets détectés, la rigidité matérielle de ces objets, ainsi que leur positionnement dans la scène (objet à gauche, objet en face, objet à droite).

Un système d'aide à la navigation doit averti l'utilisateur en temps réel pour éviter les collisions éventuelles avec des objets nuisibles. Ainsi, le temps de traitement doit être réduit autant que possible pour avoir un système fonctionnant en temps réel. À cette fin, nous avons considéré deux contraintes pour réduire le temps de traitement. Tout d'abord, un seul capteur ultrason est utilisé. Deuxièmement, nous avons procédé à utiliser et à analyser des signaux de courte durée. Cela signifie que notre processus n'est basé que sur un petit nombre de valeurs de données acquises. Cependant, plus la taille de l'échantillon est petite, plus l'information est inexacte, ambiguë, incomplète et incertaine. Dans ce contexte, la théorie de possibilités présente un cadre approprié pour la modélisation et le traitement de telles informations.

## VI.1.2 Bilans et conclusions

### A. Interprétation de scène & théorie de possibilités

Dans le premier chapitre de ce manuscrit, le cadre applicatif, le cadre théorique et les problématiques concernées par nos objectifs de recherche dans la conception d'un système d'interprétation de scène perçue par un système multi-capteurs ont été présentés.

Dans la première partie de ce chapitre, quatre études ont été menées : une étude bibliographique sur les systèmes d'interprétation de scène utilisés par les systèmes d'aide à la navigation dans un environnement intérieur, une étude des techniques de localisations qu'ils adoptent, une étude de capteurs dont ils sont équipés, et une étude sur différents modèles de modélisation des environnements intérieurs. En fonction de l'objectif final, un modèle générique d'interprétation de scène basé sur la fusion des informations de différentes sources de connaissances a été proposé.

Les concepts de base liés à l'information et les différentes formes d'imperfection qui peuvent y être affectées ont été présentées dans la deuxième partie de ce chapitre. Étant donné que la théorie de possibilités a été adoptée comme le cadre théorique pour le développement du système d'interprétation proposé, différents outils de base de cette théorie, tels que la distribution de possibilités, la mesure de possibilité, la mesure de nécessité, les mesures d'incertitude, les mesures de similarité, les différents opérateurs de fusion d'informations et les critères de décision possibilistes ont été présentés.

## **B. Sélection d'attributs & Classification binaire hiérarchique**

Le deuxième chapitre a abordé la sélection d'attributs dans un cadre possibiliste afin de réduire la dimensionnalité de l'espace de représentation des données, réduisant ainsi la complexité de notre système. Face au problème de sélection d'attributs, trois approches basées sur le pouvoir discriminant des attributs ont été proposées. Les résultats de l'évaluation ont montré un taux de précision intéressant pour les différentes approches. Cette évaluation a été réalisée d'abord en utilisant une base de donnée synthétique, puis en utilisant des bases de données de références.

Dans la deuxième partie ce chapitre, deux classifieurs binaires hiérarchiques, l'un monocritère et l'autre multicritère ont été proposés. Ces deux classifieurs visent non seulement à simplifier le problème d'entrée massive pour la classification multi-classes, mais visent également à traiter l'incertitude liée aux attributs. Les résultats de l'évaluation ont montré un taux de reconnaissance intéressant pour les classifieurs proposés. Enfin, après avoir conçu les deux classifieurs binaires hiérarchiques, nous avons procédé à les comparer à d'autres classifieurs appliqués à différentes bases de données.

## **C. Détection et reconnaissance de la rigidité matérielle des objets présentés dans une scène en se basant sur des données ultrasonores**

Le troisième chapitre a été dédié au traitement du signal ultrasonore. La première partie de ce chapitre est consacrée à la présentation d'une étude bibliographique, qui implique des méthodes utilisées pour éviter les obstacles. Dans ce contexte, une approche de détection et d'évitement d'obstacles basée sur un signal ultrasonore a été proposée.

La deuxième partie de ce chapitre est consacrée à la reconnaissance de la rigidité matérielle des objets détectés utilisant l'écho ultrasonore réfléchi par ces objets. À cette fin, l'influence physique de la rigidité du matériau de l'obstacle sur l'onde ultrasonore a été étudiée. Deux approches de reconnaissance de la rigidité matérielle ont été proposées. La première est dite statistique car les classes sont représentées par des vecteurs d'attributs statistiques, alors que la deuxième approche est dite possibiliste car les classes sont représentées par des modèles possibilistes (distributions de possibilités).

Dans la dernière partie de ce chapitre, nous nous concentrons sur la distinction des objets durs les uns des autres. Ainsi, la reconnaissance de la rigidité matérielle est considérée comme un problème de classification multi-classes. De ce fait, les deux classifieurs hiérarchiques proposés dans le deuxième chapitre ont été appliqués. À notre connaissance, aucune base de données de

signaux ultrasonores n'est fournie pour identifier la rigidité matérielle des objets, nous avons donc créé notre propre base de données afin d'évaluer les différentes approches proposées.

#### **D. Simplification de scène en se basant sur un capteur de vision**

Les systèmes d'interprétation de scène d'un environnement intérieur fournissent généralement de nombreuses informations, qui ne sont parfois pas nécessaires pour une navigation autonome. Ainsi, le quatrième chapitre se consacre à proposer une stratégie de perception sélective permettant de simplifier autant que possible la scène tout en assurant une navigation autonome. La stratégie proposée repose sur deux étapes principales. La première consiste à mettre en évidence les objets saillants pour supprimer les objets d'arrière-plan et la seconde consiste à effectuer des opérations morphologiques pour simplifier l'interprétation de la scène.

#### **E. Interprétation de scène par la fusion et le recalage de données issues d'une caméra et de deux capteurs ultrasons**

Le cinquième chapitre présente le cadre d'application du système d'interprétation proposé, qui consiste à aider les personnes à déficience visuelle à naviguer dans leurs environnements intérieurs en toute sécurité. La première partie de ce chapitre présente une revue de la littérature des systèmes utilisés pour aider les personnes à déficience visuelle. Ensuite, une étude historique est menée sur le prototype "NA\_System", développé par l'équipe "Cybernetics team" du laboratoire "CEM\_Lab" de l'École nationale d'ingénieurs de Sfax (ENIS). Les résultats obtenus sont encourageants et montrent l'efficacité de la théorie de possibilités comme cadre de représentation de données acquises à partir de différents capteurs. La stratégie d'interprétation de scène proposée s'est avérée efficace pour recalibrer et fusionner les informations provenant de multiples sources de connaissances.

## **VI.2 Perspectives**

Nos travaux de recherches ont permis d'ouvrir un certain nombre de perspectives, qui peuvent être partagées en deux catégories :

### **VI.2.1 Perspectives à court terme**

- La mise en place d'une coopération entre les méthodes utilisées dans notre système d'interprétation et les méthodes de reconnaissances des objets (sol, porte, croisement de couloir, etc.) peut être envisagée. Une telle coopération ne peut qu'améliorer la performance de notre système d'interprétation de scène. En effet, l'utilisateur sera informé de la distance qui le sépare de l'objet, de sa rigidité matérielle, de son positionnement dans la scène et

également de quoi il s'agit.

- Le choix des valeurs de seuils (pour la sélection d'attributs) a été réalisé de manière empirique. Ce choix vaut encore la peine d'être étudié afin de développer une méthodologie où les seuils peuvent être automatiquement sélectionnés en fonction de l'application envisagée (en se basant sur les valeurs des pouvoirs discriminants des différents attributs, par exemple).
- Construction d'une base de donnée authentique de signaux ultrasonores et vidéos tout en considérant différentes conditions de prise de vues à différentes distances, afin d'évoluer vers un système d'interprétation de scène commercialisable.
- Dans certains cas, l'objet détecté se constitue d'un mélange de matériaux avec différentes rigidités matérielles (par exemple, métal et mousse). Ainsi, nous envisageons d'ajouter une classe "Autre" en plus de la classe "Dur" et la classe "Mou" pour caractériser la rigidité matérielle des objets.
- Lors du recalage de signaux ultrasonores et l'image panoramique, la vitesse de balayage est considérée constante. Cependant, ce n'est pas toujours le cas dans des situations réelles. Par conséquent, une fonction de recalage dépendant de la vitesse de balayage peut être proposée.
- Dans son état actuel, la stratégie d'interprétation de scène par une caméra monoculaire est convertie en un problème d'interprétation d'une image panoramique obtenue à l'aide d'un logiciel (ICE). L'extension envisagée consiste de permettre au système proposé d'interpréter le flux de données présenté dans la vidéo de la scène observée.

### VI.2.2 Perspectives à long terme

- Nos travaux de recherche sur la modélisation possibiliste des connaissances soulèvent des questions méthodologiques. En fait, la théorie de possibilités offre une certaine flexibilité pour la représentation des données, de sorte que son modèle est très adapté à la modélisation d'informations incomplètes. Leur flexibilité dans la représentation des données les rend plus appropriées à la modélisation d'informations incomplètes. De nombreuses questions peuvent se poser à ce niveau, à savoir : "Combien d'échantillons pouvons-nous fournir tout en conservant une bonne représentation de la classe ?", "Comment s'assurer de la qualité de ces échantillons", "Comment évaluons-nous la représentativité de la classe?", Ces questions soulèvent également d'autres questions importantes, telles que "Comment mesurer la qualité des données ?", "Comment évaluer la qualité du modèle?" et "Comment

s'assurer de la performance d'une méthode vis-à-vis les imperfections?".

- À l'aide d'un algorithme de reconnaissance faciale basé sur une caméra monoculaire RVB, le NA\_System peut être formé pour reconnaître les personnes que l'utilisateur rencontre souvent dans son environnement intérieur, telles que ses amis et les membres de sa famille.
- Notre système peut être équipé d'un module de reconnaissance vocale capable de reconnaître différentes commandes vocales émises par l'utilisateur. Ces commandes sont utiles pour sélectionner quelle fonctionnalité du système l'utilisateur désire.
- Actuellement, le NA\_System fournit une sortie en deux langues : français et anglais. Nous envisageons d'étendre cette fonctionnalité pour inclure d'autres langues afin de permettre aux utilisateurs de nationalités différentes d'utiliser ce système.
- L'une des contraintes que doit satisfaire un système d'aide à la navigation est la sûreté de son fonctionnement. De même, il semble très opportun et intéressant d'enrichir notre système avec des concepts spécifiques à l'analyse de la sûreté de son fonctionnement et de sa tolérance aux fautes.
- Le NA\_System est équipé par une carte Arduino pour configurer les capteurs ultrasons et collecter leurs données, puis les envoyer à la carte Raspberry, qui est responsable de la configuration d'autres capteurs, tels que la caméra, les servomoteurs, etc. Ainsi, nous envisageons de remplacer la carte Arduino et la carte Raspberry par une autre carte capable de gérer tous les capteurs en même temps pour réduire davantage la taille du notre système.
- Des mises à jour matérielles et logicielles à NA\_System peuvent réduire la consommation d'énergie.

# Publications

## Conférences

1. S. A. Bouhamed, M. Chakroun, I. K. Kallel and H. Derbel, "Haralick feature selection for material rigidity recognition using ultrasound echo," 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2018, pp. 1-6, doi : 10.1109/ATSIP.2018.8364483.
2. M. Chakroun, S. A. Bouhamed, I. K. Kallel, H. Derbel, and B. Solaiman, "Ultrasonic material recognition for safe navigation", the 34nd International Business Information Management Association Conference (IBIMA'20), Madrid, Spain, November 13-14, 2019.
3. M. Chakroun, A. Charfi, S. A. Bouhamed, I. K. Kallel, B. Solaiman and H. Derbel, "Binary hierarchical multiclass classifier for uncertain numerical features," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2020, pp. 1-6, doi : 10.1109/ATSIP49331.2020.9231804.

## Chapitre de livre

1. M. Chakroun, S. A. Bouhamed, I. K. Kallel, H. Derbel, and B. Solaiman, "Indoor scene simplification for safe navigation using saliency map for the benefit of visually impaired people", *Human Biometrics Biorobotics*, Springer, 2020.

## Article dans une revue

1. M. Chakroun, S. A. Bouhamed, I. K. Kallel, H. Derbel, and B. Solaiman, " Feature selection based on discriminative power under uncertainty for computer vision applications", *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 20(2) :83-93 ; 2021

## Revue en cours

1. M. Chakroun, S. A. Bouhamed, I. K. Kallel, H. Derbel, and B. Solaiman, "Possibilistic

---

modeling for material rigidity recognition using an ultrasonic signal", *Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*

# Bibliographie

- [1] G. J. Sonnenberg, *Radar and electronic navigation*, Elsevier, 2013.
- [2] B. Hofmann-Wellenhof, K. Legat and M. Wieser, *Navigation : Principles of positioning and guidance*, Springer Science & Business Media, 2003.
- [3] R. Prasad and M. Ruggieri, *Applied satellite navigation using GPS, GALILEO, and augmentation systems*, Artech House, 2005.
- [4] E. Kaplan and C. Hegarty, *Understanding GPS : principles and applications*, Artech house, 2005.
- [5] W. E. Sakpere and M. O. Adeyeye, "Can Near Field Communication Solve The Limitations in Mobile Indoor Navigation ?," In *RFID Technology Integration for Business Performance Improvement*, IGI Global, 52–79 (2015).
- [6] B. Alsahwa, *Représentation d'un environnement par un système multi-capteurs : fusion et interprétation de scène. Vision par ordinateur et reconnaissance de formes*. Télécom Bretagne : Université de Rennes 1, 2014.
- [7] I. K. Kallel, *Mécanismes de raisonnement possibiliste pour l'aide à la décision et l'interprétation de scènes*. Traitement du signal et d'image. Ecole Nationale des Ingénieurs de Sfax : Université de Sfax, 2019.
- [8] P. Sobrevilla and E. Montseny, "Fuzzy sets in computer vision : an overview," *Mathware and Soft Computing*, **10**(2/3), 71–83(2003).
- [9] M. Montemerlo, S. Thrun, D. Koller, *et al.*, "FastSLAM : A factored solution to the simultaneous localization and mapping problem," *Aaai/iaai*, **593598**,(2002).
- [10] C. Cadena, L. Carlone, H. Carrillo, *et al.* "Past, present, and future of simultaneous localization and mapping : Toward the robust-perception age," *IEEE Transactions on robotics*, **32**(6), 1309–1332(2016).
- [11] J. Aulinas, Y. R. Petillot, J. Salvi, *et al.*, "The slam problem : a survey.," *CCIA*, **184**(1), 363–371(2008).

- 
- [12] M. Zaffar, S. Ehsan, R. Stolkin, *et al.*, "Sensors, SLAM and Long-term Autonomy : A Review," In *NASA/ESA Conference on Adaptive Hardware and Systems(AHS)*, 285–290(2018).
- [13] A. Borges, "Cartographie de l'environnement et localisation robuste pour la navigation de robots mobiles," *Journal européen des systèmes automatisés*, **37**(10), 1299–1304(2003).
- [14] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM) : Part II," *IEEE robotics & automation magazine*, **13**(3), 108–117(2006).
- [15] G. Grisetti, C. Stachniss and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE transactions on Robotics*, **23**(1), 34–46(2007).
- [16] P. Hébert, S. Betgé-Brezetz and R. Chatila, "Probabilistic map learning : Necessity and difficulties," In *Reasoning with Uncertainty in Robotics*, Springer, Berlin, 307–321(1995).
- [17] K. R. Beevers and W. H. Huang, "An Embedded Implementation of SLAM with Sparse Sensing," In *IEEE International Conference on Robotics and Automation*, 2008.
- [18] C. M. Gifford, R. Webb, J. Bley, *et al.*, "Low-cost multi-robot exploration and mapping," In *IEEE International Conference on Technologies for Practical Robot Applications*, 74–79(2008).
- [19] M. Krekovic , I. Dokmanic and M. Vetterli, "EchoSLAM : Simultaneous localization and mapping with acoustic echoes," In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 11-15(2016).
- [20] T. J. Chong, X. J. Tang, C. H. Leng, *et.al*, "Sensor technologies and simultaneous localization and mapping (SLAM)," *Procedia Computer Science*, **76**, 174–179(2015).
- [21] B. Vincke, A. Elouardi and A. Lambert, "Design and evaluation of an embedded system based SLAM applications," In *IEEE/SICE International Symposium on System Integration*, 224–229(2010).
- [22] V. Bonato, R. Peron, D. F. Wolf, *et al.*, "An FPGA implementation for a kalman filter with application to mobile robotics," In *2007 International Symposium on Industrial Embedded Systems*, 148-155(2007).
- [23] J. Nikolic, J. Rehder, M. Burri, *et al.*, "A synchronized visual-inertial sensor system with FPGA preprocessing for accurate real-time SLAM," In *2014 IEEE international conference on robotics and automation (ICRA)*, 431–437(2014).
- [24] I. Ismail, "Survey of SLAM in Low-Resourced Hardware," *IJAIT (International Journal of Applied Information Technology)*, **3**(01), 1–9(2019).

- 
- [25] G. Duceux, *Towards Unsupervised And Incremental Semantic Mapping*. Robotics. ENSTA ParisTech, 2015.
- [26] Y. Gu, A. Lo and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications surveys & tutorials*, **11**(1), 13–32(2009).
- [27] K. Al Nuaimi and H. Kamel, "A survey of indoor positioning systems and algorithms," In *2011 International Conference on Innovations in Information Technology*, 185–190(2011).
- [28] I. Amundson and X. D. Koutsoukos, "A survey on localization for mobile wireless sensor networks," In *Mobile entity localization and tracking in GPS-less environments*, 235–254(2009).
- [29] L. Brás, N. B. Carvalho, P. Pinho, *et al.*, "A review of antennas for indoor positioning systems," *International Journal of Antennas and Propagation*, **2012**,(2012).
- [30] D. Zhang, F. Xia, Z. Yang, *et al.*, "Localization Technologies for Indoor Human Tracking," In *2010 5th international conference on future information technology*, 1–6(2010).
- [31] F. Subhan, H. Hasbullah, A. Rozyyev, *et al.*, "Indoor positioning in Bluetooth networks using fingerprinting and lateration approach," In *2011 International Conference on Information Science and Applications*, 1–9(2011).
- [32] H. Liu, H. Darabi, P. Banerjee, *et al.*, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **37**(6), 1067–1080(2007).
- [33] J. Xiao, Z. Liu, Y. Yang, *et al.*, "Comparison and analysis of indoor wireless positioning techniques," In *2011 International conference on computer science and service system (CSSS)*, 293–296(2011).
- [34] D. Zhang, F. Xia, Z. Yang, *et al.*, "Localization technologies for indoor human tracking," In *5th International Conference on Future Information Technology*, 1–6(2010).
- [35] R. Mautz, *Indoor positioning technologies*, 2012.
- [36] Z. Tidjani, A. Mekhmoukh, G. Zidani, *et al.*, "Développement d'algorithmes d'évitement d'obstacles statiques et dynamiques," 2012.
- [37] F. Chenavier, I. Talbi, I Lecoer, *et al.*, "Estimation de la Position d'un Robot par Odométrie et Vision Monoculaire," *Traitement du Signal*, **12**(3), 255–267(1995).
- [38] G. Frappier, "Système inertiels de navigation pour robots mobiles," *Paris : Séminaire Les robots mobiles, EC2*, Paris, 1990.

- 
- [39] I. Maalej, M. Ouali and N. Derbel, "Modular ultrasonic sensor platform for mobile robot," In *2012 First International Conference on Renewable Energies and Vehicular Technology*, 477-483(2012).
- [40] E. Aitenbichler and M. Muhlhauser, "An IR local positioning system for smart items and devices," In *23rd IEEE International Conference on Distributed Computing Systems Workshops*, 334-339(2003).
- [41] L. Yenilmez and H. Temeltas, "Enabling an autonomous mobile robot to determine its actual position and orientation by using an ultrasonic ring," In *4th International Conference on Automation, Robotics and Vision*, 1125-1129(1996).
- [42] Y. Fukuju, M. Minami, H. Morikawa, *et al.*, "DOLPHIN : An autonomous indoor positioning system in ubiquitous computing environment," In *Proceedings IEEE Workshop on Software Technologies for Future Embedded Systems*, 53-56(2003).
- [43] M. Minami, Y. Fukuju, K. Hirasawa, *et al.*, "DOLPHIN : A practical approach for implementing a fully distributed indoor ultrasonic positioning system," In *International Conference on Ubiquitous Computing*, 347-365 (2004).
- [44] N. B. Priyantha, "The cricket indoor location system Doctoral dissertation," Electrical Engineering and Computer Science. Massachusetts Institute of Technology, University of Moratuwa, 2005.
- [45] O. J. Woodman and R. K. Harle, "Concurrent scheduling in the Active Bat location system," In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 431-437(2010).
- [46] A. Lindo, E. García, J. Ureña, *et al.*, "Multiband waveform design for an ultrasonic indoor positioning system," *IEEE Sensors Journal*, **15**(12), 7190–7199(2015).
- [47] A. De Angelis, A. Moschitta, P. Carbone, *et al.*, "Design and characterization of a portable ultrasonic indoor 3-D positioning system," *IEEE Transactions on Instrumentation and Measurement*, **64**(10), 2616–2625(2015).
- [48] U. Yayan, H. Yucel and A. Yazici, "A low cost ultrasonic based positioning system for the indoor navigation of mobile robots," *Journal of Intelligent & Robotic Systems*, **78**(3), 541–552(2015).
- [49] S. Yang and S. Scherer, "Monocular Object and Plane SLAM in Structured Environments," In *IEEE Robotics and Automation Letters*, **4**(4), 3145–3152(2019).
- [50] W. Huang, Y. Li and F. Hu, "Real-Time 6-DOF Monocular Visual SLAM based on ORB-SLAM2," In *2019 Chinese Control And Decision Conference (CCDC)*, 2929-2932(2019).

- 
- [51] H. Kuang, K. Zhang, R. Li, *et al.*, "Monocular SLAM Algorithm Based on Improved Depth Map Estimation and Keyframe Selection," In *10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 350-353(2018).
- [52] P. Araujo, R. Miranda, D. Carmo, *et al.*, "Air-SSLAM : A Visual Stereo Indoor SLAM for Aerial Quadrotors," *IEEE Geoscience and Remote Sensing Letters*, **14**(9), 1643–1647(2017).
- [53] D. Pohl, S. Dorodnicov and M. Achtelik, "Depth Map Improvements for Stereo-based Depth Cameras on Drones," In *2019 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 341–348(2019).
- [54] Y. Xiao, X. Lei and S. Liao, "Research on UAV Multi-Obstacle Detection Algorithm based on Stereo Vision," In *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 1241-1245(2019).
- [55] <https://sites.google.com/site/3dourelief/vision-3d/stereoscopie>
- [56] V. Lemonde, *Stéréovision Embarquée sur Véhicule : de l'Auto-Calibrage à la Détection d'Obstacles*. INSA-LAAS-CNRS : Toulouse, France, 2005.
- [57] L. Paya, A. Gil and O. Reinoso, "A State-of-the-Art Review on Mapping and Localization of Mobile Robots Using Omnidirectional Vision Sensors," *Journal of Sensors*, **2017**, 1-20(2017).
- [58] Y. W. Choi, J. W. Choi, S. G. Im, *et al.*, "Collision avoidance using omni vision slam based on fisheye image," *Journal of Institute of Control, Robotics and Systems*, **22**(3), 210–216(2016).
- [59] <https://www.abcclim.net/la-thermographie.html>
- [60] H. Rebecq, T. Horstschäfer, G. Gallego, *et al.*, "Evo : A geometric approach to event-based 6-dof parallel tracking and mapping in real time," *IEEE Robotics and Automation Letters*, **2**(2), 593–600(2016).
- [61] E. Mueggler, G. Gallego, H. Rebecq, *et al.*, "Continuous-time visual-inertial odometry with event cameras," *IEEE Transactions on Robotics*, **34**(6), 1425–1440(2018).
- [62] S. Choi, Q. Y. Zhou and V. Koltun, "Robust reconstruction of indoor scenes," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5556–5565(2015).
- [63] V. Lehtola, H. Kaartinen, A. Nüchter, *et al.*, "Comparison of the selected state-of-the-art 3D indoor scanning and point cloud generation methods," *Remote sensing*, **9**(8), 796(2017).

- 
- [64] G. T. Michailidis and R. Pajarola, "Bayesian graph-cut optimization for wall surfaces reconstruction in indoor environments," *The Visual Computer*, **33**(10), 1347–1355(2017).
- [65] R. Wang, L. Xie and D. Chen, "Modeling indoor spaces using decomposition and reconstruction of structural elements," *Photogrammetric Engineering & Remote Sensing*, **83**(12), 827–841(2017).
- [66] Y. Pang, C. Zhang, L. Zhou, *et al.*, "Extracting indoor space information in complex building environments," *ISPRS International Journal of Geo-Information*, **7**(8), 321(2018).
- [67] S. Thrun, "Learning metric-topological maps for indoor mobile robot navigation," *Artificial Intelligence*, **99**(1), 21–71(1998).
- [68] S. Thrun, "Robotic mapping : A survey," In *Exploring Artificial Intelligence in the New Millenium*, 1–35(2002).
- [69] <https://david-filliat.developpez.com/tutoriel/robotique-mobile/?page=navigation-utilisant-une-carte>
- [70] M. Bosse, P. Newman, J. Leonard, *et al.*, "An atlas framework for scalable mapping.," In *2003 IEEE International Conference on Robotics and Automation*, 1899-1906(2003).
- [71] T. Bailey, *Mobile robot localisation and mapping in extensive outdoor environments*, PhD thesis. Australian Center for Field Robotics : University of Sydney, Sydney, Australie, 2002.
- [72] J. Guivant, E. Nebot, J. Nieto, *et al.*, "Navigation and mapping in large unstructured environments," *The International Journal of Robotics Research*, **23**(4-5), 449–472(2004).
- [73] C. Zins, "Conceptual approaches for defining data, information, and knowledge," *Journal of the American society for information science and technology*, **58**(4), 479–493(2007).
- [74] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy sets and systems*, **1**(1), 3–28(1978).
- [75] D. Dubois and H. Prade, "Possibility theory : qualitative and quantitative aspects," In *Quantified representation of uncertainty and imprecision*, 169-226(1988).
- [76] B. Solaiman and É. Bossé, *Possibility Theory for the Design of Information Fusion Systems*, Springer International Publishing, 2019.
- [77] D. Dubois and H. Prade, "On several representations of an uncertain body of evidence," In *Fuzzy Information and Decision Processes*, 167-181(1982).
- [78] B. B. Devi and V. V. S. Sarma, "Estimation of fuzzy memberships from histograms," *Information sciences*, **35**(1), 43–59(1985).

- 
- [79] G. Jumarie, "A general approach to approximate reasoning via probability-possibility conversion," *Proceedings of IEEE Systems Man and Cybernetics Conference-SMC*, **4**, 656–661(1993).
- [80] D. Dubois and H. Prade, "On possibility/probability transformations," In *Fuzzy Logic*, 103-112(1993).
- [81] D. Dubois and H. Pradei, "La fusion d'informations imprécises," *Traitement du signal*, **11**(6), 447–458(1994).
- [82] G. J. Klir and J. F. Geer, "Information-preserving probability-possibility transformations : recent developments," In *Fuzzy Logic*, 417-428(1993).
- [83] M. H. Masson and T. Denoeux, "Inferring a possibility distribution from empirical data," *Fuzzy Sets and Systems*, **157**(3), 319–340(2006).
- [84] M. S. Bouguelid, *Contribution à l'application de la reconnaissance des formes et la théorie des possibilités au diagnostic adaptatif et prédictif des systèmes dynamiques*, Thèse de Doctorat. Université De Reims Champagne-Ardenne, 2007.
- [85] D. Dubois and H. Prade, "Panorama des nouvelles méthodes de traitement de l'incertitude et de l'imprécision," *Institut de Recherche en Informatique de Toulouse (IRIT)*, CNRS et Université P. Sabatier, France, 2010.
- [86] D. Dubois and H. Prade, "Teoria Vozmojnosti (Traduction Russe de "Théorie des Possibilités")," *Radio i Sviiaz*, Masson, Paris, 1990.
- [87] R. R. Yager, "On the specificity of a possibility distribution," *Fuzzy sets and systems*, **50**(3), 279–292(1992).
- [88] M. Higashi and G. J. Klir, "On the notion of distance representing information closeness : Possibility and probability distributions," *INTERNATIONAL JOURNAL OF GENERAL SYSTEM*, **9**(2), 103–115(1983).
- [89] M. Higashi and G. J. Klir, "Measures of uncertainty and information based on possibility distributions," *International journal of general systems*, **9**(1), 43–58(1982).
- [90] I. Jenhani, N. B. Amor, Z. Elouedi, *et al.*, "Information Affinity : a new similarity measure for possibilistic uncertain information," In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, 840-852(2007).
- [91] A. Charfi, S. A. Bouhamed, E. Bossé, *et al.*, "Possibilistic Similarity Measures for Data Science and Machine Learning Applications," *IEEE Access*, **8**, 49198–49211(2020).

- 
- [92] R. R. Yager, "Entropy and specificity in a mathematical theory of evidence," In *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer, Berlin, Heidelberg, 291–310(2008).
- [93] A. Tversky, "Features of similarity," *Psychological review*, **84**(4), 327(1977).
- [94] S. Kikuchi and V. Perincherry, "Handling Uncertainty in Large Scale Systems with Certainty and Integrity," In *MIT Engineering Systems Symposium*, 2004.
- [95] F. Portet and A. Gatt, "Towards a possibility-theoretic approach to uncertainty in medical data interpretation for text generation," In *International Workshop on Knowledge Representation for Health Care*, Springer, Berlin, Heidelberg, 155–168(2009).
- [96] F. Sha and F. Pereira, "Shallow parsing with conditional random fields," In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 213-220(2003).
- [97] S. A. Bouhamed, I. K. Kallel, D. S. Masmoudi *et al.*, "Feature selection in possibilistic modeling," *Pattern Recognition*, **48**(11), 3627–3640(2015).
- [98] <https://archive.ics.uci.edu/ml/index.php>
- [99] L. C. Chang, M. Z. M. Amin, S. N. Yang, *et al.*, "Building ANN-Based Regional Multi-Step-Ahead Flood Inundation Forecast Models," *Water*, **10**(9), 1283(2018).
- [100] H. Alaiz-Moreton, J. Aveleira-Mata, J. Ondicol-garcia, *et al.*, "Multiclass Classification Procedure for Detecting Attacks on MQTT-IoT Protocol," *Complexity*, **2019**,(2019).
- [101] S. Itani, F. Lecron and P. Fortemps, "A multi-level classification framework for multi-site medical data : Application to the ADHD-200 collection," *Expert Systems with Applications*, **91**, 36–45(2018).
- [102] Q. Zhou, "Multi-layer affective computing model based on emotional psychology," *Electronic Commerce Research*, **18**(1), 109–124(2018).
- [103] S. Begum and R. S. Aygun, "Greedy hierarchical binary classifiers for multi-class classification of biological data," *Network Modeling Analysis in Health Informatics and Bioinformatics*, **3**(1), 1–15(2014).
- [104] A. K. Mourya, H. Kaur and M. Uddin, "A Novel Approach to Heterogeneous Multi-class SVM Classification," In *New Paradigm in Decision Science and Management*, Springer, Singapore, 39–47(2020).
- [105] V. N. Vladimir, "An overview of statistical learning theory," *IEEE transactions on neural networks*, **10**(5), 988–999(1999).

- 
- [106] J. R. Quinlan, *C4. 5 : programs for machine learning*, Elsevier, 2014.
- [107] I. Guyon, S. Gunn, M. Nikravesh, *et al.*, *Feature extraction : foundations and applications*, Springer, **207**, 2008.
- [108] M. M. Saritas and A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *International Journal of Intelligent Systems and Applications in Engineering*, **7**(2), 88–91(2019).'
- [109] S. Chen, G. I. Webb, L. Liu, *et al.*, "A novel selective naïve Bayes algorithm," *Knowledge-Based Systems*, **192**, 105361(2020).
- [110] J. Hu, H. Peng, J. Wang, *et al.*, "KNN-P : A KNN classifier optimized by P systems," *Theoretical Computer Science*, **817**, 55–65(2020).
- [111] T. Lan, H. Hu, C. Jiang, *et al.*, "A Comparative Study of Decision Tree, Random Forest, and Convolutional Neural Network for Spread-F Identification," *Advances in Space Research*, **65**(8), 2052–2061(2020).
- [112] Z. Hosenie, R. J. Lyon, B. W. Stappers, *et al.*, "Comparing Multiclass, Binary, and Hierarchical Machine Learning Classification schemes for variable stars," *Monthly Notices of the Royal Astronomical Society*, **488**(4), 4858–4872(2019).
- [113] G. Douzas, F. Bacao and F. Last, "Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE," *Information Sciences*, **465**, 1–20(2018).
- [114] S. Rajan and J. Ghosh, "An Empirical Comparison of Hierarchical vs. Two-Level Approaches to Multiclass Problems," In *International Workshop on Multiple Classifier Systems*, Springer, Berlin, Heidelberg, 283–292(2004).
- [115] N. Matloff, *Statistical regression and classification : From linear models to machine learning*, CRC Press, 2017.
- [116] J. M. Puche, J. M. Benítez, J. Castro, *et al.*, "Fuzzy Pairwise Multiclass Support Vector Machines," In *Mexican International Conference on Artificial Intelligence*, Springer, Berlin, Heidelberg, 562–571(2006).
- [117] J. Bi and T. Zhang, "Support vector classification with input data uncertainty," *Advances in neural information processing systems*, **17**(1), 161–168(2005).
- [118] C. C. Aggarwal and S. Y. Philip, "A survey of Uncertain Data Algorithms and Applications," *IEEE Transactions on knowledge and data engineering*, **21**(5), 609–623(2008).

- 
- [119] K. Soundararajan, S. S. Kumar and C. Anusuya, "Classification of Uncertain Data Using Selection Algorithm," *International Journal of Modern Engineering Research (IJMER)*, **2**, 1066–1072(2012).
- [120] M. Li, H. Xu and Y. Deng, "Evidential Decision Tree Based on Belief Entropy," *Entropy*, **21**(9), 897(2019).
- [121] J. Xue, C. Wu, Z. Chen, *et al.*, "Modeling human-like decision-making for inbound smart ships based on fuzzy decision trees," *Expert Systems with Applications*, **115**, 172–188(2019).
- [122] M. Javid and J. Hamidzadeh, "An active multi class classification using privileged information and belief function," *International Journal of Machine Learning and Cybernetics*, **11**(3), 511–524(2020).
- [123] R. Agrawal and B. Ram, "A Modified K-Nearest Neighbor Algorithm to Handle Uncertain Data," In *5th International Conference on IT Convergence and Security (ICITCS)*, 1–4(2015).
- [124] K. Baati, T. M. Hamdani, A. M. Alimi, *et al.*, "A new classifier for categorical data based on a possibilistic estimation and a novel generalized minimum-based algorithm," *Journal of Intelligent & Fuzzy Systems*, **33**(3), 1723–1731(2017).
- [125] K. Baati, T. M. Hamdani, A. M. Alimi, *et al.*, "A new possibilistic classifier for mixed categorical and numerical data based on a bi-module possibilistic estimation and the generalized minimum-based algorithm," *Journal of Intelligent & Fuzzy Systems*, **36**(4), 3513–3523(2019).
- [126] N. B. Amor, S. Benferhat, C. Cnrs, *et al.*, "Qualitative classification and evaluation in possibilistic decision trees," In *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No. 04CH37542)*, **2**, 653–657(2004).
- [127] I. Jenhani, S. Benferhat and Z. Elouedi, "On the Use of Clustering in Possibilistic Decision Tree Induction," In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Springer, Berlin, Heidelberg, 505–517(2009).
- [128] <http://www.cs.waikato.ac.nz/ml/weka/>
- [129] Y. T. Win, N. Afzulpurka and C. Punyasai, "Shape detection of object behind thin medium using ultrasonic sensors," *International Journal of Advanced Robotic Systems*, **15**(2), 1–13(2018).
- [130] K. Ohtani and M. Baba, "Ultrasonic Robot Eye for Shape Recognition Employing a Genetic Algorithm," In *Proc. of IMEKO*, pp. 2021–2025(2003).

- 
- [131] G. Benet, M. Martínez, F. Blanes, *et al.*, "Differentiating walls from corners using the amplitude of ultrasonic echoes," *Robotics and Autonomous Systems*, **50**(1), 13–25(2005).
- [132] M. Martínez, G. Benet, F. Blanes, *et al.*, "Using the Amplitude of Ultrasonic Echoes to Classify Detected Objects in a Scene," In *Proceedings of the 11th International Conference on Advanced Robotics (ICAR'03)*, 1136-1142(2003).
- [133] R. Visvanathan, S. M. Mamduh, K. Kamarudin, *et al.*, "Multi-channel ultrasonic sensing system for wall features extraction," In *2014 2nd International Conference on Electronic Design (ICED)*, 351-355(2014).
- [134] W. F. Sahnoun, *Contribution à la description d'un obstacle détecté par un capteur ultrason*, Mastère de recherche I2M, ISIMS : Université de Sfax, 2013.
- [135] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on pure and applied mathematics*, **41**(7), 909–996(1988).
- [136] L. Mažeika, R. Šlitteris and A. Vladišauskas, "Measurement of velocity and attenuation for ultrasonic longitudinal waves in the polyethylene samples," *Ultragarsas" Ultrasound"*, **65**(4), 12–15(2010).
- [137] J. Bercoff, M. Tanter, M. Muller, *et al.*, "The role of viscosity in the impulse diffraction field of elastic waves induced by the acoustic radiation force," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, **51**(11), 1523–1536(2004).
- [138] NDT Education Resource Center, Disponible sur : <https://www.nde-ed.org/EducationResources/CommunityCollege/Ultrasonics/Physics/attenuation.htm>
- [139] Y. M. Zhu, F. Peyrin and R. Goutte, "The wigner-ville transform : description of a new tool for signal and image processing," In *Annals of Telecommunications*, **42**, 105–118(1987).
- [140] N. Zayed and H. A. Elnemr, "Statistical Analysis of Haralick Texture Features to Discriminate Lung Abnormalities," *International journal of biomedical imaging*, **2015**,(2015).
- [141] A. Chaddad, C. Tanougast and A. Dandache, "Extraction of haralick features from segmented texture multispectral bio-images for detection of colon cancer cells," In *2011 First International Conference on Informatics and Computational Intelligence*, 55–59(2011).
- [142] B. Kishore, R. V. Arjunan, R. Saha, *et al.*, "Using Haralick Features for the Distance Measure Classification of Digital Mammograms," *International Journal of Computer Applications*, **6**(1), 2014.
- [143] X. Zhang, J. Cui, W. Wang, *et al.*, "A study for texture feature extraction of high-

resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm," *Sensors*, **17**(7), 1474(2017).

- [144] R. M. Haralick, K. Shanmugam and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621(1973).
- [145] S. A. Bouhamed, M. Chakroun, I. K. Kallel, *et al.*, "Haralick feature selection for material rigidity recognition using ultrasound echo," In *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 1–6(2018).
- [146] C. Konig, Disponible sur : <https://www.futura-sciences.com/sante/dossiers/medecine-oeil-vision-dela-vision-667/page/5/>
- [147] A. Boucher and T. Lee, "Comment extraire la sémantique d'une image?," In *Conference Internationale Sciences Electroniques, Technologies de l'Information et des Telecommunications (SETIT'05)*, 295-306(2005).
- [148] R. Hmida, *Vision et reconstruction 3D : application à la robotique mobile*, Robotique. Université Montpellier : Laboratoire d'électronique et de micro-électronique, 2016.
- [149] L. Mennillo, *Reconstruction 3D de l'environnement dynamique d'un véhicule à l'aide d'un système multi-caméras hétérogène en stéréo wide-baseline*. Automatique / Robotique. Université Clermont : Auvergne, Français, 2019.
- [150] H. N. Do, M. Jadalila, J. Choi, *et al.*, "Feature selection for position estimation using an omnidirectional camera," *Image and Vision Computing*, **39**, 1–9(2015).
- [151] T. Goedemé, M. Nuttin, T. Tuytelaars, *et al.*, "Omnidirectional vision based topological navigation," *International Journal of Computer Vision*, **74**(3), 219–236(2007).
- [152] C. L. Shih and Y. Lee, "A Simple Robotic Eye-In-Hand Camera Positioning and Alignment Control Method Based on Parallelogram Features," *Robotics*, **7**(2), 31(2018).
- [153] L. Y. Xu, Z. Q. Cao, P. Zhao, *et al.*, "A New Monocular Vision Measurement Method to Estimate 3D Positions of Objects on Floor," *International Journal of Automation and Computing*, **14**(2), 159–168(2017).
- [154] G. Bresson, T. Féraud, R. Aufrère, *et al.*, "Real-time monocular SLAM with low memory requirements," *IEEE Transactions on Intelligent Transportation Systems*, **16**(4), 1827–1839(2015).
- [155] Y. J. Yin, D. Xu, Z. T. Zhang, *et al.*, "Plane measurement based on monocular vision," *Journal of electronic measurement and instrument*, **27**(4), 347–352(2013).

- 
- [156] J. Zhou and B. Li, "Robust ground plane detection with normalized homography in monocular sequences from a robot platform," *2006 International Conference on Image Processing*, 3017–3020(2006).
- [157] N. A. Othman, M. U. Salur, M. Karakose, *et al.*, "An Embedded Real-Time Object Detection and Measurement of its Size," *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, 1–4(2018).
- [158] L. M. Lorigo, R. Brooks and W. E. Grimsou, "Visually-Guided Obstacle Avoidance in Unstructured Environments," In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications*, **1**, 373–379(1997).
- [159] I. Ulrich and I. Nourbakhsh, "Appearance-Based Obstacle Detection with Monocular Color Vision," In *AAAI/IAAI*, 866–871(2000).
- [160] T. Naito, T. Ito and Y. Kaneda, "The Obstacle Detection Method Using Optical Flow Estimation at the Edge Image," In *2007 IEEE Intelligent Vehicles Symposium*, 817-822(2007).
- [161] J. Lalonde, R. Laganiere and L. Martel, "Single-Siew Obstacle Detection for Smart Backup Camera Systems," In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1–8(2012).
- [162] Z. Chen and S. T. Birchfield, "Visual detection of lintel-occluded doors from a single image," In *Computer Vision and Pattern Recognition Workshops*, 1–8(2008).
- [163] S. Kim, H. Cheong, D. H. Kim, *et al.*, "Context-based object recognition for door detection," In *15th International Conference on Advanced Robotics (ICAR)*, 155-160(2011).
- [164] U. G. Adar and L. Bayindir, "Door detection using camera images obtained from indoor environments," In *23rd Signal Processing and Communications Applications Conference (SIU)*, 2005-2008(2015).
- [165] D. Kim, D. Kim and R. Nevatia, "A method for recognition and localization of generic objects for indoor navigation", In *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*, 280–288(1994).
- [166] R. M. Noz-Salinas, E. Aguirre, M. García-Silvente, *et al.*, "Door-detection using computer vision and fuzzy logic," *WSEAS Transactions on Systems*, **10**(3), 2004.
- [167] R. Sekkal, *Techniques visuelles pour la détection et le suivi d'objets 2D*. Traitement du signal et de l'image. INSA de Rennes, NNT : 2014ISAR0032. tel-00981107v2, 2014.

- 
- [168] M. Seki and Y. Sugaya, "Floor-wall boundary detection from projected ellipses for autonomous robot navigation," In *International Conference On Advanced Informatics : Concepts, Theory And Application (ICAICTA)*, 1–4(2016).
- [169] R. Gallea, E. Ardizzone and R. Pirrone, "Physical Metaphor for Streaming Media Retargeting," *IEEE transactions on multimedia*, **16**(4), 971–979(2014).
- [170] G. Li, Y. Xie, L. Lin, *et al.*, "Instance-Level salient object segmentation," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2386–2395(2017).
- [171] Y. F. Ma and H. J. Zhang, "Contrast-based image attention analysis by using fuzzy growing," In *Proceedings of the eleventh ACM international conference on Multimedia*, 374–381(2003).
- [172] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, **20**(11), 1254–1259(1998).
- [173] C. Koch and S. Ullman, "Shifts in selective visual attention : towards the underlying neural circuitry," In *Matters of Intelligence*, 115–141(1987).
- [174] J. M. Wolfe, K. R. Cave and S. L. Franzel, "Guided search : an alternative to the feature integration model for visual search," *Journal of Experimental Psychology : Human perception and performance*, **15**(3), 419–433(1989).
- [175] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive psychology*, **12**(1), 97–136(1980).
- [176] J. Harel, C. Koch and P. Perona, "Graph-based visual saliency," In *International Conference on Neural Information Processing Systems*, 545–552(2006).
- [177] R. Achanta, S. Hemami, F. Estrada, *et al.*, "Frequency-tuned salient region detection," In *2009 IEEE conference on computer vision and pattern recognition*, 1597–1604(2009).
- [178] E. Erdem and A. Erdem, "Visual saliency estimation by nonlinearly integrating features using region covariance," *Journal of vision*, **13**(4), 1–20(2013).
- [179] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Vision research*, **49**(10), 1295–1306(2009).
- [180] S. Lu, C. Tan and J. H. Lim, "Robust and efficient saliency modeling from image co-occurrence histograms," *IEEE transactions on pattern analysis and machine intelligence*, **36**(1), 195–201(2013).

- 
- [181] B. Schauerte and R. Stiefelhagen, "Quaternion-based spectral saliency detection for eye fixation prediction," In *European Conference on Computer Vision*, Springer, Berlin, Heidelberg, 116–129(2012).
- [182] J. Li, M. D. Levine, X. An, *et al.*, "Visual saliency based on scale-space analysis in the frequency domain," *IEEE transactions on pattern analysis and machine intelligence*, **35**(4), 996–1010(2012).
- [183] H. Jiang, J. Wang, Z. Yuan, *et al.*, "Salient object detection : A discriminative regional feature integration approach," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2083-2090(2013).
- [184] F. Perazzi, P. Krahenbuhl, Y. Pritch, *et al.*, "Saliency filters : Contrast based filtering for salient region detection," In *2012 IEEE conference on computer vision and pattern recognition*, 733-740(2012).
- [185] C. Yang, L. Zhang, H. Lu, *et al.*, "Saliency detection via graph-based manifold ranking," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3166-3173(2013).
- [186] H. Yoon, B. H. Kim, M. Mukhriddin, *et al.*, "Salient Region Extraction based on Global Contrast Enhancement and Saliency Cut for Image Information Recognition of the Visually Impaired," *THIS*, **12**(5), 2287–2312(2018).
- [187] J. Lou, M. Ren and H. Wang, "Regional principal color based saliency detection," *PloS one*, **9**(11), 1–13(2014).
- [188] M. M. Cheng, N. J. Mitra, X. Huang, *et al.*, "Global contrast based salient region detection," *IEEE transactions on pattern analysis and machine intelligence*, **37**(3), 569–582(2014).
- [189] Q. Jiang, F. Shao and G. Jiang, "Efficient saliency detection using color contrast and similarity distribution information," In *Proc. of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 9–12(2014).
- [190] Q. Zhou, J. Chen, S. Ren, *et al.*, "On Contrast Combinations for Visual Saliency Detection," In *2013 IEEE International Conference on Image Processing*, 2665-2669(2013).
- [191] J. Zhou, J. Zhai and Y. Ren, "Salient Object Detection via Adaptive Region Merging," *KSII Transactions on Internet & Information Systems*, **10**(9),4386–4404(2016).
- [192] X. Qian, J. Han, G. Cheng, *et al.*, "Optimal contrast based saliency detection," *Pattern Recognition Letters*, **34**(11), 1270–1278(2013).

- 
- [193] W. Zhu, S. Liang, Y. Wei, *et al.*, "Saliency optimization from robust background detection," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2814–2821(2014).
- [194] Y. Wei, F. Wen, W. Zhu, *et al.*, "Geodesic saliency using background priors," In *European Conference on Computer Vision*, 29–42(2012).
- [195] Y. Li, X. Hou, C. Koch, *et al.*, "The secrets of salient object segmentation," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 280–287(2014).
- [196] S. Lu, V. Mahadevan and N. Vasconcelos, "Learning optimal seeds for diffusion-based salient object detection," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2790-2797(2014).
- [197] R. Zhao, W. Ouyang, H. Li, *et al.*, "Saliency detection by multi-context deep learning," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1265-1274(2015).
- [198] R. Achanta, A. Shaji, K. Smith, *et al.*, "Slic superpixels compared to state-of-the art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, **34**(11), 2274–2282(2012).
- [199] L. Mai, Y. Niu and F. Liu, "Saliency aggregation : A datadriven approach," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1131-1138(2013).
- [200] P. S. Liao, T. S. Chen and P. C. Chung, "A Fast Algorithm for Multilevel Thresholding," *J. Inf. Sci. Eng.*, **17**(5), 713–727(2001).
- [201] J. Serra, "Biomedical image analysis by mathematical morphology (author's transl)," *Pathologie-biologie*, **27**(4), 205–207(1979).
- [202] E. Rahtu, J. Kannala, M. Salo, *et al.*, "Segmenting salient objects from images and videos," In *European Conference on Computer Vision*, 366-379(2010).
- [203] N. Murray, M. Vanrell, X. Otazu, *et al.*, "Saliency estimation using a non-parametric low-level vision model," In *CVPR*, 433-440(2011).
- [204] X. Hou, J. Harel and C. Koch, "Image signature : Highlighting sparse salient regions," *IEEE transactions on pattern analysis and machine intelligence*, **34**(1), 194–201(2011).
- [205] [https://www.who.int/health-topics/blindness-and-vision-loss#tab=tab\\_1](https://www.who.int/health-topics/blindness-and-vision-loss#tab=tab_1)
- [206] <http://www.netwalbert.com/vuedensemble/definition.htm>

- 
- [207] E. Foulke, "The perceptual basis for mobility," *DOCUMENT RESUME ED*, **84**(738), 93(1970).
- [208] V. Bhoge and S. Y. Chinchulikar, "Review on GPS Navigation System for blind people," *IJESRT*, **5**(6), 570–574(2016).
- [209] S. Vatansever and I. Butun, "A broad overview of GPS fundamentals : Now and future," In *IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, 1-6(2017).
- [210] B. Kaur and J. Bhattacharya, "A scene perception system for visually impaired based on object detection and classification using multi-modal DCNN," *arXiv preprint arXiv :1805.08798*, 2018.
- [211] Q. K. Dang, Y. Chee, D. D. Pham, *et al.*, "A virtual blind cane using a line laser-based vision system and an inertial measurement unit," *Sensors*, **16**(1), 95(2016).
- [212] A. S. Al-Fahoum, H. B. Al-Hmoud and A. A. Al-Fraihat, "A Smart Infrared Microcontroller-Based Blind Guidance System," *Active and Passive Electronic Components*, 2013.
- [213] D. Jain, "Path-guided indoor navigation for the visually impaired using minimal building retrofitting," In *Proceedings of the 16th international ACM SIGACCESS conference on Computers accessibility*, 225–232(2014).
- [214] M. Medhioub, I. Khanfir Kallel, S. A. Bouhamed, *et al.*, "Electronic Embedded System for Stair Recognition Based on Possibilistic Modeling of Ultrasonic Signal", *IEEE SENSORS JOURNAL*, **21**(5), pp. 5787-5797 (2021).
- [215] R. Agarwal , N. Ladha, M. Agarwal, *et al.*, "Low cost ultrasonic smart glasses for blind, In Information Technology," In *2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 210–213(2017).
- [216] A. Pereira, N. Nunes, D. Vieira, *et al.*, "Blind Guide : An Ultrasound Sensor-based Body Area Network for Guiding Blind People," *Procedia Computer Science*, **67**, 403–408(2015).
- [217] J. Sohl-Dickstein, S. Teng, B. M. Gaub, *et al.*, "A device for human ultrasonic echolocation," *IEEE Transactions on Biomedical Engineering*, **62**(6), 1526–1534(2015).
- [218] V. V. Nahar, J. L. Nikam and P. K. Deore, "Smart blind walking stick," *International Journal of Modern Trends in Engineering and Research*, **2349**, 645–652(2016).
- [219] A. V. Nandini, A. Dwivedi, N. A. Kumar, *et al.*, "Smart Cane for Assisting Visually

- Impaired People," In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, 546-551(2019).
- [220] P. Grant, L. Spencer, A. Arnoldussen, *et al.*, "The Functional Performance of the Brain-Port V100 Device in Persons Who Are Profoundly Blind," *Journal of Visual Impairment & Blindness*, **110**(2), 77–88(2016).
- [221] M. Gori, G. Cappagli, A. Tonelli, *et al.*, "Devices for visually impaired people : High technological devices with low user acceptance and no adaptability for children", *Neuroscience & Biobehavioral Reviews*, **69**, 79–88(2016).
- [222] N. Mahmud, R. K. Saha, R. B. Zafar, *et al.*, "Vibration and Voice Operated Navigation System for Visually Impaired Person," In *2014 international conference on informatics, electronics vision (ICIEV)*, 1–5(2014).
- [223] S. Shoval, J. Borenstein and Y. Koren, "The navbelt a computerized travel aid for the blind based on mobile robotics technology," *IEEE Transactions on Biomedical Engineering*, **45**(12), 1376-1386(2014).
- [224] B. Siddhartha, A. P. Chavan and B. V. Uma, "An Electronic Smart Jacket for the Navigation of Visually Impaired Society," *Materials Today : Proceedings*, **5**(4), 10665–10669(2018).
- [225] A. Nada, S. Mashelly, M. A. Fakhr, *et al.*, "Effective Fast Response Smart Stick for Blind People," In *Proceedings of the second nternational Conference on Advances in bioinformatics and environmental engineering-ICABEE*, 2015.
- [226] M. Kumar, F. Kabir and S. Roy, "Low Cost Smart Stick for Blind and Partially Sighted People," *International Journal of Advanced Engineering and Management*, **2**(3), 65–68(2017).
- [227] R. Want, A. Hopper, V. Falcao, *et al.*, "The Active Badge Location System," *ACM Transactions on Information Systems (TOIS)*, **10**(1), 91–102(1992).
- [228] M. I. Islam, M. M. H.Raj, S. Nath, *et al.*, "An Indoor Navigation System for Visually Impaired People Using a Path Finding Algorithm and a Wearable Cap," In *2018 3rd International Conference for Convergence in Technology (I2CT)*, 1–6(2018).
- [229] Y. Lin, K. Wang, W. Yi, *et al.*, "Deep learning based wearable assistive system for visually impaired people," In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.
- [230] J. Bai, S. Lian, Z. Liu, *et al.*, "Smart guiding glasses for visually impaired people in indoor environment," *IEEE Transactions on Consumer Electronics*, **63**(3), 258–266(2017).

- 
- [231] A. Nishajith, J. Nivedha, S. S. Nair, *et al.*, "Smart Cap - Wearable Visual Guidance System for Blind," In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, 275-278(2018).
- [232] K. Chaccour and G. Badr, "Computer vision guidance system for indoor navigation of visually impaired people," In *2016 IEEE 8th International Conference on Intelligent Systems (IS)*, 449-454(2016).
- [233] K. Niveditha, P. D. Kavya, P. Nivedha, *et al.*, "Virtual Eye for Blind using IOT," *International Journal of Engineering Research & Technology*, **8**, 116–121(2020).
- [234] N. A. Kumar, Y. H. Thangal and K. S. Beevi, "IoT Enabled Navigation System for Blind," In *IEEE R10 Humanitarian Technology Conference (R10-HTC)*, 186-189(2019).
- [235] D. P. Khairnar, R. B. Karad and A. Kapse, "PARTHA : A Visually Impaired Assistance System," In *3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, 32-37(2020).
- [236] H. Zhang and C. Ye, "An indoor navigation aid for the visually impaired," In *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 467–472(2016).
- [237] C. Ye, S. Hong, X. Qian, *et al.*, "Co-Robotic Cane : A New Robotic Navigation Aid for the Visually Impaired," *IEEE systems, man, and cybernetics magazine*, **2**(2), 33–42(2016).
- [238] D. Jeon, N. Ickes, P. Raina, *et al.*, "A. 24 1 A 0.6V 8mW 3D vision processor for a navigation device for the visually impaired," In *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 416–417(2016).
- [239] K. A. Kaczmarek, "The tongue display unit (tdu) for electrotactile spatiotemporal pattern presentation," *Scientia Iranica*, **18**(6), 1476–1485(2011).
- [240] Y. Matsuda, I. Sakuma, Y. Jimbo, *et al.*, "Finger braille recognition system for people who communicate with deaf-blind people," In *2008 IEEE International Conference on Mechatronics and Automation*, 268–273(2008).
- [241] L. A. Jones, B. Lockyer and E. Piatieski, "Tactile display and vibrotactile pattern recognition on the torso," *Advanced Robotics*, **20**(12), 1359–1374(2006).
- [242] J. Sohl-Dickstein, S. Teng, B. M. Gaub, *et al.*, "A device for human ultrasonic echolocation," *IEEE Transactions on Biomedical Engineering*, **62**(6), 1526–1534(2015).
- [243] L. Kay, "Auditory perception of objects by blind persons, using a bioacoustics high resolution air sonar," *The Journal of the Acoustical Society of America*, **107**(6), 3266–3275(2000).

- 
- [244] M. Auvray, S. Hanneton and J. K. O'Regan, "Learning to perceive with a visuoauditory substitution system : localisation and object recognition with the voice," *Perception*, **36**(3), 416–430(2007).
- [245] S. Abboud, S. Hanassy, S. Levy-Tzedek, *et al.*, "Eye-music : Introducing a visual colorful experience for the blind using auditory sensory substitution," *Restorative neurology and neuroscience*, **32**(2), 247–257(2014).
- [246] J. Villanueva and R. Farcy, "Optical device indicating a safe free path to blind people," *IEEE transactions on instrumentation and measurement*, **61**(1), 170–177(2011).
- [247] M. C. Kang, S. H. Chae, J. Y. Sun, *et al.*, "A novel obstacle detection method based on deformable grid for the visually impaired," *IEEE Transactions on Consumer Electronics*, **61**(3), 376–383(2015).
- [248] V. N. Hoang, T. H. Nguyen, T. L. Le, *et al.*, "Obstacle detection and warning for visually impaired people based on electrode matrix and mobile kinect," In *2015 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS)*, 54–59(2015).
- [249] S. Wu, J. Wieland, O. Farivar, *et al.*, "Automatic alt-text : Computer generated image descriptions for blind users on a social network service," In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1180–1192(2017).
- [250] S. Jaiswal, J. Warriar, V. Sinha, *et al.*, "Small sized vision based system for blinds," *Int. J. Eng. Sci*, **15968**, (2018).
- [251] R. K. Katzschmann, B. Araki and D. Rus, "Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **26**(3), 583–593(2018).
- [252] <https://www.arduino.cc/>
- [253] <https://www.raspberrypi.org/documentation/hardware/camera/>

**Titre :** Perception de scène par un système multi-capteurs application à la navigation dans des environnements intérieurs structurés

**Mots clés :** Interprétation de scène; modélisation de scène; théorie de possibilité; localisation et détection des objets; reconnaissance de la rigidité matérielle des objets; classification; recalage et fusion des données hétérogènes

**Résumé :** Cette thèse s'inscrit dans le cadre de la navigation perceptuelle. Notre objectif est d'étudier et de concevoir un système d'interprétation d'une scène d'environnement intérieur, observée par un système multi-capteurs réunissant un capteur ultrason et une caméra RVB. Le système proposé peut être employé pour équiper un dispositif intelligent d'assistance aux non voyants, ou encore un robot opérant dans des espaces meublés. Dans un système d'interprétation de scène, les acquisitions faites par les capteurs, présentent des restrictions du monde réel et se trouvent affectées d'imperfections, qu'il convient de prendre en compte au lieu de les ignorer. Leur prise en compte dans notre système d'interprétation a été effectuée par l'emploi de la théorie des possibilités lors de la modélisation des données acquises. Les modèles adoptés sont des distributions de possibilités. L'analyse et l'interprétation de la scène acquise s'est en suite basée sur ces connaissances possibilistes. Le système d'aide à la navigation proposé dans ce travail, présente une description de la scène environnante selon un modèle simpliste, partageant le champ intercepté par

les capteurs en trois zones majeures, à savoir : face, gauche et droite. Il fournit à l'utilisateur des informations concernant la distance qui le sépare des objets détectés, la rigidité matérielle de ces objets, ainsi que leur positionnement dans la scène (objet à gauche, objet en face, objet à droite). Les performances du système d'interprétation proposé sont évaluées en utilisant le prototype "NA\_System", développé par l'équipe "Cybernetics team" du laboratoire "CEM\_Lab" de l'École nationale d'ingénieurs de Sfax (ENIS). Les résultats obtenus sont encourageants et montrent l'efficacité de la théorie des possibilités comme cadre de représentation de données acquises de différents capteurs. La stratégie d'interprétation de scène proposée s'est montrée efficace pour intégrer les informations issues de multiples sources de connaissances. Dans la chaîne de traitement de données adoptées pour l'analyse et l'interprétation de la scène, de nouvelles approches ont été proposées, notamment pour la sélection d'attributs, la détection d'objets saillants, la classification, la fusion et le recalage de données issues de deux sources.

**Title:** Scene perception by a multi-sensor system application to navigation in structured interior environments.

**Keywords:** Scene interpretation; Scene modeling; Theory of possibility; Object detection and localization; Recognition of material rigidity of objects; Classification; Registration and fusion of heterogeneous data.

**Abstract:** This thesis falls within the framework of perceptual navigation. Our goal is to study and design an interpretation system for indoor environment scenes, observed by a multi-sensor system that combines ultrasonic sensors and RGB camera. The proposed system can be used to equip smart device for assisting blind people, or for robot operating in furnished spaces. In a scene interpretation system, the acquisitions made by the sensors present real world restrictions and are affected by imperfections, which should be taken into account rather than ignored. Their inclusion in our interpretation system was carried out by using the possibility theory when modeling the acquired data. The adopted models are possibility distributions. The analysis and interpretation of the acquired scene was then based on this possibilistic knowledge. The navigation aid system proposed in this work describes the surrounding scene according to a simplistic model, and divides the intercepted field by the sensors into three main areas, namely: front, left and right.

It provides the user with information about the distance between him and the detected objects, the material rigidity of these objects, as well as their position in the scene (object on the left, object in front, object on the right). The performance of the proposed interpretation system is evaluated using the "NA\_System" prototype, developed by the "Cybernetics team" of the "CEM\_Lab" laboratory of the National Engineering School of Sfax (ENIS). The obtained results are encouraging and show the effectiveness of the possibility theory as a framework for representing data acquired from different sensors. The proposed scene interpretation strategy has been proved to be effective in integrating information from multiple knowledge sources. In the data processing chain for scene analysis and interpretation, new approaches are proposed, especially in attribute selection, salient object detection, classification, fusion and registration data from two sources.